



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE CIENCIAS

TESINA DE GRADUACIÓN PARA OSTENTAR EL TÍTULO DE
LICENCIADO EN BIOQUÍMICA

ANÁLISIS DE LOS MICROSATÉLITES AUTOSÓMICOS DE UNA MUESTRA DE LA POBLACIÓN URUGUAYA CON FINES FORENSES.

AUTOR: LORENA SILVA ORGEN

**TUTOR: Prof. Agdo. PhD. BERNARDO
BERTONI.**

*Laboratorio de Epidemiología Genética- Departamento de Genética
Facultad de Medicina - UDELAR.*

*Laboratorio de Registro Nacional de Huellas Genéticas
Dirección Nacional de Policía Científica.*

URUGUAY – MONTEVIDEO

2022

A mi madre, padre y hermano

AGRADECIMIENTOS

A través de estas líneas quiero expresar mi más sincero agradecimiento a todas las personas que con su soporte científico y humano han colaborado en la realización de este trabajo.

Agradezco muy especialmente a mi tutor, Prof. Agdo. PhD. Bernardo Bertoni, por darme la oportunidad de realizar mi tesina de graduación e integrarme a su grupo de trabajo en el Laboratorio de Epidemiología de la Facultad de Medicina de la Universidad de la República. Sobre todo, agradecerle por la orientación recibida, su paciencia y disponibilidad en todos los momentos en que lo necesite, lo cual me permitió un buen aprovechamiento del trabajo realizado.

Agradecer al Ministerio del Interior y a la Dirección Nacional de Policía Científica, por la posibilidad brindada al permitirme realizar esta tesina en conjunto con la Facultad de Ciencias y agradecer especialmente la confianza otorgada al haberseme concedido el honor de poder contribuir a dicha Dirección con una base de datos de referencia de la población uruguaya, objetivo del presente trabajo.

A la encargada del Laboratorio de Registro Nacional de Huellas Genéticas de la Dirección Nacional de Policía Científica, MSc. Lic. Ciencias Biológicas Natalia Sandberg, por la guía recibida, su apoyo y por compartir sus conocimientos y experiencias en el área forense.

A todos los integrantes del Laboratorio de Epidemiología Genética, Departamento de Genética de la Facultad de Medicina y del Laboratorio Registro Nacional de Huellas Genéticas, gracias por todo el apoyo recibido.

Quiero agradecer a toda mi familia, pero particularmente a mis padres y hermano por el apoyo incondicional que siempre me han brindado en todos los proyectos y demás metas que me he trazado. A pesar de la distancia, siempre los he sentido aquí a mi lado compartiendo mis alegrías y angustias. Gracias por enseñarme que los sueños se logran a base de esfuerzo, perseverancia y dedicación.

ÍNDICE

LISTA DE TABLAS	V
LISTA DE FIGURAS	VI
LISTA DE ABREVIATURAS	VIII
RESUMEN	IX
1. INTRODUCCIÓN	1
1.1 Organización del ADN.....	4
1.2 El ADN en la Identificación Humana.....	4
1.3 Polimorfismos del ADN en la identificación humana.....	6
1.3.1 Los Microsatélites (STRs).....	7
1.4 Nomenclatura de los STRs.....	7
1.5 Antecedentes de estudios de variabilidad y caracterización genética en la población uruguaya.....	9
1.6 Antecedentes de estudios sociodemográficos y genéticos de la población uruguaya.....	10
1.6.1 Contribución ancestral en la población uruguaya.....	11
1.7 La investigación en el ámbito forense; actualidad y propósitos.....	14
2. OBJETIVOS	15
2.1 Objetivo general.....	15
2.2 Objetivos específicos.....	15
3. MATERIALES Y MÉTODOS	17
3.1 Obtención del perfil genético.....	18
3.2 Análisis estadístico.....	20

3.2.1 Cálculo de las frecuencias alélicas.....	21
3.2.2 Equilibrio de Hardy-Weinberg.....	21
3.3 Parámetros poblacionales.....	21
3.3.1 Heterocigosidad.....	21
3.3.2 Poder de Exclusión (PE).....	22
3.3.3 Probabilidad de Coincidencia o <i>match</i> (PM).....	23
3.3.4 Poder de Discriminación (PD).....	23
3.3.5 Índice de paternidad (IP).....	24
3.3.6 Contenido de Información Polimórfico (PIC).....	25
3.4 Comparación de la frecuencia alélica entre la población uruguaya y la población disponible en CODIS.....	26
3.5 Diferenciación genética.....	27
3.5.1. Índice de fijación de Wright (F_{ST}).....	29
3.5.2 Coeficiente de diferenciación genética de Nei (G_{ST}).....	30
3.6 Métodos empleados para el análisis de estructuración de la población.....	31
3.6.1 STRUCTURE.....	31
3.6.2 Análisis de Componentes Principales (PCA).....	33
3.7 Análisis de mestizaje de la población uruguaya.	34
3.7.1 Poblaciones empleadas para el análisis del mestizaje de la Población uruguaya.	34
3.7.2 Métodos utilizados para el Análisis de Mestizaje.....	35
3.7.2.1 ADMIX 95.....	35
3.7.2.2 VARMIX 3.0.....	36

4. RESULTADOS	37
4.1 Frecuencias alélicas y parámetros estadísticos en la muestra de la población de Uruguay.....	37
4.2 Comparación de las frecuencias alélicas entre la población uruguaya y los datos poblaciones disponibles en CODIS.....	42
4.2.1 Comparación para el marcador Penta E.....	43
4.2.2 Comparación para el marcador D18S51.....	44
4.2.3 Comparación para el marcador FGA.....	45
4.2.4 Comparación para el marcador TPOX.....	46
4.4 Diferenciación genética.....	48
4.4.1 Índice de fijación de Wright (F_{ST}).....	48
4.4.2 Coeficiente de diferenciación genética de Nei (G_{ST}).....	49
4.5 Análisis de la estructura genética en la muestra poblacional uruguaya empleando el Software STRUCTURE v 2.3.4.....	50
4.6 Análisis de Componentes Principales (PCA).....	57
4.7 Resultados obtenidos con el método de identidad génica (ADMIX 95).....	58
4.8 Resultados obtenidos con el programa VARMIX 3.0.....	59
5. DISCUSIÓN	60
6. CONCLUSIONES	80
7. REFERENCIAS BIBLIOGRÁFICAS	82
8. ANEXOS	101
Anexo 1: Protocolo de Amplificación directa de ADN presente en tarjetas de almacenamiento.....	101
Anexo 2: Electroferograma de la escalera alélica PowerPlex® Fusion.....	103
Anexo 3: Perfil genético de un individuo.....	104

Anexo 4: Frecuencias alélicas de la población caucásica disponible en CODIS.....	105
Anexo 5: Frecuencias alélicas para 15 marcadores STRs obtenidas a partir de una muestra población de 200 individuos uruguayos.....	107

LISTA DE TABLAS

Tabla 1a:	Frecuencias alélicas observadas en los 23 loci STRs analizados para los alelos comprendidos entre el 2.2 y 18.3.....	39
Tabla 1b:	Frecuencias alélicas observadas en los 23 loci STRs analizados para los alelos comprendidos entre el 19 y 46.2.....	40
Tabla 2:	Parámetros estadísticos forenses de los 23 loci STRs analizados.....	41
Tabla 3:	Análisis de varianza (Test de Bartlett).....	47
Tabla 4:	Índice de fijación de Wright (F_{ST}) entre poblaciones.....	48
Tabla 5:	Diferenciación genética de Nei (G_{ST}) entre poblaciones.....	49
Tabla 6:	Mezcla de amplificación de la PCR para amplificación directa de ADN de las tarjetas perforadas de almacenamiento.....	102
Tabla 7:	Protocolo del ciclo térmico.....	102
Tabla 8a:	Frecuencias alélicas obtenidas por el FBI en una muestra poblacional caucásica de 361 individuos estadounidenses disponibles en CODIS para los alelos comprendidos entre el 2.2 y 18.3.....	105
Tabla 8b:	Frecuencias alélicas obtenidas por el FBI en una muestra poblacional caucásica de 361 individuos estadounidenses disponibles en CODIS para los alelos comprendidos entre el 19 y 36.....	106
Tabla 9:	Frecuencias alélicas obtenidas para 15 marcadores STR en una muestra poblacional uruguaya de 200 individuos.....	107

LISTA DE FIGURAS

Figura 1:	Nomenclatura del marcador D18S51.....	7
Figura 2:	Configuración del sistema PowerPlex Fusion.	18
Figura 3:	Electroferograma del size standard WEN ILS 500 IDX y ejemplo de asignación a un perfil genético.	19
Figura 4:	Distribución de las frecuencias alélicas para el marcador Penta E.....	43
Figura 5:	Distribución de las frecuencias alélicas para el marcador D18S51.....	44
Figura 6:	Distribución de las frecuencias alélicas para el marcador FGA.	45
Figura 7:	Distribución de las frecuencias alélicas para el marcador TPOX.....	46
Figura 8:	Árbol de distancias genéticas obtenido a partir del Índice de fijación de Wright (F_{ST}) utilizando el método de Neighbor-Joining.	48
Figura 9:	Árbol de distancias genéticas obtenido a partir del Coeficiente de diferenciación genética de Nei (G_{ST}) utilizando el método de Neighbor-Joining.....	49
Figura 10:	Valores de Delta K obtenidos a partir de Harvester Structure para la población uruguaya.....	50
Figura 11:	Gráfica que representa una simulación con $K=1$, $K=2$ y $K=3$	51
Figura 12:	Valores de Delta K obtenidos a partir de Harvester Structure para la población uruguaya y las parentales.....	52
Figura 13:	Gráfico que representa una simulación con $K=2$	53
Figura 14:	Gráfica que representa una simulación con $K=3$	54
Figura 15:	Gráfico de triangulo que representa un análisis en el Software STRUCTURE de 3 poblaciones parentales.....	55

Figura 16:	Resultados obtenidos del análisis del Software STRUCTURE de 4 poblaciones para K=2; K=3; K=4; K=5.....	56
Figura 17:	Ilustración de los 5 clústers en los que el PCA agrupó a los individuos de la población uruguaya y aquellos individuos de las poblaciones empleadas por CODIS.....	57
Figura 18:	Representación gráfica de los resultados de ADMIX 95.....	58
Figura 19:	Representación gráfica de los resultados de VARMIX 3.0.....	59
Figura 20:	Residencias Definitivas otorgadas entre el año 2013 y 2019.....	67
Figura 21:	Residencias Definitivas por Nacionalidad otorgadas entre el año 2013 y 2019.....	68
Figura 22:	Proporción de ascendencia afroamericana, nativo americana y europea estimada para 6 poblaciones latinoamericanas.....	72
Figura 23:	Representación gráfica de la contribución ancestral observada para la población uruguaya a partir del análisis de distintos marcadores.....	77
Figura 24:	Electroferograma de la escalera alélica PowerPlex® Fusion.....	103
Figura 25:	Perfil genético de un individuo.....	104

LISTA DE ABREVIATURAS

ADN - Ácido desoxirribonucleico.

CODIS - Combined DNA Index System.

EHW - Equilibrio de Hardy-Weinberg.

ESS - European Standard Set

F_{ST} - Índice de fijación de Wright.

G_{ST} - Coeficiente de diferenciación genética de Nei.

He – Heterocigosidad esperada.

Ho – Heterocigosidad observada.

ISFG - International Society of Forensic Genetics.

LRNHG - Laboratorio de Registro Nacional de Huellas Genéticas.

N – número de individuos.

PCR - Reacción en cadena de la polimerasa.

PD – Poder de discriminación.

PE - Poder de exclusión.

PI - Índice de Paternidad.

PIC - Contenido de Información Polimórfica.

PM - Probabilidad de Coincidencia o *match*.

STR - Repeticiones Cortas en Tándem.

RESUMEN

La Genética Forense proporciona sistemas de análisis de ADN a fin de resolver casos de criminalística forense, estudios de paternidad y casos judiciales, para lo cual se requiere que cada población obtenga su propia base de datos genéticos. En el presente estudio, se analizaron 4000 individuos correspondientes a una muestra de la población uruguaya. Se amplificaron 23 microsatélites autosómicos y la Amelogenina, mediante el sistema PowerPlex® Fusion 6C System (Promega Corporation). A través del análisis de los microsatélites se encontró que luego de aplicar la corrección de Bonferroni a los marcadores D8S1179, D12S391 y D13S317, todos los marcadores estudiados estuvieron en equilibrio de Hardy-Weinberg a excepción del D8S1179. Con los datos obtenidos, se realizó un análisis de los parámetros estadísticos forenses, como; el poder de discriminación (PD), poder de exclusión (PE), contenido de información polimórfica (PIC), heterocigosidad observada (Ho) y heterocigosidad esperada (He). El marcador SEE33 resultó ser el más polimórfico con un índice de 93,91%, un índice Ho de 94,55% y He de 94,58%, presentando a su vez un PD y PE total mayor a 99,999999%, mientras que el menor valor de PIC lo registró el marcador TPOX con el 62%. Se analizó la distribución de la variación genética entre poblaciones empleando el Índice de Fijación de Wright (F_{ST}) y el Coeficiente de diferenciación genética de Nei (G_{ST}), a partir de las frecuencias alélicas obtenidas y datos poblacionales de la región. Se efectuó un análisis de estructuración poblacional utilizando el software STRUCTURE y un Análisis de Componentes Principales (PCA), empleando para ello los datos de las poblaciones parentales disponibles en CODIS, desarrollado por el FBI. A su vez, se observó el grado de mestizaje poblacional desde

dos enfoques diferentes, uno basado en las frecuencias génicas (ADMIX 95) y otro en la información genética de los microsatélites (VARMIX 3.0).

Los resultados obtenidos mostraron que la población uruguaya presenta escasa diferenciación genética con las poblaciones regionales (Argentina, Bolivia, Brasil, Chile y Paraguay), no se observó subestructuración poblacional y presenta gran similitud con la población caucásica e hispana disponible en el sistema de referencia utilizado en genética forense. Los resultados obtenidos en el análisis de mestizaje, concuerdan con estudios previos realizados para otros marcadores moleculares, observándose, en promedio para ambos métodos, una contribución genética parental mayoritaria proveniente de europeos (65%), con un aporte menor de africanos (21,6%) y nativo americanos (13%).

Este trabajo constituye un aporte significativo para la identificación humana y la genética forense en Uruguay, brindando una base de datos de referencia imprescindible en la realización de cálculos estadísticos en casos de criminalística forense y filiación, logrando así la obtención de resultados más confiables.

Palabras claves: ADN, marcadores autosómicos, criminalística forense, población uruguaya, diferenciación genética, estructuración de la población, mestizaje.

1. INTRODUCCIÓN

El empleo de métodos fiables en la identificación humana ha sido siempre una necesidad para la sociedad, tanto en resolver cuestiones del ámbito civil como criminal. A lo largo de los años, diferentes técnicas han sido utilizadas a fin de caracterizar e individualizar a las personas, entre ellas; el análisis de huellas dactilares, los análisis antropológicos, los análisis de los grupos sanguíneos y también los análisis moleculares de los polimorfismos del sistema HLA (antígeno leucocitario humano). Con el descubrimiento de la estructura del ADN en 1953 y de sus regiones polimórficas, la identificación humana, se convirtió en una herramienta indispensable en la genética forense.

El uso del análisis de ADN, está fundamentado en la individualidad y exclusividad del perfil genético de un individuo. La excepción a esta regla es la presencia de gemelos monocigóticos o univitelinos que genéticamente son idénticos (Quintero et al., 2020). Se denomina perfil genético de un individuo, a la combinación genotípica de los marcadores que son heredados de sus progenitores, estos corresponden a secuencias de ADN que varían entre individuos, denominadas regiones hipervariables o polimórficas, y se encuentran distribuidas en cromosomas autosómicos y sexuales (Butler, 2005). Estas regiones del ADN tienen un alto poder de discriminación, lo cual permite, con gran fiabilidad, excluir a un individuo de ser falsamente acusado en un crimen o, en casos de paternidad; permitir su inclusión o exclusión (Lorente, 2004).

Los primeros polimorfismos de longitud encontrados en el genoma humano, fueron descritos por Jeffreys y cols. (1985a), los minisatélites; que junto con los microsatélites o STRs (*Short Tandem Repeats* - repetidos cortos en tándem), son polimorfismos de longitud denominados VNTRs (Repeticiones en Tándem de Número Variable), los cuales se caracterizan por presentar una organización semejante: segmentos de ADN similares se encuentran dispuestos en tándem (uno detrás de otro) en número variable de veces (Jeffreys et al., 1985a). Los STRs consisten en elementos repetitivos de 2 a 6 nucleótidos, son numerosos en el genoma humano, tienen una longitud muy polimórfica y también pueden variar en las secuencias de los elementos repetitivos, haciendo que estos marcadores sean muy efectivos en la identificación de individuos. El tamaño de los STRs, en comparación con los minisatélites empleados por Jeffreys y cols. (1985a), hace que los marcadores STRs sean mejores candidatos para su uso en Genética Forense.

El surgimiento de la técnica de Reacción en Cadena de la Polimerasa (PCR), ha permitido generar perfiles genéticos a partir de una muestra de referencia o de una evidencia de ADN en los Laboratorios Forenses, utilizando reacciones multiplex que permiten amplificar muchos STRs en conjunto, haciendo que el procedimiento sea más rápido, fácil y de menor costo (Mullis et al., 1986). El número de marcadores que pueden amplificarse simultáneamente ha aumentado hasta 27 STRs valiéndose del empleo de fluorocromos de diferentes colores (Promega Corporation, 2018).

Los avances en la tecnología del ADN, facilitaron la creación de bancos de datos de perfiles de ADN con fines tanto civiles como criminales. Estados Unidos fue pionero en su creación, en 1994 (<https://www.fbi.gov/services/laboratory/biometric->

analysis/codis/codis-and-ndis-fact-sheet), surgió el sistema CODIS (*Combined DNA Index System*), desarrollado por el FBI (Federal Bureau of Investigation). En 1997, fueron seleccionados por su gran poder discriminante y por estar localizados en regiones no codificantes, 13 marcadores genéticos STR (Butler, 2005), estos son: D21S11, D18S51, D16S539, D13S317, D8S1179, D7S820, D5S818, D3S1358, vWA , TH01, CSF1PO, FGA, TPOX, posteriormente, en enero del 2017, se adicionaron 7 loci en las bases, los cuales son los marcadores; D22S1045, D19S433, D1S1656, D12S391, D10S1248, D2S441 y D2S1338; estudiando un total 20 marcadores STRs autosómicos, además de la Amelogenina que tipifica el sexo, y que debe incluirse en las bases de datos, con el fin de reducir la probabilidad de coincidencias aleatorias (Moretti et al., 2016). Estos marcadores presentan alto polimorfismo y tienen un gran poder discriminante en la identificación humana (Lorente et al., 2007).

El uso de los marcadores genéticos STR en el campo forense, ha aumentado con el transcurso de los años, generando la necesidad en los laboratorios del uso de varios de ellos con fines técnicos y de investigación, para obtener un mayor acercamiento a los valores estadísticos poblacionales de las zonas de trabajo (Butler, 2006).

La evaluación estadística de los resultados de un análisis de ADN, ya sean en el establecimiento de vínculos biológicos de parentesco o identificación de evidencias, se realiza a partir de las frecuencias alélicas, empleando algoritmos matemáticos validados. Estas frecuencias alélicas surgen de la base de datos de referencia, que deben ser locales (Jeffreys et al., 1985a). La utilización de estos marcadores con fines de identificación exige como requisito básico que estos sean previamente estudiados y

caracterizados en la población que posteriormente servirá como referencia, para obtener resultados y conclusiones de alta confiabilidad (Hakim et al., 2020; Toscanini et al., 2015).

1.1 Organización del ADN

El Proyecto del Genoma Humano, logro caracterizar la secuencia de todo el genoma. El Genoma Humano haploide se puede clasificar en: ADN codificante y ADN no codificante, que tiene funciones estructurales y reguladoras, siendo esta parte donde se encuentra la mayor parte de la variabilidad genética entre los individuos (Pellegrino & Crespillo 2021). El ADN no codificante, no está sujeto a presión selectiva y soporta mayores niveles de variación entre individuos, por lo que tiene gran interés en la identificación humana debido a su alto polimorfismo (Butler, 2005).

1.2 El ADN en la Identificación Humana

En sus inicios, la identificación humana se basaba en parámetros morfológicos; pero los avances científicos y tecnológicos permitieron el desarrollo de la Genética Forense para la identificación de personas, a través del análisis de vestigios biológicos recogidos de la escena del crimen, restos óseos o cadavéricos e investigaciones de paternidad (Jobling & Gill, 2004).

En el siglo XX, se logró procesar muestras biológicas con el fin de identificar personas, para ello se utilizó marcadores eritrocitarios ABO (conocidos como polimorfismos clásicos), descubiertos por Karl Landsteiner en 1900 (Landsteiner,

1990). Estos marcadores fueron utilizados en casos de investigación biológica de paternidad y en el análisis de vestigios biológicos de interés criminal como manchas de sangre (Jobling & Gill, 2004).

Posteriormente se fueron descubriendo diversos polimorfismos en las proteínas séricas leucocitarias y eritrocíticas que al ser analizadas a la vez, resultaban eficaces para la identificación de personas (Carracedo, 1999), sin embargo, su potencial resultó ser limitado debido a la poca variabilidad para estos marcadores (Fernández et al., 1999).

La descripción de un nuevo marcador biológico en 1954, marcó la segunda fase del desarrollo de los marcadores de identificación humana. Fue demostrada la existencia de un sistema de histocompatibilidad, el sistema antígeno leucocitario humano o HLA (Trujillo et al., 2018). Pero presentaban grandes limitaciones cuando se trataba de analizar muestras degradadas o que presentaban escasa cantidad de ADN, lo que sucede con mucha frecuencia en el trabajo forense (Carracedo, 2013).

Por otro lado, el análisis de los polimorfismos de ADN marcó la tercera fase de la identificación humana en ciencias forense. En 1985a, Jeffreys y cols., describen unos marcadores, los repetidos de ADN en tándem, extremadamente variables y heredables que producían una especie de “huellas genéticas” de ADN o “ADN fingerprinting” (Jeffreys et al., 1985a).

El desarrollo de la PCR (Mullis et al., 1986), permitió utilizar los polimorfismos de ADN, para la identificación humana, teniendo un incremento, en más del 99% de los

rangos de probabilidad para excluir a los investigados de casos civiles y penales (Jeffreys et al., 1985b). A partir del año 1987 esta técnica empezó a tomar protagonismo cuando se utilizó el “ADN fingerprinting” o “Análisis de Huellas Genéticas” en los procesos criminales (Lander et al., 2001). En la actualidad, los estudios genéticos de STRs para la identificación humana son considerados esenciales cuando se tienen restos biológicos y son utilizados en los tribunales de justicia en el mundo por su gran confiabilidad científica (González, 2006).

1.3 Polimorfismos del ADN en la identificación humana

En la identificación humana se puede definir un marcador genético como un polimorfismo que permite diferenciar individuos, así como establecer sus relaciones de parentesco (Rangel, 2010). Los polimorfismos expresan la variabilidad existente en un fragmento de ADN, por lo que se puede inferir que, a mayor número de alelos, hay mayor polimorfismo y mayor poder de identificación de un individuo (Lorente y Lorente, 1995).

A nivel de ADN, un polimorfismo puede ser desde la mutación de una sola base hasta el cambio en el número de unidades repetidas en tándem en ciertas regiones del ADN. Este ADN repetitivo en tándem, está formado por bloques de ADN que se repiten consecutivamente y se suele dividir en: a) ADN satélite, b) ADN minisatélites, c) ADN microsátélites (López & Garrido, 2012).

1.3.1 Los Microsatélites (STRs)

Son por excelencia los polimorfismos empleados en forense actualmente, corresponden a la repetición en tándem de entre 2 y 6 nucleótidos. Los microsatélites presentan dos características que los hacen ideales para ser utilizados: I) están distribuidos de forma casi homogénea por todo el genoma, II) presentan un número elevado de alelos con frecuencias similares entre sí, de forma que la probabilidad de que un individuo sea heterocigoto es muy elevada (Weber y Wong, 1993). Secuencias nucleotídicas repetidas cortas (por ejemplo: ACTG), dan lugar a diversos alelos que se nombran de acuerdo al número de veces que se encuentre la secuencia repetida.

1.4 Nomenclatura de los STRs

Los marcadores STRs que quedan por fuera de la región codificante, pueden designarse por su posición cromosómica, su nomenclatura está representada por códigos compuestos por letras y números, donde “D” significa DNA, le sigue el número que representa al cromosoma, luego la letra “S” por secuencia de copia única y el número final indica el orden en el cual el marcador fue descubierto y categorizado (Butler, 2005). A modo de ejemplo, en la Figura 1 se esquematiza la nomenclatura del marcador D18S51.

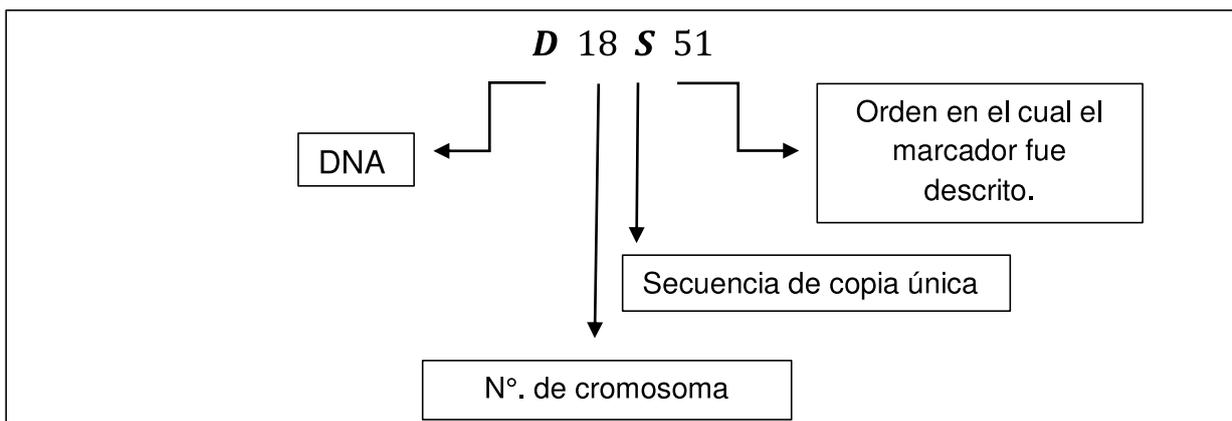


Figura 1: Nomenclatura del marcador D18S51 (Fuente adaptada de Butler (2005)).

Mientras que para los loci STRs o marcadores genéticos que se encuentran dentro de una región codificante de proteínas (así como el intrón de los genes) se utiliza el nombre del gen para designarlo, esto se puede aplicar, por ejemplo: a STRs tales como vWA (GeneBank M25858), su denominación viene de su ubicación en el sitio que codifica para el factor de von Willebrand y para el TPOX (GeneBank M68651), que se encuentra ubicado en el gen 10 de la peroxidasa tiroidea humana y de ahí recibe su nombre (Butler, 2005).

Según el patrón de repetición, los STRs se clasifican en categorías, dependiendo del número de nucleótidos que forman la unidad de repetición, pudiendo ser dinucleótidos, trinucleótidos, tetranucleótidos, pentanucleótidos y hexanucleótidos (Butler, 2005).

Las técnicas moleculares se van sofisticando cada vez más, en sistemas multiplex, permitiendo así, en prácticamente un solo paso, el análisis de los STRs recomendados para su uso en el software CODIS. Estos marcadores son seleccionados cuando presentan alto polimorfismo y tienen alto poder discriminante en la identificación humana (Lorente et al., 2007).

El sistema por ser multiplex, brinda información de todos los loci en una reacción, disminuyendo así el tiempo del análisis y la construcción de base de datos es rápida usando este tipo de sistemas (Lins et al., 1996).

1.5 Antecedentes de estudios de variabilidad y caracterización genética en la población uruguaya.

Existen estudios realizados en varios países del mundo utilizando marcadores STRs autosómicos con la finalidad de caracterizar a los grupos poblacionales. La población uruguaya registra una serie de estudios:

En el año 2001, Pagano y cols., publicaron un estudio realizado en 188 individuos no emparentados entre sí, pertenecientes al departamento de Montevideo, empleando el kit Power Plex 1.2™, tipificándose 8 loci STRs autosómicos, donde se observó que todos los loci analizados, excepto el CSF1P0 cumplen con las expectativas de EHW (Pagano et al, 2001).

Lorente y cols., (2000) realizaron un estudio titulado “Análisis de la distribución alélica de 15 STR en la población uruguaya”, presentado en el III Simposio Latino-Americano de Identificación Humana, Curitiba, Brasil, en junio del 2001, donde se analiza una muestra de la población uruguaya de 200 individuos, empleando el kit Power Plex 16. Se tipificaron 15 loci STRs autosómicos (CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta E, Penta D, TH01, TPOX, y vWA), y se observó que el Penta E fue uno de los loci más polimórficos del conjunto de loci analizados (Lorente et al, 2000).

1.6 Antecedentes de estudios sociodemográficos y genéticos de la población uruguaya.

Las poblaciones latinoamericanas surgieron del desigual aporte de indígenas, africanos y europeos, proceso del cual la población uruguaya no quedó exenta (Sans, 2000). La población de Uruguay, se constituyó esencialmente a partir del aporte de diferentes oleadas migratorias. Los pueblos indígenas, provenientes de tres o cuatro etnias (charrúa, chana, arachanes y guaraní) eran demográficamente pequeños y fueron desapareciendo debido principalmente a exterminios ocurridos en la época (Sans, 2009).

Con relación a los africanos y sus descendientes, las referencias más antiguas los relacionan con el inicio de la colonia, ya que se mencionan a negros esclavos trabajando en la fundación portuguesa de Colonia del Sacramento en 1680 y con el tráfico regular de esclavos que comenzó en el año 1743 (Sans, 2009).

A mediados del siglo XIX, se observó una masiva inmigración europea, que conllevó al remplazo y mestizaje de los pueblos nativos por poblaciones de origen europeo.

En Uruguay el mestizaje ha tenido gran impacto en la conformación de las poblaciones actuales, generando así poblaciones híbridas que dependiendo de la región o grupo analizado, podrán tener el aporte de 2 o 3 de estos grupos; los nativos americanos, europeos y africanos (Sans, 2000).

El análisis del mestizaje, tiene como objetivo comprender el proceso de colonización, deducir las características originales de las poblaciones nativas tempranas y la relación entre grupos étnicos y determinadas enfermedades (Sans, 2000).

Teniendo en cuenta los datos históricos de cómo fue el proceso de mestizaje en Uruguay, se realizaron varios estudios, con la finalidad de conocer los componentes ancestrales de la población:

1.6.1 Contribución ancestral en la población uruguaya

Tacuarembó, presenta una población resultado de varias oleadas migratorias de Europa y Oriente Próximo, así como africanos y afrodescendientes en su mayoría provenientes de Brasil, estas oleadas migratorias se fueron asentando con las diversas etnias nativas del territorio (Sans, 1994). Esta particularidad, lo hace una región interesante desde el punto de vista de la genética de poblaciones, motivo por el cual se han realizado diferentes estudios de mezcla sobre esta población, que van desde el estudio de los rasgos morfológicos (Sans, 1994), marcadores clásicos (grupos sanguíneos, proteínas séricas y sistema HLA). Los cuales mostraron una contribución europea de 65%, 20% nativos americanos y 15% africanos (Sans et al., 1997) y ADN hereditario uniparental (ADN mitocondrial y del cromosoma Y) (Bonilla et al., 2004; Bertoni et al., 2005). El análisis de ADN mitocondrial reveló que la principal contribución a través de la línea materna es el nativo americano con un 62% (Bonilla et al., 2004) mientras que el análisis de cromosoma Y mostró una contribución europea casi exclusiva para la línea paterna y una contribución nativo americana entre 1,6% y 8,3% (Bertoni et al., 2005).

Sans y cols. (1997), a partir de 10 sistemas sanguíneos estimaron las contribuciones parentales en una muestra de la población de Montevideo, donde se estimó un aporte africano de 7%, indígena de 1% y el restante 92%, europeo. Luego, se publicó el primer análisis de ADN mitocondrial (ADNmt) para esta región del país, que determinaba un aporte indígena de 20,4%; dicho estudio no discriminaba los otros posibles aportes (Gascue et al., 2005). Más recientemente, Bonilla y cols. (2015) publicaron los resultados de un estudio sobre cáncer de mama en una muestra de gran parte del país, incluyendo Montevideo. Ese estudio incluyó, entre otros, ADNmt y marcadores individuales de ancestría (ancestry informative markers – AIMs) en autosomas del ADN nuclear (ADNn), donde se observaron aportes de 9,3% africano, 15,3 % indígena y 75,4% europeo (Bonilla et al., 2015). Sans y cols. (2020), realizaron un análisis de la contribución parental de marcadores de ADNmt sobre la muestra poblacional utilizada por Bonilla y cols. (2015), donde se observó que en la muestra de mujeres nacidas en Montevideo, los aportes fueron de 33,9% indígena americano, 7,4% africano y 58,2% europeo/mediterráneo, mientras que en la muestra de residentes fue 36,6%, 8,3% y 55,1% respectivamente. En análisis de ancestría biparental estimada por AIMs, mostró que las mujeres nacidas en Montevideo tuvieron, en promedio, aportes de 12,6% indígena, 78,6% europeo, y 8,8% africano, mientras que, para las residentes los valores promedio fueron similares, 13,8% de aporte indígena, 76,8% de aporte europeo y 9,4% de aporte africano.

Vega-Requena y cols. (2020), publicaron un estudio titulado “Genetic Admixture Analysis in the Population of Tacuarembó-Uruguay Using *Alu* Insertions”. Esta caracterización genética se realizó analizando 8 inserciones *Alu* en una muestra poblacional conformada por 58 individuos descendientes de vascos residentes en

Trinidad, 55 individuos autodefinidos de ascendencia vasca de la ciudad de Montevideo y por ultimo un subconjunto de 70 individuos de ascendencia europea pero sin ascendencia vasca conocida. Estas inserciones se emplean para proporcionar información de ancestralidad y diferenciación genética en la población. Los autores observaron que la contribución ancestral que predomina en Tacuarembó es la europea con un 63%, seguido por la contribución africana con un 30% y la contribución ancestral más pequeña se corresponde con el aporte indígena con un 7% (Vega-Requena et al., 2020).

Las contribuciones ancestrales para la población de Uruguay en conjunto, fueron estimadas por Hidalgo y cols. (2005) en 84,1% europea, 10,4% nativo americana y 5,6% africana. En su trabajo analizó la frecuencia de alelos y fenotipos en 5 loci autosómicos en una muestra de 85 individuos no emparentados, de ambos sexos y nacidos en varias regiones de Uruguay, y realizó una comparación a través de un análisis de distancias genéticas de la población uruguaya con otras poblaciones americanas, donde se observó que la menor distancia se encuentra entre la muestra de la población de Uruguay y una muestra de la población de Buenos Aires, lo cual concuerda con la historia demográfica de ambas regiones (Hidalgo et al., 2005).

1.7 La investigación en el ámbito forense; actualidad y propósitos

Actualmente en Uruguay, en el ámbito forense se emplean los datos de diferentes poblaciones para realizar estimaciones estadísticas, como por ejemplo: datos de la población caucásica de Estados Unidos. Los datos se encuentran incluidos en la base de datos CODIS, y permiten realizar los cálculos estadísticos requeridos (Moretti et al., 2016).

La población uruguaya cuenta con información insuficiente para los 20 marcadores STRs autosómicos recomendados actualmente por el FBI para su uso en el software CODIS, contando únicamente con un estudio realizado por Lorente y cols. (2000), referido a la distribución alélica de 15 marcadores STR autosómicos realizado sobre un muestreo poblacional de 200 individuos.

Consideramos que sería un gran aporte analizar los 20 STRs autosómicos recomendados a partir de 2017 por el FBI, en una muestra de 4000 individuos para conocer con mayor detalle la diversidad y estructura de la población uruguaya. Esto facilita la creación de una base de datos, la cual permitirá a los laboratorios forenses obtener resultados más confiables y representativos en la identificación de personas a través del análisis de vestidos biológicos recogidos de la escena del crimen, restos óseos o cadavéricos y en las investigaciones de paternidad.

2. OBJETIVOS

2.1. Objetivo general

Caracterizar la estructura genética de la población uruguaya.

2.2. Objetivos específicos

1. Analizar 23 loci STRs: CSF1PO, D1S1656, D2S441, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D10S1248, D12S391, D13S317, D16S539, D18S51, D19S443, D21S11, D22S1045, FGA, Penta E, Penta D, TH01, TPOX, SE33 y vWA en una muestra de 4000 individuos de la población uruguaya.

2. Determinar los parámetros estadísticos forenses: equilibrio de Hardy - Weinberg (EHW), la heterocigosidad observada (H_o), la heterocigosidad esperada (H_e), el poder de exclusión (PE), la probabilidad de coincidencia o *match* (PM), el poder de discriminación (PD), el índice de paternidad (IP) y el contenido de información polimórfica (PIC).

3. Realizar una comparación de la diversidad genética uruguaya con la de poblaciones mestizas de la región (Argentina, Bolivia, Brasil, Chile y Paraguay).

4. Analizar la subestructuración de la población en una muestra de 4000 individuos mediante dos aproximaciones diferentes, una basada en inferencia bayesiana (STRUCTURE) y otro analítico (PCA).

5. Analizar la contribución de las poblaciones parentales a la población híbrida en estudio.

3. MATERIALES Y MÉTODOS

El tipo de muestra empleado en este trabajo, se encuentra en una espátula recolectora; “Bode Buccal DNA Collection Systems”, el cual es un sistema de recolección de células bucales del interior de la mejilla (mucosa yugal) para el análisis de identificación de ADN humano. Es un método simple y directo, no invasivo, fácil de usar, diseñado para maximizar la cantidad de ADN recolectado y facilitar la automatización, permitiendo la amplificación por PCR directo, obteniendo resultados de alta calidad en el genotipado. Este diseño facilita el procesamiento automatizado con Sistemas de Perforación BSD 600 Duet puncher (BSD Robotics) (<https://www.bodetech.com/pages/bode-buccal-dna-collection.systems>).

Las muestras de mucosa yugal de personas condenadas o formalizadas por la justicia competente (Ley N°. 18.849), que ingresan a la Dirección Nacional de Policía Científica, previo a su envío al Laboratorio Registro Nacional de Huellas Genéticas, son identificadas con un código alfanumérico, lo que no nos permite conocer la identidad del individuo como tampoco su procedencia dentro del territorio Nacional. El objetivo del empleo de estas muestras completamente anónimas, es generar una Base de Datos de frecuencias alélicas de referencia, imprescindible en la realización de cálculos estadísticos en casos de criminalística forense o filiación.

Se realizó un muestreo completamente al azar entre aquellos perfiles obtenidos en el año 2018 que presentaban los 23 STRs de interés, obteniéndose una submuestra de 4000 individuos compuesta por 501 mujeres (12,5%) y 3499 hombres (87,5%).

3.1 Obtención del perfil genético

El perfil genético fue obtenido mediante amplificación directa de los marcadores genéticos del ADN no codificante con el kit PowerPlex® Fusion 6C System siguiendo las especificaciones del fabricante (Promega Corporation). Este sistema provee todos los materiales necesarios para la co-amplificación y detección fluorescente de 24 loci (23 STRs y Amelogenina), incluyendo los loci requeridos por CODIS y por el conjunto de Estándares Europeos (ESS) (ver Figura 2). Con 24 loci el sistema PowerPlex Fusion ofrece un alto poder de discriminación en una única reacción y proporciona más información para casos forenses difíciles de paternidad y de parentesco (Promega Corporation, 2018).

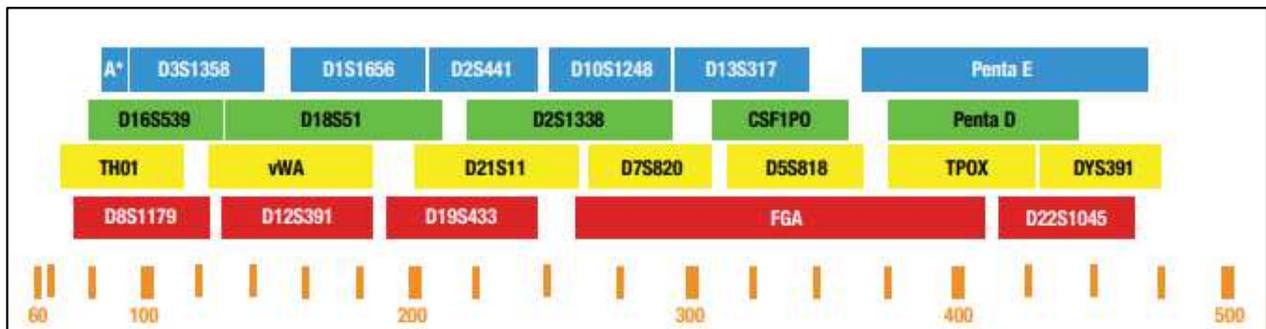


Figura 2: Configuración del sistema PowerPlex Fusion. Este sistema permite la co-amplificación de Amelogenina, D3S1358, D1S1656, D2S441, D10S1248, D13S317 y Penta E marcados con fluoresceína; D16S539, D18S51, D2S1338, CSF1PO y Penta D marcados con JOE; TH01, vWA, D21S11, D7S820, D5S818, TPOX y DYS391 marcados con TMR-ET; y D8S1179, D12S391, D19S443, FGA y D22S1045 marcados con CXR-ET. El estándar interno (ILS 500) está marcado con el colorante CC5 y contiene 21 fragmentos de ADN de 60, 65, 80, 100, 120, 140, 160, 180, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475 y 500 bases (Promega Corporation, 2018).

La amplificación de los STRs, se realizó mediante la técnica de PCR multiplex. El proceso de amplificación se encuentra detallado en el **Anexo 1**. Los productos amplificados fueron detectados por el Analizador Genético Applied Biosystems™ 3500, por medio de electroforesis capilar con detección por fluorescencia.

Luego de la electroforesis capilar, los datos fueron analizados en el software [GeneMapper® ID-X Software v1.2](#), para la tipificación genética, con este se obtuvo y editó los perfiles genéticos. Se revisó la presencia de picos en los controles negativos de cada corrida, para descartar contaminación así como también se observaron los controles positivos para asegurar una correcta amplificación. El kit empleado utiliza un marcador estándar interno de referencia (WEN ILS 500), que se incluye en cada muestra y sobre el cual se interpola el tamaño de los alelos (ver Figura 3). Los pesos moleculares de los fragmentos del ADN que componen el estándar interno empleado son revisados en cada muestra y asignados desde los 60 pb a los 500 pb.

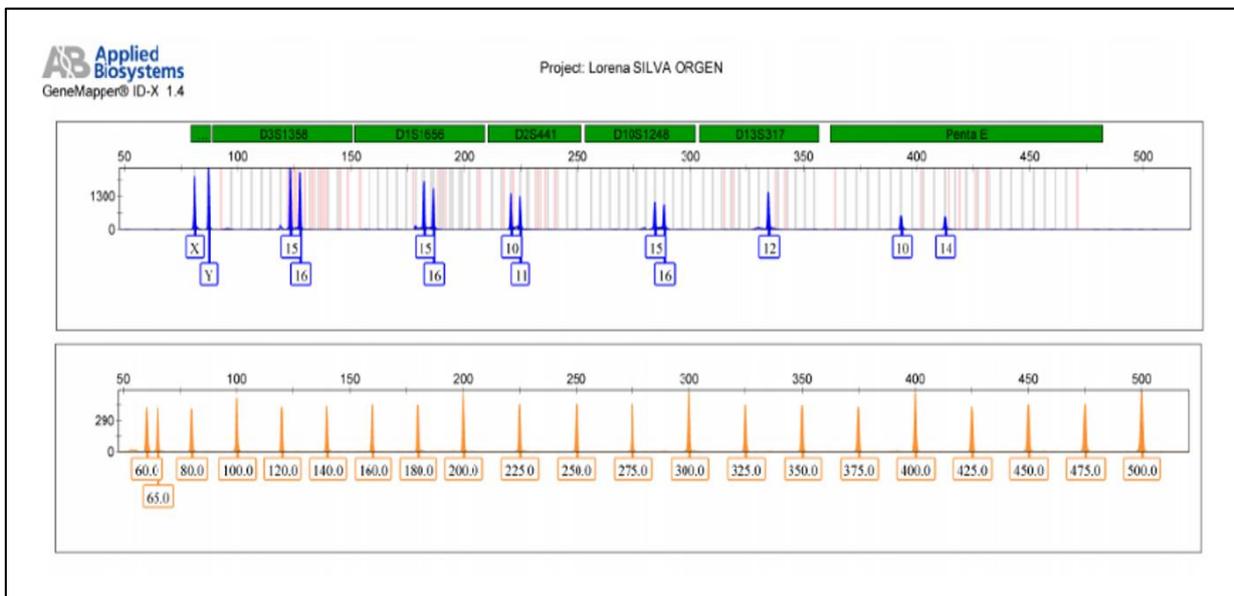


Figura 3: Electroferograma del size standard WEN ILS 500 IDX y ejemplo de asignación a un perfil genético (azul). Se muestran las longitudes de los fragmentos en pares de bases de 60 a 500pb.

Los picos observados son evaluados por los tamaños que refieren en el electroferograma y se consideraron los que estuvieron por encima de 100 RFU y no presentaron artefactos. La asignación de alelos se realizó siguiendo como guía la escalera alélica que utiliza el kit empleado (Ver **Anexo 2**).

Finalmente, luego de que los perfiles genéticos (ver ejemplo en **Anexo 3**) pasaran satisfactoriamente las pruebas de control de calidad, se procedió a exportar los mismos en forma de códigos numéricos, desde el Software GeneMapper® ID-X Software v1.2 en formato de texto (csv) compatible con planillas tipo Excel, donde se reordeno al azar los individuos y se re-codificaron los mismos.

3.2 Análisis estadístico

A fin de determinar la variabilidad de los 23 marcadores STR autosómicos de la población uruguaya, se elaboró una base de datos con los genotipos de los 4000 individuos. Mediante el programa PGDSpider versión 2.1.1.5 (Lischer & Excoffier, 2012), se generaron los archivos necesarios para los distintos programas utilizados en el estudio. Se estimaron las frecuencias alélicas, se testeó el equilibrio de Hardy - Weinberg (EHW) y los siguientes parámetros estadísticos forenses: la heterocigosidad observada (H_o) y la heterocigosidad esperada (H_e) fueron obtenidos empleando el paquete estadístico Arlequin 3.5.2.2 (Excoffier & Lischer, 2010). Por otro lado el programa Forensic Markets (Peña & Alfonso-Sánchez, 2016) nos permitió el cálculo del poder de exclusión (PE), la probabilidad de coincidencia o *match* (PM), el poder de discriminación (PD) y del índice de paternidad (IP). Por último el contenido de información polimórfica (PIC) se realizó mediante el uso de R (<http://www.R-project.org/>) en RStudio (<http://www.rstudio.com/>).

3.2.1 Cálculo de las frecuencias alélicas.

La frecuencias alélicas se calculan mediante el recuento de cada alelo y dividiendo este número por el número total de alelos analizados (Cavalli-Sforza et al., 1981).

3.2.2 Equilibrio de Hardy-Weinberg

La Ley de Hardy-Weinberg, postulada en 1908 por George Hardy y Wilhelm Weinberg, establece que en toda población lo suficientemente grande, en la que los apareamientos se producen al azar y que no se encuentra sometida a ninguna fuerza evolutiva como ser la mutación, selección y migración, las frecuencia génicas y genotípicas se mantienen constantes de una generación a otra. De esta manera, una población que cumpla dichas características se dice que está en equilibrio de Hardy-Weinberg (Chakraborty y cols., 1992). Para poder comprobar esto, se comparó la frecuencia genotípica que se observó en la muestra con respecto a la frecuencia esperada obtenida del producto de combinar las frecuencias alélicas. Para cada loci se determinó si existía ajuste al EHW mediante una prueba exacta de Fisher.

3.3 Parámetros poblacionales

3.3.1 Heterocigosidad

Es estudiada como heterocigosidad observada (H_o) y heterocigosidad esperada (H_e).

La H_o , refleja la frecuencia de individuos heterocigotos encontrados en la muestra poblacional para cualquiera de los loci estudiados (Crow et al., 1970), y se obtiene de acuerdo a la siguiente formula:

$$H_o = 1 - \sum p_i$$

Donde p_i es la frecuencia de cada alelo.

La H_e , se refiere a la posibilidad de que resulten diferentes dos alelos tomados al azar de la población. Se calcula a partir de las frecuencias alélicas (Nei, 1977), ya sea como la sumatoria de los heterocigotos esperados, o la forma más usual:

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

Donde p_i es la frecuencia de cada alelo.

3.3.2 Poder de Exclusión (PE)

El poder de exclusión, se define como la capacidad del panel de marcadores de excluir un genotipo en particular. En casos forenses, es la probabilidad de excluir a un individuo que esté involucrado en el caso y en casos de paternidad, es la probabilidad de excluir un padre erróneamente inculcado (Jamieson, 1994; Krenke et al., 2002; Lee et al., 1998).

El PE se puede calcular siguiendo la siguiente formula:

$$PE = h^2 (1 - 2hH^2)$$

Donde h es la proporción de individuos heterocigotos y H es la proporción de individuos homocigotos en la muestra poblacional (Tillmar, 2010).

También se analizó el poder de exclusión total (PE total) de acuerdo a la siguiente fórmula:

$$PE\ total = 1 - (1 - pe_1)(1 - pe_2) \dots (1 - pe_n)$$

Donde “ pe ” representa el poder de exclusión de cada marcador y “ n ” representa el número total de loci involucrados (Rosado, 2008).

3.3.3 Probabilidad de Coincidencia o match (PM)

La probabilidad de coincidencia o *match* (PM) se define como; la probabilidad de que dos individuos elegidos al azar en una población, tengan idénticos genotipos en uno o más sistemas genéticos, es decir, es la probabilidad de que dos individuos no relacionados compartan un mismo perfil de ADN (Jones, 1972).

Esta probabilidad puede aumentar cuando tenemos un perfil de ADN parcial debido a la degradación o cuando tenemos una reducción del número de locus informativos. Además, también puede aumentar si el sospechoso y el culpable tienen una relación de parentesco o si se originaron a partir de la misma subpoblación. Se excluyen de estos valores, los casos de gemelos monocigóticos o univitelinos, ya que comparten el mismo perfil genético (Quintero et al., 2020).

3.3.4 Poder de Discriminación (PD)

El poder de discriminación permite determinar la probabilidad de que un marcador o un conjunto de marcadores sean capaces de diferenciar genéticamente a

individuos no relacionados entre sí, tomados al azar en la población (Quintero et al., 2020).

Se puede calcular siguiendo la siguiente formula:

$$PD = 1 - PM$$

Donde *PM* la probabilidad de coincidencia o *match* (Tillmar, 2010).

También se analizó el poder de discriminación total (PD total), por ser indicativo de la probabilidad de discriminación para una serie de alelos.

Se calculó mediante la siguiente formula:

$$PD\ total = 1 - (PM_1.PM_2...PM_n)$$

Donde "*PM*" representa el poder de coincidencia y "*n*" representa el número total de loci involucrados.

3.3.5 Índice de paternidad (IP)

Para estimar el grado de certeza con que se asigna una paternidad es necesario confrontar la probabilidad que los alelos presentes en el hijo provengan del presunto padre versus la probabilidad de que los haya recibido de cualquier otro individuo al azar en la población. De acuerdo a esto, se calcula la razón entre ambas probabilidades,

denominado índice de paternidad, que expresa cuántas veces es más probable que el presunto padre sea el padre biológico a que lo sea cualquier otro sujeto en la población.

Se puede calcular siguiendo la siguiente fórmula, donde:

H_0 : El presunto padre es el padre del niño.

H_1 : Un hombre al azar, no relacionado con el presunto padre es el padre del niño.

$$IP = \frac{\Pr(G_C, G_M, G_{PA} | H_0)}{\Pr(G_C, G_M, G_{PA} | H_1)}$$

Lo que significa, la probabilidad de ver el perfil de ADN del niño (G_C), de la madre (G_M) y del padre presunto padre (G_{PA}) cuando PA es el padre, en comparación con la probabilidad de ver los mismos perfiles de ADN cuando PA no es el padre (Tillmar, 2010).

El IP se modificará principalmente en función de la frecuencia con que se observen los alelos heredados del padre en la población general; a menor frecuencia alélica, mayor el IP y mayor la probabilidad de que sea el padre biológico (Lagos et al., 2011).

3.3.6 Contenido de Información Polimórfico (PIC)

El contenido de información polimórfica (PIC), se puede interpretar como la probabilidad de que los alelos materno y paterno de un niño sean deducibles, es decir

la probabilidad de deducir que alelo ha trasmitido un padre al niño (Botstein et al., 1980; Guo & Elston, 1999).

Se puede calcular siguiendo la siguiente formula:

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

Donde p_i y p_j son las frecuencias alélicas.

Es un índice que evalúa la informatividad de un marcador en la población de acuerdo con las frecuencias de los alelos (Botstein et al., 1980). Marcadores con valores de PIC superiores a 0,5 son considerados muy informativos, con valores entre 0,25 y 0,50 medianamente informativos y con valores inferiores a 0,25, poco informativos (Botstein et al., 1980).

3.4. Comparación de la frecuencia alélica entre la población uruguaya y la población disponible en CODIS

La comparación de las frecuencias alélicas de la población uruguaya con la población empleada por CODIS, se realizó con el objetivo de detectar la existencia de diferencias estadísticas significativas entre las poblaciones para cada marcador analizado. De este modo, se trabajó con los datos correspondientes a 2 muestras de la población uruguaya; los correspondientes a este trabajo (URU_4000) y datos obtenidos de un estudio realizado previamente para 15 loci STRs en 200 individuos (URU_200)

pertenecientes a la población uruguaya (ver **Anexo 5**) (Lorente et al, 2000) y los datos de 361 individuos caucásicos de la base de datos CODIS (CODIS_361) del FBI (ver **Anexo 4**) (Hares, 2015; Hill et al., 2013).

Se realizó una prueba de Bartlett para probar si las muestras provienen de poblaciones con la misma varianza a un nivel de confianza determinado. Este test realiza un contraste de hipótesis, donde:

H_0 : las muestras presentan varianzas iguales.

H_1 : las muestras presentan varianzas distintas.

Para verificar este contraste de hipótesis, se adoptó, en todas las comparaciones, un nivel significativo de $p < 0,05$ (Arias, 2017).

3.5. Diferenciación genética

La diferenciación genética, analiza la distribución de la variación genética entre poblaciones, nos permite medir cuan diferente son las poblaciones analizadas entre sí.

A modo de comparar los datos obtenidos para la población uruguaya, con datos publicados de poblaciones regionales se estimó el grado de diferenciación genética entre diferentes poblaciones empleando el Índice de Fijación de Wright (F_{ST}) (Wright, 1951) y Coeficiente de diferenciación genética (G_{ST}) (Nei, 1973).

Se realizó un análisis cualitativo de la bibliografía ya existente, lo que permitió obtener datos de frecuencias alélicas de STRs autosómicos, para las poblaciones de interés.

Se estableció como poblaciones de interés, por un lado aquellas poblaciones latinoamericanas que por su cercanía pudieran presentar o no diferenciación genética con la población uruguaya: Argentina (Buenos Aires, Mendoza y Santa Fe (Marino et al., 2006)), Bolivia (Quispe, 2018), Brasil (Rio de Janeiro (De Souza Góes et al., 2004), San Pablo (Martins et al., 2011), Santa Catarina (Ocampos et al., 2009), Espirito Santo (Silva et al., 2011)), Chile (Vergara et al., 2012), Paraguay (Vullo et al., 2016). Por otro lado, sabiendo que la población uruguaya se caracteriza por ser una población mestiza, que surgió del aporte desigual de nativos americanos, africanos traídos como esclavos y europeos (Sans, 2000), se incluyeron datos poblacionales de Guinea Ecuatorial (África) (Alves et al., 2005) y de España (Camacho et al., 2007).

Los cálculos de F_{ST} y G_{ST} se hicieron tomando en cuenta las frecuencias alélicas de las diferentes poblaciones mencionadas, las cuales posteriormente fueron insertadas en el programa Excel® (Microsoft Office 2013). Posteriormente se estimó el F_{ST} según Weir y Cockerham (1984) y G_{ST} en RStudio (<http://www.rstudio.com/>), empleando el paquete Polysat versión 1.7-4 (Clark & Jasieniuk, 2011), de donde se obtuvo la matriz con la que se construyó un árbol de distancias genéticas utilizando el método de Neighbor-Joining que refleja la relación genética entre las poblaciones analizadas (Saitou & Nei, 1987).

Existen diferentes métodos para estimar la diferenciación genética:

3.5.1. Índice de fijación de Wright (F_{ST})

El estadístico F_{ST} de Wright, es una medida de la correlación entre los genes de una subpoblación en relación a la población entera (Wright, 1951). Permite determinar las diferencias genéticas que existe entre y dentro de las subpoblaciones, mediante los valores de heterocigosidad esperada.

$$F_{ST} = 1 - \frac{H_S}{H_T} = \frac{H_T - H_S}{H_T}$$

Donde H_S es la heterocigosidad esperada media de un grupo de subpoblaciones y H_T heterocigosidad esperada total de la población.

Puede tomar valores que van entre 0 y 1, donde un valor de 0 indica que no hay diferenciación genética entre poblaciones, mientras que valores diferentes a 0 estadísticamente significativos son un indicativo de diferenciación poblacional, es decir presencia de estructura genética.

Las reglas propuestas por Wright en 1978 para la interpretación del estadístico F_{ST} son las siguientes:

$0 \leq F_{ST} \leq 0,05$ indica: “poca diferenciación genética”.

$0,05 < F_{ST} \leq 0,15$ indica: “diferenciación moderada”.

$0,15 < F_{ST} \leq 0,25$ indica: “gran diferenciación”.

$0,25 < F_{ST}$ indica: “diferenciación muy grande”.

3.5.2 Coeficiente de diferenciación genética de Nei (G_{ST})

Nei (1973) propuso una metodología alternativa para analizar la diferenciación genética entre poblaciones, que trabaja directamente con frecuencias alélicas en términos de heterocigosidades esperadas (diversidad genética) dentro y entre poblaciones.

El total de la diversidad genética (H_T), medida en términos del total de la heterocigosidad esperada, puede ser desglosada con el objetivo de determinar la proporción de la diversidad genética de las subpoblaciones que están presentes dentro de una población (H_S) y entre poblaciones (D_{ST} , Nei, 1973), indicando la diferenciación genética entre poblaciones.

$$H_T = D_{ST} + H_S$$

Estos índices de diversidad pueden ser utilizados para calcular medidas de diferenciación genética (G_{ST}).

$$G_{ST} = \frac{D_{ST}}{H_T} = \frac{H_T - H_S}{H_T}$$

El coeficiente de diferenciación genética, G_{ST} se define como la proporción de diversidad genética que existe entre las poblaciones (Nei, 1973). Es equivalente al F_{ST} de Wright (1951) cuando hay solamente dos alelos en un locus y en el caso de alelos múltiples G_{ST} equivale al promedio ponderado de F_{ST} para todos los alelos (Nei, 1973). G_{ST} proporciona una medida de la contribución que hace la heterocigosidad entre poblaciones, D_{ST} , a la diversidad genética total, H_T .

Los valores de G_{ST} pueden variar entre 0 y 1. G_{ST} es 0 cuando el total de la diversidad genética (H_T) es igual a la heterocigosidad esperada (H_S) en cada población, lo que significa que las frecuencias alélicas son idénticas a través de todas las poblaciones. G_{ST} es 1 cuando la diversidad genética de las subpoblaciones que están presentes dentro de una población (H_S) es 0, es decir que no hay variación dentro de las poblaciones.

3.6 Métodos empleados para el análisis de estructuración de la población

3.6.1 STRUCTURE

De manera preliminar, se realizó un análisis de agrupación bayesiana como método de clasificación, implementado en el software STRUCTURE v 2.3.4 (Pritchard et al., 2000), con el objetivo de inferir la estructuración poblacional y el número más probable de poblaciones para la población de Uruguay. Este método analiza las diferencias en la distribución de variantes genéticas entre las poblaciones, identifica las diferentes poblaciones a partir de los datos y asigna a estas poblaciones, mediante un índice de membresía a aquellos individuos que comparten patrones de variación similares (Pritchard et al., 2000).

Para el cálculo, este programa utiliza un enfoque bayesiano sistemático que aplica la estimación de *Markov Chain Monte Carlo* (MCMC). Comienza asignando aleatoriamente individuos a un número predeterminado de grupos, luego se estiman las frecuencias en cada grupo y los individuos son reasignados en función de esas estimaciones de frecuencia (Pritchard et al., 2000). Esto se repite muchas veces, en este trabajo se realizaron 10.000 iteraciones en el proceso de burnin y 100.000 pasos

Markov Chain Monte Carlo luego del burnin, a fin de obtener estimaciones confiables de frecuencias alélicas en cada población y probabilidades de membresía de individuos a una población. *Structure* realiza análisis individuales para cada número de población, aplicando el modelo de K poblaciones o grupos genéticos, y los individuos son asignados aleatoriamente de 1 a K grupos, cada una de los cuales está caracterizado por las frecuencias alélicas en cada locus (Pritchard et al., 2000).

El valor de K no se define fácilmente, y como este valor debe preseleccionarse, los valores definidos por el usuario deben considerarse cuidadosamente teniendo en cuenta las características de las poblaciones muestreadas (Porras-Hurtado et al., 2013). Para la muestra de la población uruguaya se realizó una corrida evaluando el número de poblaciones o grupos genéticos (K), de entre 1 (asumiendo una sola población) y 6 (asumiendo 6 poblaciones). Para cada K se repitió el análisis 10 veces de forma de verificar la consistencia de los resultados. En este análisis se asumió un modelo de mezclas, y de frecuencias alélicas correlacionadas (Porras-Hurtado et al., 2013). Los resultados de esta corrida fueron analizados en STRUCTURE HARVESTER (Earl & VonHoldt, 2012), que permite detectar el número de grupos para los datos simulados usando el estadístico ΔK de Evanno, pero se debe tener en cuenta que este estadístico no mostrará estimaciones para la primera y la última K que se pruebe (Evanno et al., 2005).

En base a los resultados obtenidos, se efectuó una segunda corrida incluyendo a la población uruguaya y los datos de las poblaciones que se encuentran disponibles en el software CODIS (afroamericanos, asiáticos y caucásicos), empleando valores de entre 1 y 6 (abarcando el K óptimo obtenido en la corrida para la población uruguaya).

En este análisis se asumió un modelo de mezclas, y de frecuencias alélicas correlacionadas (Porras-Hurtado et al., 2013). La falta de información en CODIS para la población nativo americana necesaria para realizar el análisis, obligó a utilizar como variable proxy a la población asiática. Los resultados de esta segunda corrida fueron analizados en STRUCTURE HARVESTER para determinar los valores óptimos de K definitivos.

Existe una alternativa a los análisis bayesianos de STRUCTURE, que a diferencia de éste no asume el equilibrio de Hardy-Weinberg: el Análisis de Componentes Principales o PCA.

3.6.2. Análisis de Componentes Principales (PCA)

A modo de efectuar un análisis exploratorio de los datos y observar la existencia de subgrupos para las variables observadas, se realizó un análisis de componentes principales o PCA. Este método busca separar a los individuos analizados en grupos (o *clusters*) de acuerdo a los genotipos que presentan los loci analizados y a la variación que presentan en éstos. Este análisis se efectuó empleando R en RStudio, aplicando la función *dapc* (paquete *adegenet* (Jombart & Ahmed, 2011)) con los datos de las 4 poblaciones empleadas por CODIS y de la población uruguaya. Esto arrojó el agrupamiento de los individuos analizados en diferentes clústers.

3.7 Análisis de mestizaje de la población uruguaya.

3.7.1 Poblaciones empleadas para el análisis del mestizaje de la población uruguaya.

Para realizar un análisis de mestizaje, es necesario elegir aquellas poblaciones que dieron origen a la población mestiza en estudio. Se debe tener en cuenta, por un lado, los datos históricos de cómo fue el proceso de mestizaje y por otro lado, trabajar con aquellos marcadores que presentan frecuencias significativamente distintas entre las poblaciones parentales (Shriver et al., 1997).

Para estudiar el pasado de una población y los cambios que ha experimentado a lo largo del tiempo, se emplean marcadores moleculares de ADN mitocondrial, de cromosomas autosómicos o del cromosoma Y; aportando diferentes resultados, que dependen del marcador utilizado. Los marcadores moleculares presentes en cromosomas autosómicos nos permiten conocer el aporte de ambas líneas; paterna y materna, indistintamente de la población; mientras que los marcadores presentes en el cromosoma mitocondrial como el cromosoma Y, nos brindaran información proveniente de la línea materna o paterna, respectivamente (Butler, 2005).

En primer lugar, se realizó un análisis exhaustivo en la bibliografía (Alves et al., 2005; Sala et al., 2010; Crossetti et al., 2008; Paredes et al., 2003) ya existente, con el propósito de obtener las frecuencias alélicas con que se encuentran los microsatélites en las distintas poblaciones. Se definió el uso de 15 marcadores STRs, de los cuales se dispone información para aquellas poblaciones parentales de interés. Posteriormente, estas poblaciones fueron divididas en dos grupos, donde el primero de

ellos está formado por tres poblaciones parentales y el otro está constituido por la población de Uruguay; considerada como población mestiza.

Para el caso de las poblaciones africanas, se trabajó con individuos de la población de Guinea Ecuatorial (Alves et al., 2005), para las poblaciones nativo americanas se analizaron individuos pertenecientes a comunidades Guaraní de la localidad de Misiones - Argentina (Sala et al., 2010) e individuos pertenecientes a la población nativa del Gran Chaco - Argentina (Crossetti et al., 2008). Y para el caso de la población europea, se analizaron individuos pertenecientes a la población del noreste de España (Paredes et al., 2003), todas ellas para un total de 15 marcadores (CSF1P0, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX, vWA).

Para cada grupo parental formado por dos o más poblaciones, el cálculo de las frecuencias alélicas de cada grupo, se realizó sumando todos los individuos de las distintas poblaciones correspondientes a europeos, nativo americanos y africanos y se calcularon las frecuencias alélicas y su correspondiente conteo.

3.7.2 Métodos utilizados para el Análisis de Mestizaje.

3.7.2.1 ADMIX 95

ADMIX 95 (Chakraborty, 1985) es un método de identidad génica diseñado para el análisis de mezclas en poblaciones híbridas. En base a las similitudes genéticas entre las poblaciones, este método establece una relación entre la población híbrida y las poblaciones parentales como una función lineal (Chakraborty, 1985). Calcula tantas

estimaciones como poblaciones parentales se hayan introducido, permitiendo así evaluar la calidad de las poblaciones parentales (Bertoni, 2005).

3.7.2.2 VARMIX 3.0

VARMIX 3.0 (Bertoni, 2005) es un método de varianzas alélicas diseñado para el análisis de mezclas. Este método calcula la contribución de cada población parental a la población híbrida en estudio. Para esto emplea la información molecular de los microsatélites, descomponiendo la varianza alélica de la población híbrida en las distintas varianzas de las poblaciones parentales que la componen, permitiendo así estimar el peso relativo o histórico de los individuos de una población parental en la población híbrida (Bertoni et al., 2012).

4. RESULTADOS.

4.1 Frecuencias alélicas y parámetros estadísticos en la muestra de la población de Uruguay.

Las frecuencias alélicas obtenidas de los 23 marcadores STRs autosómicos en la población estudiada se muestran en la Tabla 1a y 1b. Algunos marcadores presentan un número mayor de alelos, como es el caso del marcador SE33, el cual presentó 58 alelos, en tanto que el marcador TPOX presentó el menor número de alelos (9 alelos).

Al analizar el test de EHW se obtuvieron valores de “p” mayores a 0,05, observándose una desviación del equilibrio en tres marcadores genéticos (D8S1179, D13S317 y D12S391), sin embargo, para dos de estos marcadores (D12S391 y D13S317) esta desviación resultó no significativa tras aplicar la corrección de Bonferroni, obteniéndose valores de “p” mayores a 0,02, mientras que el marcador D8S1178 mostro desviación del EHW.

En la Tabla 2, se presentan los parámetros estadísticos para los 23 STRs analizados, donde se puede observar que el STRs más informativos por su alto contenido de información polimórfica (PIC), es el SE33 (0,9391), mientras que el marcador que presenta el índice más bajo es el TPOX (0,6200).

En la muestra poblacional uruguaya, el marcador SE33 presentó el PD más alto (0,9944), mientras que el TPOX con 0,8416 fue el marcador con el PD más bajo. En

cuanto al poder de exclusión (PE), el marcador SE33 (0,8893) es el más informativo, mientras que el TPOX fue el marcador con el PE más bajo.

Analizando la Ho y He individual se observó que el marcador SE33 presenta el valor más alto (0,9455 y 0,9458 respectivamente). Por otro lado, los índices más bajos de la Ho y He, corresponden a los marcadores D22S1045 (0,6888 y 0,6944) y TPOX (0,6750 y 0,6686 respectivamente).

El PD y PE total fue de 99,999999%, por lo cual se evidencia que los 23 marcadores STRs analizados, si intervienen en conjunto en los cálculos de probabilidad pueden identificar a los individuos, discriminando y excluyendo a los falsamente implicados.

Tabla 1a: Frecuencias alélicas observadas en los 23 loci STRs analizados para los alelos comprendidos entre el 2.2 y 18.3

ALELO	CSF1PO	D1S1656	D2S1338	D2S441	D3S1358	D5S818	D7S820	D8S1179	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D21S11	D22S1045	FGA	PentaD	PentaE	SE33	TH01	TPOX	vWA	
2.2																								
3.2																								
5.0																								
6.0																								
6.1																								
6.3																								
6.4																								
7.0																								
7.3																								
8.0																								
8.2																								
8.3																								
9.0																								
9.1																								
9.2																								
9.3																								
9.4																								
10.0																								
10.1																								
10.2																								
10.3																								
11.0																								
11.2																								
11.3																								
12.0																								
12.2																								
12.3																								
13.0																								
13.2																								
13.3																								
14.0																								
14.2																								
14.3																								
15.0																								
15.2																								
15.3																								
16.0																								
16.1																								
16.2																								
16.3																								
16.4																								
17.0																								
17.1																								
17.2																								
17.3																								
18.0																								
18.2																								
18.3																								

La información contenida en la presente tabla es de dominio de la Dirección Nacional de Policía Científica – Ministerio del Interior.

Solicitar a: cientifica-ayudantia@minterior.gub.uy

Tabla 1b: Frecuencias alélicas observadas en los 23 loci STRs analizados para los alelos comprendidos entre el 19 y 46.2

	CSF1PO	D1S1656	D2S1338	D2S441	D3S1358	D5S818	D7S820	D8S1179	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D21S11	D22S1045	FGA	PentaD	PentaE	SE33	TH01	TPOX	vWA	
ALELO																								
19.0																								
19.1																								
19.2																								
19.3																								
20.0																								
20.1																								
20.2																								
20.3																								
21.0																								
21.1																								
21.2																								
22.0																								
22.2																								
22.3																								
23.0																								
23.2																								
24.0																								
24.2																								
24.3																								
25.0																								
25.2																								
26.0																								
26.2																								
27.0																								
27.2																								
28.0																								
28.2																								
29.0																								
29.1																								
29.2																								
29.3																								
30.0																								
30.2																								
31.0																								
31.2																								
32.0																								
32.2																								
33.0																								
33.1																								
33.2																								
34.0																								
34.2																								
35.0																								
35.1																								
35.2																								
36.0																								
37.0																								
43.2																								
44.2																								
45.2																								
46.2																								

**La información contenida en la presente tabla es de dominio de la Dirección Nacional de
Policía Científica – Ministerio del Interior.**

Solicitar a: cientifica-ayudantia@minterior.gub.uy

Tabla 2: Parámetros estadísticos forenses de los 23 loci STRs analizados.

	CSF1P0	D1S1656	D2S1338	D2S441	D3S1358	D5S818	D7S820	D8S1179	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D21S11	D22S1045	FGA	Penta D	Penta E	SE33	TH01	TPOX	vWA	TOTAL	
N	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	4000	-
Ho	0,7360	0,9008	0,8758	0,7605	0,7703	0,7285	0,7860	0,8003	0,7595	0,8763	0,8100	0,7770	0,8760	0,8033	0,8430	0,6888	0,8730	0,8428	0,9028	0,9455	0,7925	0,6750	0,7908	0,7908	0,809
He	0,7302	0,8950	0,8723	0,7627	0,7694	0,7283	0,7956	0,8141	0,7570	0,8875	0,8091	0,7862	0,8761	0,8119	0,8467	0,6944	0,8757	0,8452	0,9038	0,9458	0,7948	0,6686	0,7977	0,7977	0,811
EHW	0,7182	0,8174	0,3902	0,2930	0,8339	0,8233	0,7535	0,0133	0,9626	0,0373	0,0232	0,3969	0,7632	0,0571	0,9764	0,1196	0,9162	0,8189	0,5897	0,3362	0,3738	0,1199	0,2843	0,2843	-
PM	0,1213	0,0204	0,0288	0,0946	0,0897	0,1170	0,0719	0,0585	0,0980	0,0230	0,0618	0,0773	0,0282	0,0589	0,0412	0,1453	0,0286	0,0433	0,0170	0,0056	0,0744	0,1584	0,0700	0,0700	0,066
PD	0,8787	0,9796	0,9712	0,9054	0,9103	0,8830	0,9281	0,9415	0,9020	0,9770	0,9382	0,9227	0,9718	0,9411	0,9588	0,8547	0,9714	0,9567	0,9830	0,9944	0,9256	0,8416	0,9300	0,9300	1
PE	0,4764	0,7850	0,7391	0,5315	0,5428	0,4732	0,5907	0,6253	0,5217	0,7697	0,6159	0,5735	0,7466	0,6210	0,6881	0,4195	0,7459	0,6852	0,8029	0,8893	0,5892	0,3812	0,5946	0,5946	1
PI	1,8528	4,7570	3,9135	2,1060	2,1677	1,8394	2,4448	2,6883	2,0568	4,4402	2,6181	2,3378	4,0312	2,6561	3,2586	1,6355	4,0199	3,2280	5,1903	9,2033	2,4350	1,5082	2,4703	2,4703	-
PIC	0,6811	0,8612	0,8595	0,7158	0,7327	0,6850	0,7651	0,7899	0,7180	0,8577	0,7837	0,7546	0,8629	0,7123	0,7922	0,6424	0,8594	0,8246	0,8960	0,9391	0,7546	0,6200	0,7685	0,7685	0,777

N: número de individuos analizados; Ho: heterocigosidad observada; He: heterocigosidad esperada; p-value para el EHW: equilibrio de Hardy – Weinberg; PM: probabilidad de coincidencia o *match*; PD: poder de discriminación; PE: poder de exclusión; PI: índice de paternidad; PIC: contenido de información polimórfica.

4.2 Comparación de las frecuencias alélicas entre la población uruguaya y los datos poblaciones disponibles en CODIS.

Para desarrollar el estudio de comparación de las frecuencias alélicas entre la población uruguaya y los datos poblaciones disponibles en CODIS, se procedió a comparar los datos URU_4000, los datos URU_200 y los datos obtenidos por el FBI; CODIS_361.

Se observó la existencia de alelos propios tanto para las poblaciones uruguayas analizadas como para la población caucásica disponible en el software CODIS. Así mismo se detectó la presencia de alelos que se encuentran con mayor frecuencia en la población uruguaya que en la población CODIS y viceversa.

A modo de ejemplificar, se expone la comparación de las frecuencias alélicas, entre las poblaciones mencionadas para aquellos marcadores STRs que del presente trabajo resultaron más informativos, por su alto contenido de información polimórfica (PIC); el marcador Penta E (0,8960), D18S51 (0,8629), y el FGA (0,8594), así como también para el marcador que presenta el índice más bajo; TPOX (0,6200).

4.2.1 Comparación para el marcador Penta E.

En la Figura 4, se expone una comparación entre las frecuencias alélicas obtenidas para URU_4000, URU_200 y CODIS_361 para el marcador Penta E.

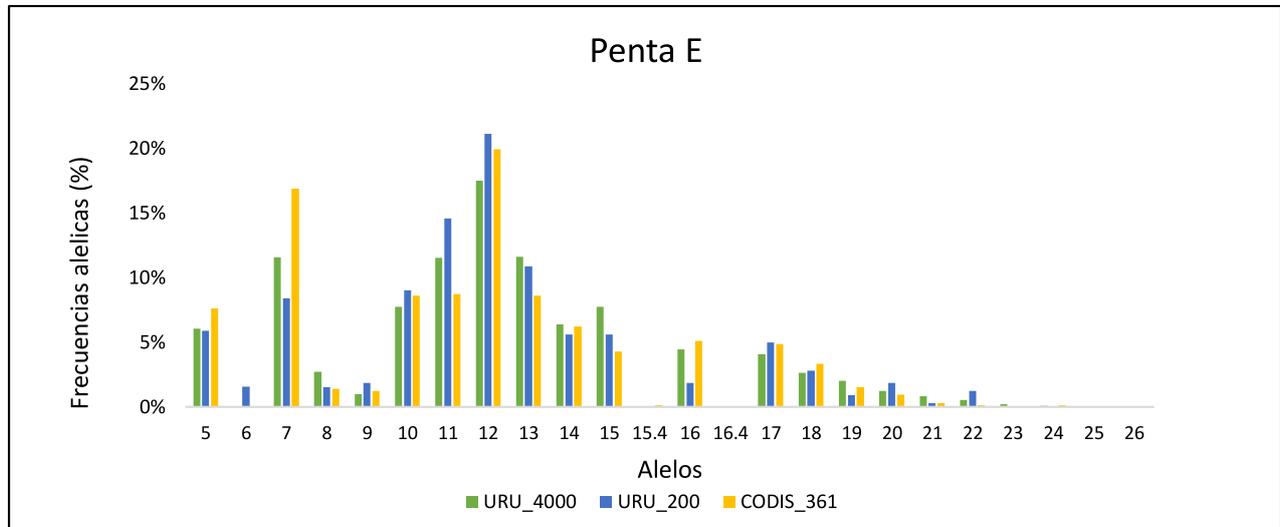


Figura 4: Distribución de las frecuencias alélicas en las poblaciones uruguayas (URU_4000 y URU_200) y en la población CODIS_361 para el marcador Penta E.

Al realizar una comparación de las frecuencias alélicas para este marcador, se pudo apreciar que en la muestra URU_4000, los alelos con mayor frecuencia alélica son el 12 con un 17,50%, seguido por los alelos 13 (11,60%), 7 (11,56), y 11 (11,55%), tendencia esta que se mantiene si observamos las frecuencias alélicas obtenidas para URU_200, con la diferencia que el alelo 11 presentó una frecuencia alélica más elevada (14,60%). Mientras tanto, los alelos con mayor frecuencia alélica en la muestra CODIS_361 son el 12 con 19,94% y el 7 con un 16,90%.

Se observó a su vez, que los alelos 16.4, 23, 25 y 26 se encuentran presentes únicamente en la muestra URU_4000, en un porcentaje que no supera el 0,21%. Por otro lado, el alelo 15.4 únicamente se observó para la muestra CODIS_361 con una frecuencia de 0,14%.

4.2.2 Comparación para el marcador D18S51

En la Figura 5, se expone una comparación entre las frecuencias alélicas obtenidas para URU_4000, URU_200 y CODIS_361 para el marcador D18S51.

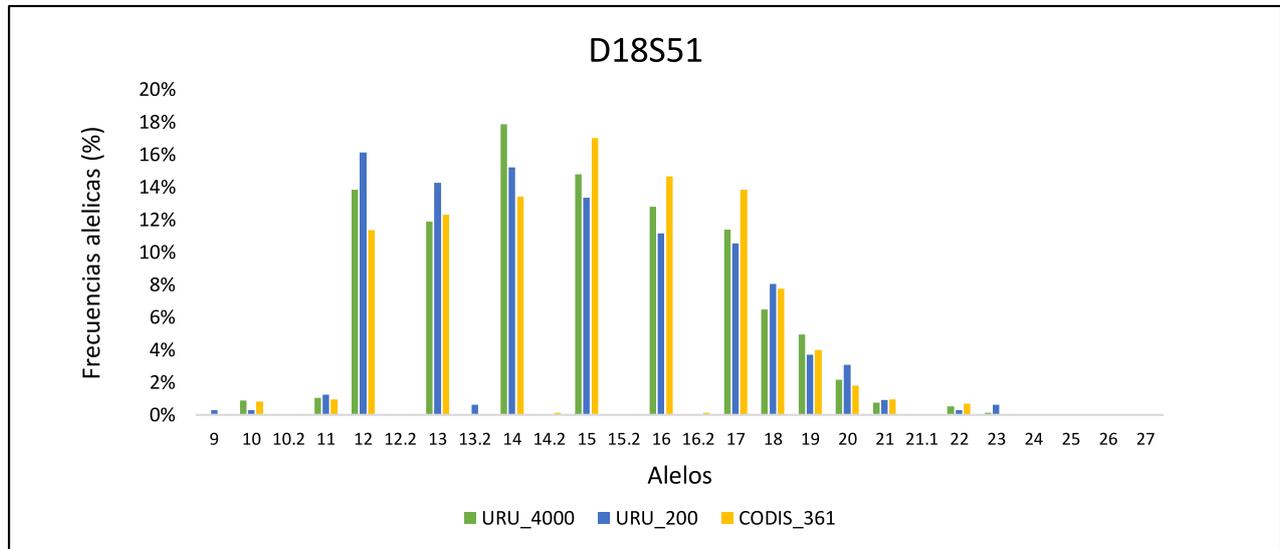


Figura 5: Distribución de las frecuencias alélicas en las poblaciones uruguayas (URU_4000 y URU_200) y en la población CODIS_361 para el marcador D18S51.

Para este marcador, los alelos que presentan mayor frecuencia alélica en la muestra URU_4000, son el 14 con un 17,89%, seguido por el 15 (14,81%), el 12 (13,84%), 16 (12,81%), 13 (11,88%) y el 17 (11,40%). En los datos obtenidos para URU_200, los alelos que presentan mayor frecuencia alélica son el 12 con un 16,15%, seguido por el alelo 14 (15,22%), 13 (14,29%), el 15 (13,35%), 16 (11,18%) y el 17 (10,56%). En la muestra CODIS_361, se observa que los alelos que presentan mayor frecuencia alélica son el 15 (17,04%), 16 (14,68%), 17 (13,85%), 14 (13,43%), 13 (12,33%) y el 12 (11,36%).

Los alelos 10.2, 12.2, 15.2, 21.1, 24, 25, 26 y 27 se obtuvieron únicamente para la muestra URU_200, no observándose en las demás poblaciones, mientras que los alelos 9, 13.2 y 23, están presentes en la muestra poblacional URU_200, pero no en la población CODIS_361. Por otro lado el alelo, 14.2 que se observa únicamente en la muestra CODIS, no se detectó en las muestras poblacionales uruguayas.

4.2.3 Comparación para el marcador FGA

En la Figura 6, se expone una comparación entre las frecuencias alélicas obtenidas para URU_4000, URU_200 y CODIS_361, para el marcador FGA.

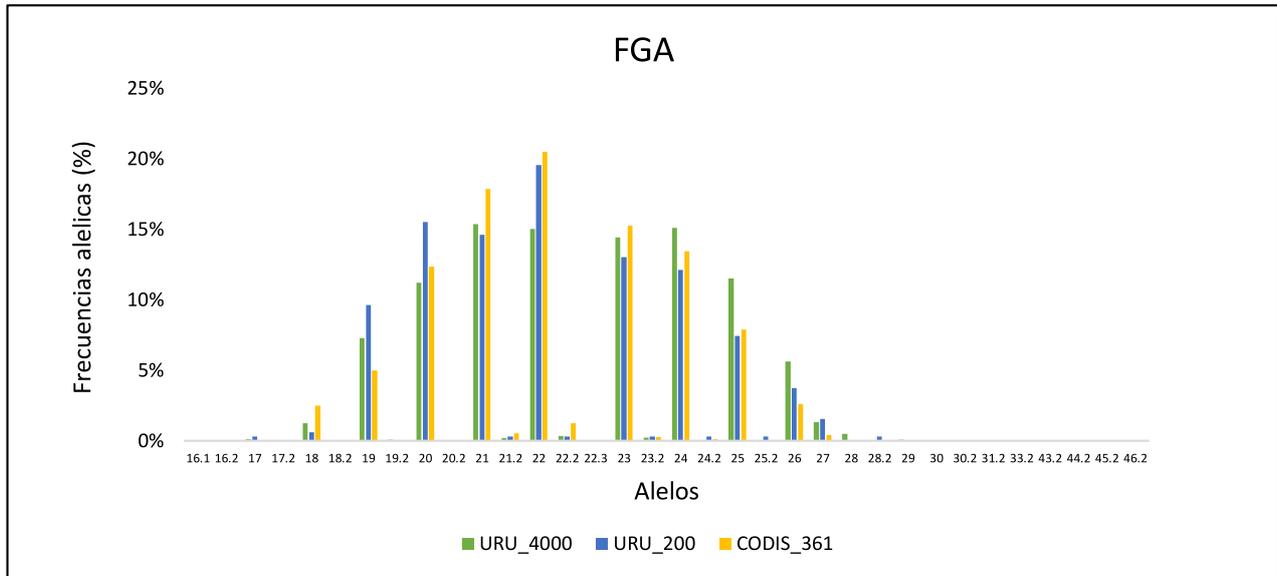


Figura 6: Distribución de las frecuencias alélicas en las poblaciones uruguayas (URU_4000 y URU_200) y en la población CODIS_361 para el marcador FGA.

Para este marcador se observó que los alelos 16.1, 17.2, 18.2, 19.2, 20.2, 22.3, 28, 29, 30, 30.2, 31.2, 33.2, 43.2, 44.2, 45.2 y 46.2 se encuentran presentes únicamente en la muestra poblacional URU_4000, mientras que el alelo 28.2 se observó únicamente en la muestra URU_200.

En la muestra URU_200, se observó que los alelos con mayor frecuencia alélica son, el 22 con un 19,57% seguido por los alelos 20 (15,53%), 21 (14,60%), 23 (13,04%), 24 (12,11%). Mientras que los alelos que presentaron menos frecuencia alélica son el 17, 21.2, 22.2, 23.2, 25.2 y 28.2 con un 0,31%. Por otro lado, los alelos con mayor frecuencia alélica en la muestra URU_4000 son el 21 con un 15,39%, seguido por los alelos 24 (15,09%), 22 (15%), 23 (14,41%), 25 (11,51%) y el 20 (11,21%), mientras que los alelos

que presentan menor frecuencia alélica se corresponden con aquellos alelos que se observaron únicamente en esta población.

Las frecuencias alélicas obtenida para la muestra poblacional CODIS_361, permiten observar que para este marcador los alelos que presentan mayor frecuencia alélica son el 22 con un 20,50%, seguidos por los alelos 21 (17,87%), 23 (15,24%), 24 (13,43%) y el 20 (12,33%). Mientras que el alelo con menor frecuencia alélica es el 24.2 con un 0.14%.

4.2.4 Comparación para el marcador TPOX

En la Figura 7, se expone una comparación entre las frecuencias alélicas obtenidas para URU_4000, URU_200 y CODIS_361, para el marcador TPOX.

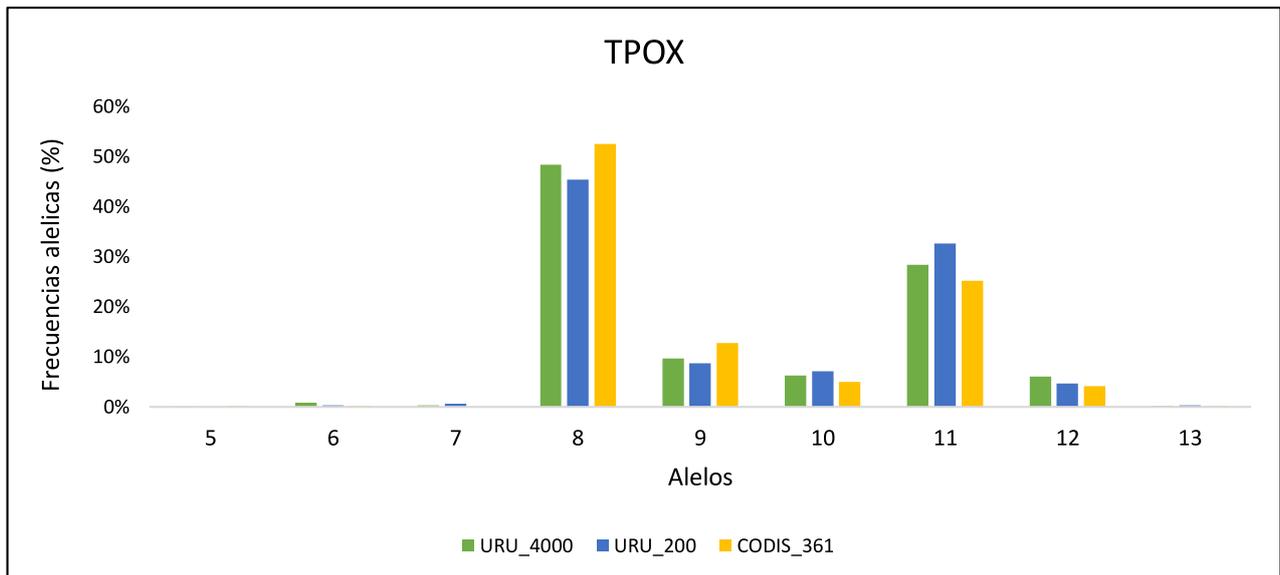


Figura 7: Distribución de las frecuencias alélicas en las poblaciones uruguayas (URU_4000 y URU_200) y en la población CODIS_361 para el marcador TPOX.

Para este marcador, los alelos con mayor frecuencia alélica observados en la muestra CODIS_361, coinciden con los observados en las muestras poblacionales de

Uruguay (URU_4000 y URU_200); el alelo 8 con un 48,40% y 45,34% respectivamente, seguido por el alelo 11 con un 28,34% y un 32,61%.

En este marcador, el alelo 7 se encuentra presente en la población uruguaya, pero no en la población CODIS_361 y su frecuencia alélica es de 0,28% en URU_4000 y 0,62% en URU_200.

Con los datos empleados en la comparación de frecuencias alélicas, se realizó una prueba de Bartlett para probar si las muestras provienen de poblaciones con la misma varianza, a un nivel de confianza de 95%. Los resultados se pueden observar en la Tabla 3. Se obtuvieron valores de $p > 0,05$, para todos los marcadores STRs analizados, lo que nos indica que no rechazamos la hipótesis nula, lo que significa que las varianzas de las poblaciones de las que se extrajeron las diferentes muestras son iguales.

Tabla 3: Análisis de varianza (Test de Bartlett)

Locus	<i>p</i>	Locus	<i>p</i>	Locus	<i>p</i>	Locus	<i>p</i>
CSF1P0	0,9092	D7S820	0,9586	D18S51	0,9431	PENTA E	0,8105
D1S1656	0,7261	D8S1179	0,9769	D19S433	0,4817	SEE33	0,8563
D2S441	0,943	D10S1248	0,9219	D21S11	0,5429	TH01	0,7614
D2S1338	0,6032	D12S391	0,9069	D22S1045	0,9114	TPOX	0,9645
D3S1358	0,9721	D13S317	0,922	FGA	0,3994	VWA	0,9768
D5S818	0,9378	D16S539	0,9262	PENTA D	0,6546		

4.4 Diferenciación genética

4.4.1 Índice de fijación de Wright (F_{ST})

Los valores F_{ST} resultantes entre las distintas poblaciones se observan en forma de matriz en la Tabla 4. Los valores obtenidos son muy cercanos a 0 ($F_{ST} \leq 0,05$), por lo que hay muy poca diferenciación genética entre las poblaciones estudiadas.

Tabla 4: Índice de fijación de Wright (F_{ST}) entre poblaciones.

	Uruguay	Guinea Ecuatorial	Buenos Aires	Santa Fe	Mendoza	Rio de Janeiro	San Pablo	Santa Catarina	Espirito Santo	Bolivia	Chile	Paraguay	España
Uruguay													
Guinea Ecuatorial	0,0015												
Buenos Aires	0,0004	0,0093											
Santa Fe	0,0019	0,0104	0,0050										
Mendoza	0,0003	0,0131	0,0014	0,0041									
Rio de Janeiro	0,0011	0,0076	0,0053	0,0079	0,0063								
San Pablo	0,0003	0,0070	0,0023	0,0051	0,0027	0,0037							
Santa Catarina	0,0003	0,0019	0,0007	0,0025	0,0006	0,0014	0,0003						
Espirito Santo	0,0003	0,0130	0,0022	0,0037	0,0041	0,0044	0,0021	0,0003					
Bolivia	0,0040	0,0283	0,0154	0,0175	0,0137	0,0179	0,0199	0,0060	0,0185				
Chile	0,0050	0,0105	0,0095	0,0053	0,0075	0,0067	0,0099	0,0065	0,0051	0,0108			
Paraguay	0,0022	0,0112	0,0057	0,0008	0,0045	0,0089	0,0064	0,0031	0,0046	0,0172	0,0057		
España	0,0018	0,0127	0,0069	0,0099	0,0079	0,0038	0,0063	0,0021	0,0055	0,0184	0,0059	0,0114	

A partir de estos datos y utilizando el algoritmo de Neighbor-Joining (NJ) se generó un árbol de distancias genéticas con las poblaciones analizadas (Figura 8).

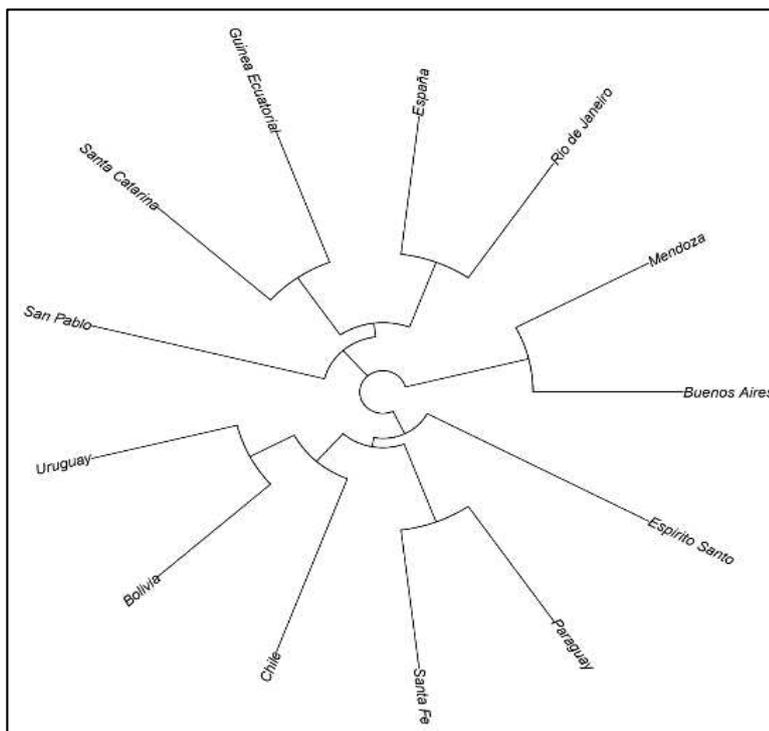


Figura 8: Árbol de distancias genéticas realizado a partir de los datos obtenidos de Índice de fijación de Wright (F_{ST}) utilizando el método de Neighbor-Joining.

4.4.2 Coeficiente de diferenciación genética de Nei (G_{ST})

Los valores G_{ST} resultantes entre las distintas poblaciones se observan en forma de matriz en la Tabla 5. Los datos negativos se pueden considerar como “0”, es decir, no hay diferenciación en las poblaciones. Los valores obtenidos son muy cercanos a 0 ($G_{ST} \leq 0,05$), por lo que hay muy poca diferenciación genética entre las poblaciones analizadas.

Tabla 5: Diferenciación genética de Nei (G_{ST}) entre poblaciones

	Uruguay	Guinea Ecuatorial	Buenos Aires	Santa Fe	Mendoza	Rio de Janeiro	San Pablo	Santa Catarina	Espírito Santo	Bolivia	Chile	Paraguay	España
Uruguay													
Guinea Ecuatorial	0,0103												
Buenos Aires	0,0005	0,0120											
Santa Fe	0,0032	0,0144	0,0035										
Mendoza	0,0002	0,0127	0,0001	0,0028									
Rio de Janeiro	0,0029	0,0061	0,0040	0,0066	0,0042								
San Pablo	0,0003	0,0071	0,0012	0,0038	0,0014	0,0019							
Santa Catarina	0,0002	0,0099	0,0008	0,0037	0,0011	0,0027	0,0001						
Espírito Santo	0,0005	0,0093	0,0012	0,0045	0,0023	0,0024	0,0001	0,0001					
Bolivia	0,0168	0,0295	0,0158	0,0185	0,0119	0,0172	0,0198	0,0197	0,0214				
Chile	0,0079	0,0178	0,0079	0,0040	0,0069	0,0065	0,0089	0,0087	0,0098	0,0131			
Paraguay	0,0042	0,0153	0,0044	0,0000	0,0032	0,0077	0,0051	0,0049	0,0064	0,0178	0,0044		
España	0,0046	0,0130	0,0053	0,0084	0,0058	0,0020	0,0043	0,0042	0,0043	0,0184	0,0055	0,0100	

A partir de estos datos y utilizando el algoritmo de Neighbor-Joining (NJ) se generó un árbol de distancias genéticas con las poblaciones analizadas (Figura 9).

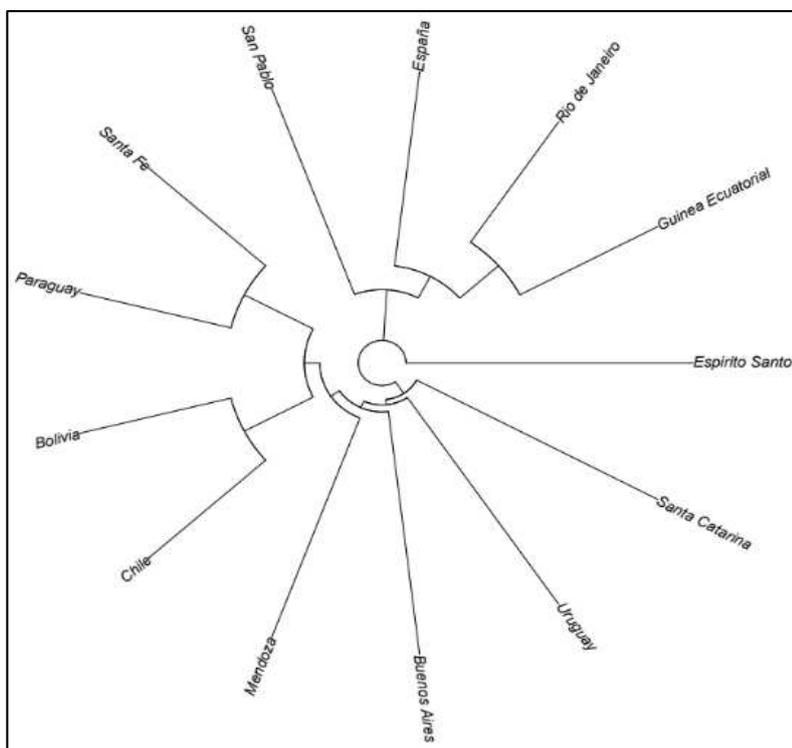


Figura 9: Árbol de distancias genéticas realizado a partir de los datos obtenidos para el Coeficiente de diferenciación genética de Nei (G_{ST}) utilizando el método de Neighbor-Joining.

4.5 Análisis de la estructura genética en la muestra poblacional uruguaya empleando el Software STRUCTURE v 2.3.4.

Evanno y cols, (2005) identificaron una magnitud (ΔK) basada en el segundo orden de cambio de magnitud de la probabilidad entre K que presenta un pico en el valor real de K. Los resultados del análisis con STRUCTURE fueron procesados con el software Harvester (v 0.6.94) (Earl & vonHoldt, 2012) para determinar el K más probable por medio del método de Evanno (Evanno et al., 2005). En el análisis donde se consideró únicamente a los individuos de la población uruguaya, el número más probable de poblaciones o grupos genéticos (K) de acuerdo con el método de Evanno, fue de 2 ($\Delta K=7.173185$) (Figura 10).

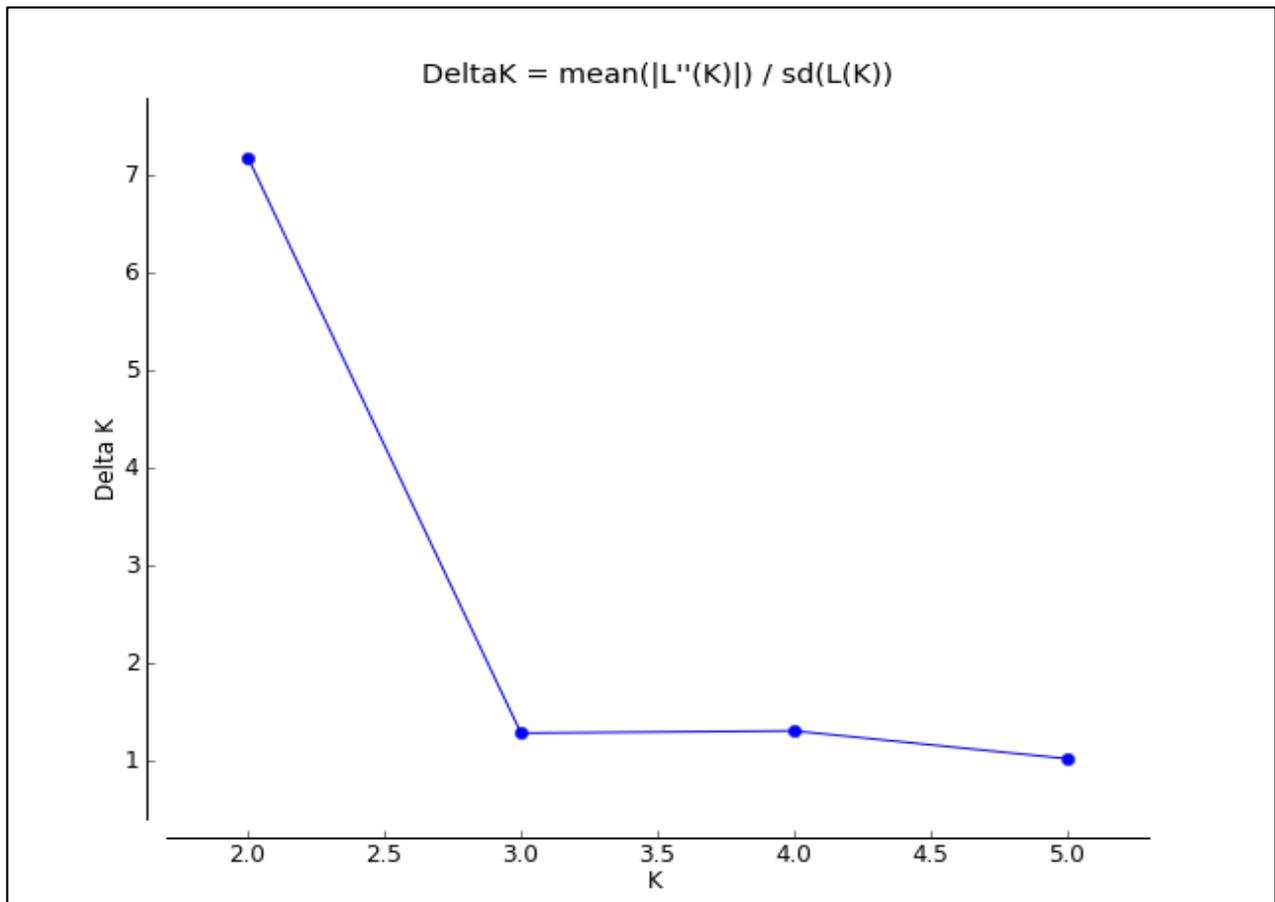


Figura 10: Valores de Delta K obtenidos a partir de Harvester Structure, calculados como la media de la probabilidad de K dividido por la desviación estándar de la probabilidad de K.

En la Figura 11, se ilustra el mapa de la población uruguaya para $K=1$, $K=2$ y $K=3$ y las proporciones de cada individuos para cada clúster.

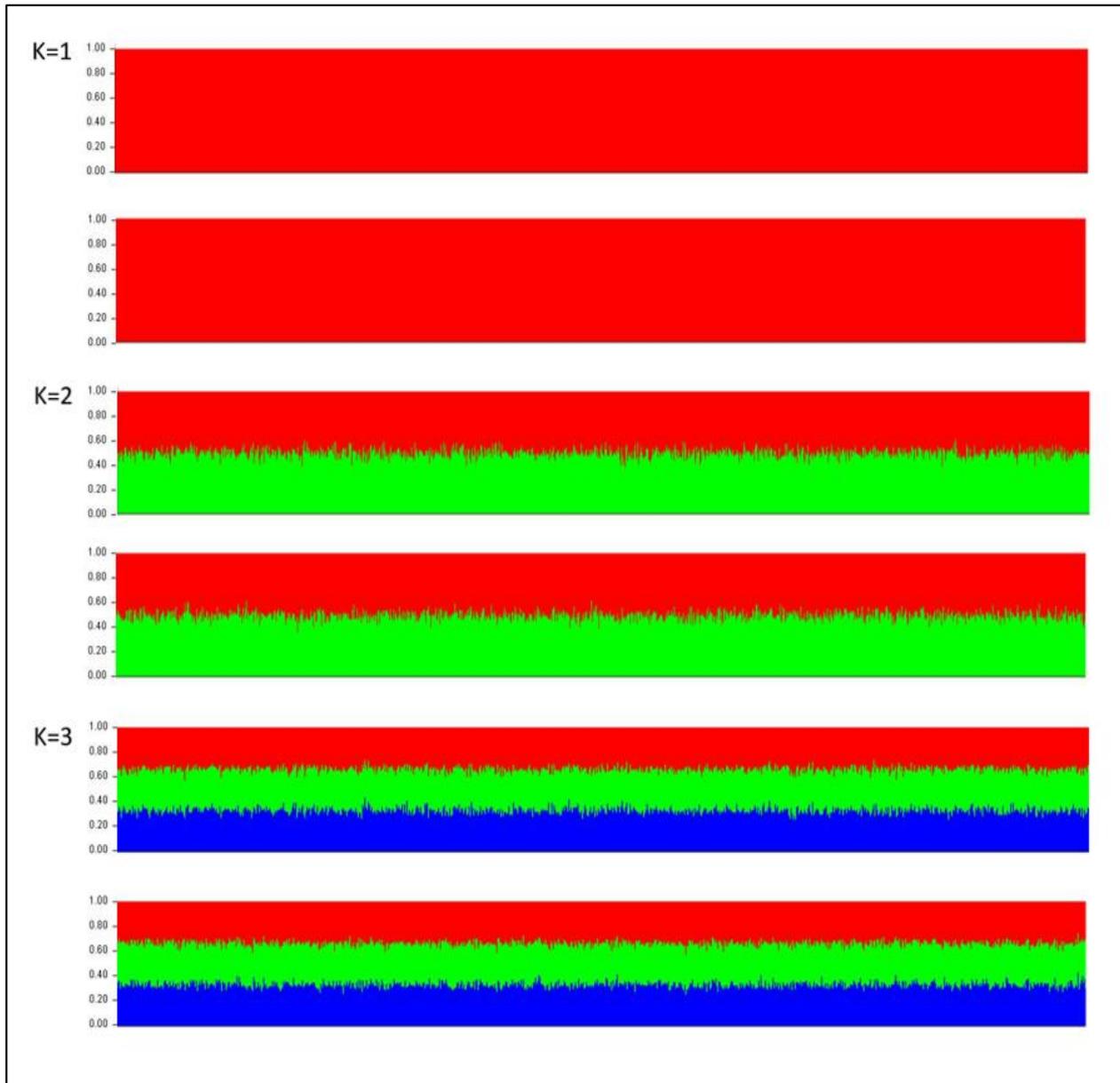


Figura 11: Gráfica que representa una simulación con $K=1$, $K=2$ y $K=3$ (Los diferentes colores representan los diferentes grupos genéticos o clúster). Para cada individuo de la población uruguaya se presenta una barra que grafica la probabilidad de éste de pertenecer al clúster rojo, verde o azul.

Se efectuó una segunda corrida incluyendo a la población uruguaya los datos de las poblaciones parentales que se encuentran disponibles en CODIS (afroamericanos, asiáticos y caucásicos), empleando valores de K entre 1 y 6. En esta corrida, el número más probable de poblaciones o grupos genéticos (K) de acuerdo con el método de Evanno (Evanno et al., 2005), fue de 2 ($\Delta K=76.468302$), seguido por K=3 ($\Delta K= 39.191098$).

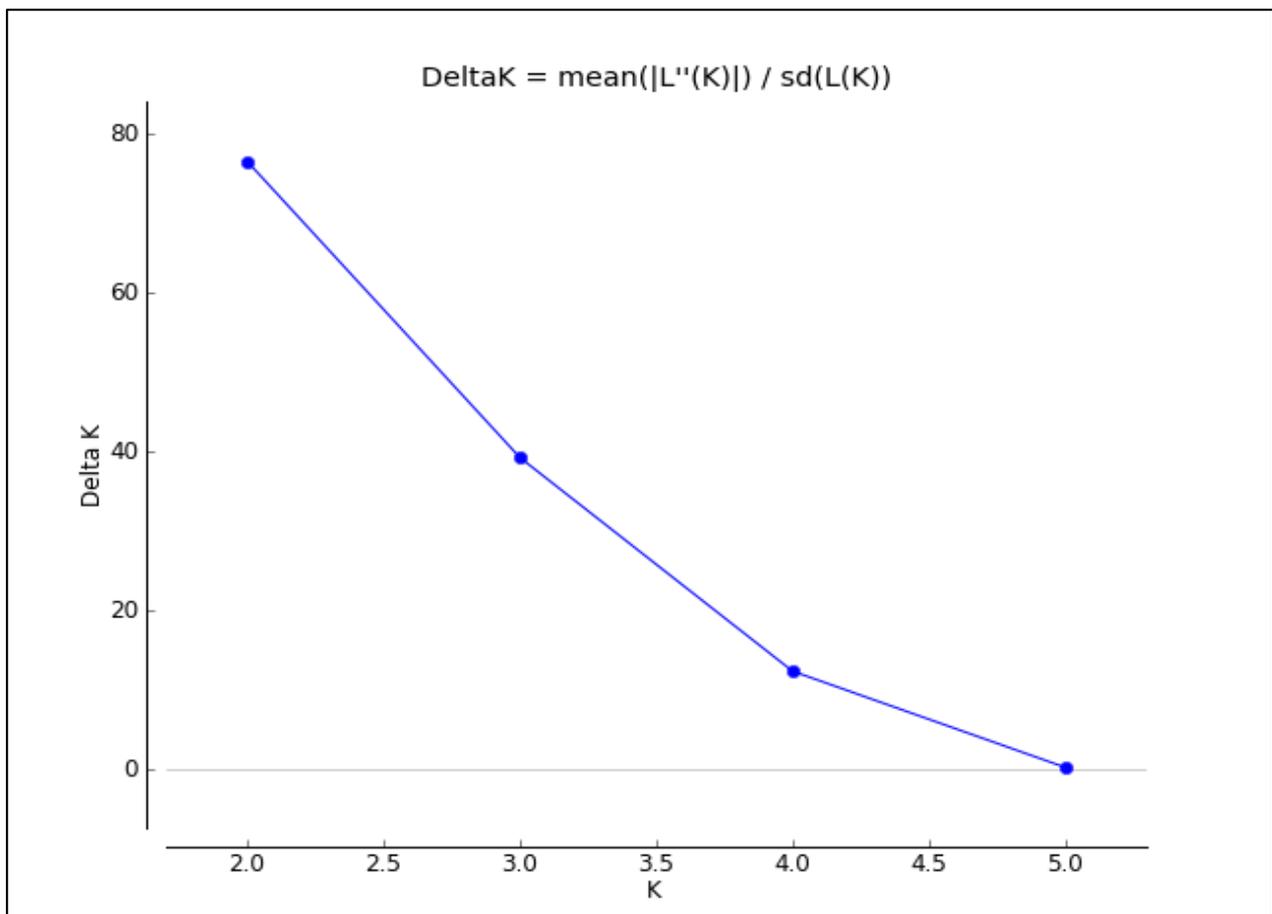


Figura 12: Valores de Delta K obtenidos a partir de Harvester Structure, calculados como la media de la probabilidad de K dividido por la desviación estándar de la probabilidad de K.

En la Figura 13, se ilustra el mapa de poblaciones analizadas para un valor de $K=2$ y las proporciones de individuos para cada clúster en las distintas poblaciones.

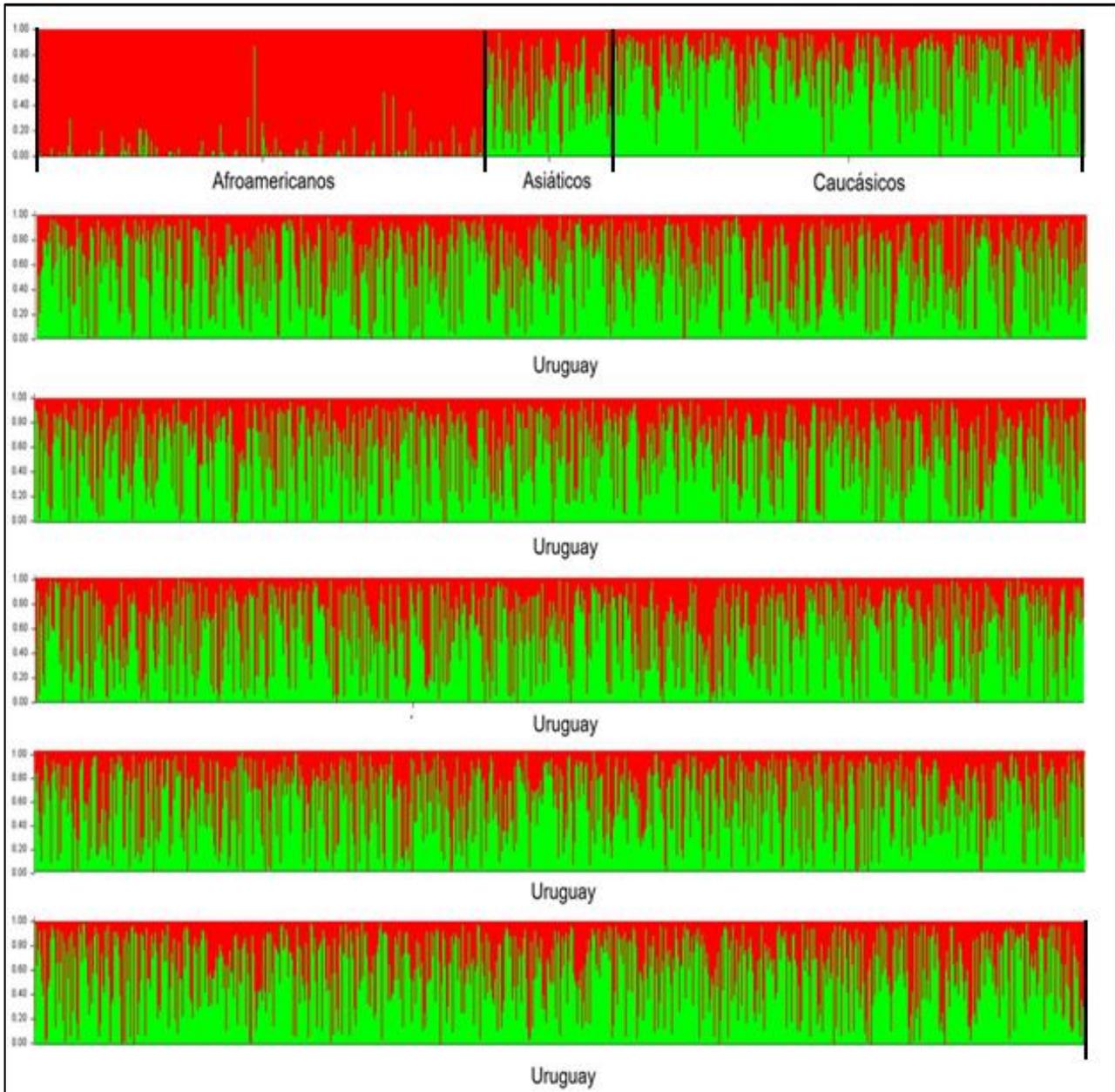


Figura 13: Gráfico que representa una simulación con $K=2$ (Los diferentes colores representan diferentes grupos genéticos). Para cada individuo se presenta una barra que grafica la probabilidad de éste de pertenecer al clúster rojo o verde. Los nombres en la abscisa representan los individuos de cada población analizada; afroamericanos ($N=342$), asiáticos ($N=97$), caucásicos ($N=361$) (disponibles en CODIS) y la población uruguaya ($N=4000$).

En la Figura 14, se ilustra el mapa de poblaciones analizadas para K=3 y las proporciones de individuos para cada clúster en las distintas poblaciones.

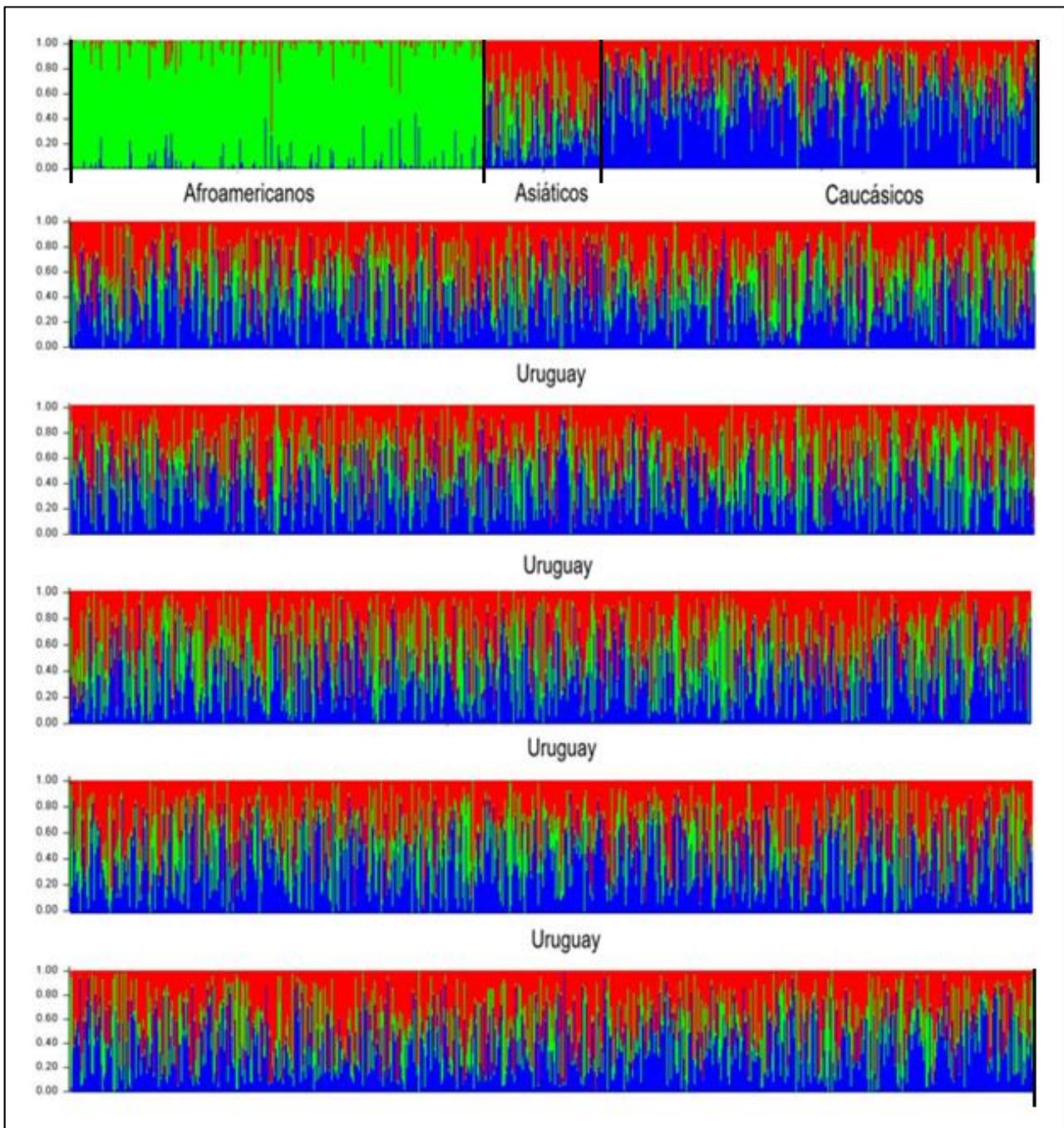


Figura 14: Gráfica que representa una simulación con K=3 (Los diferentes colores representan diferentes grupos genéticos). Para cada individuo se presenta una barra que grafica la probabilidad de éste de pertenecer al clúster rojo, verde o azul. Los nombres en la abscisa representan los individuos de cada población analizada; afroamericanos (N=342), asiáticos (N=97), caucásicos (N=361) (disponibles en CODIS) y la población uruguaya (N=4000).

Cuando asumimos $K=3$, los vectores de ascendencia se pueden graficar en un triángulo, donde cada individuo está representado por un punto de color y cada población parental está representada por un color. Para un punto dado, cada uno de los tres componentes viene dado por la distancia a un borde del triángulo. Por lo tanto, los individuos que se encuentran en una de las esquinas se asignan completamente a ese clúster.

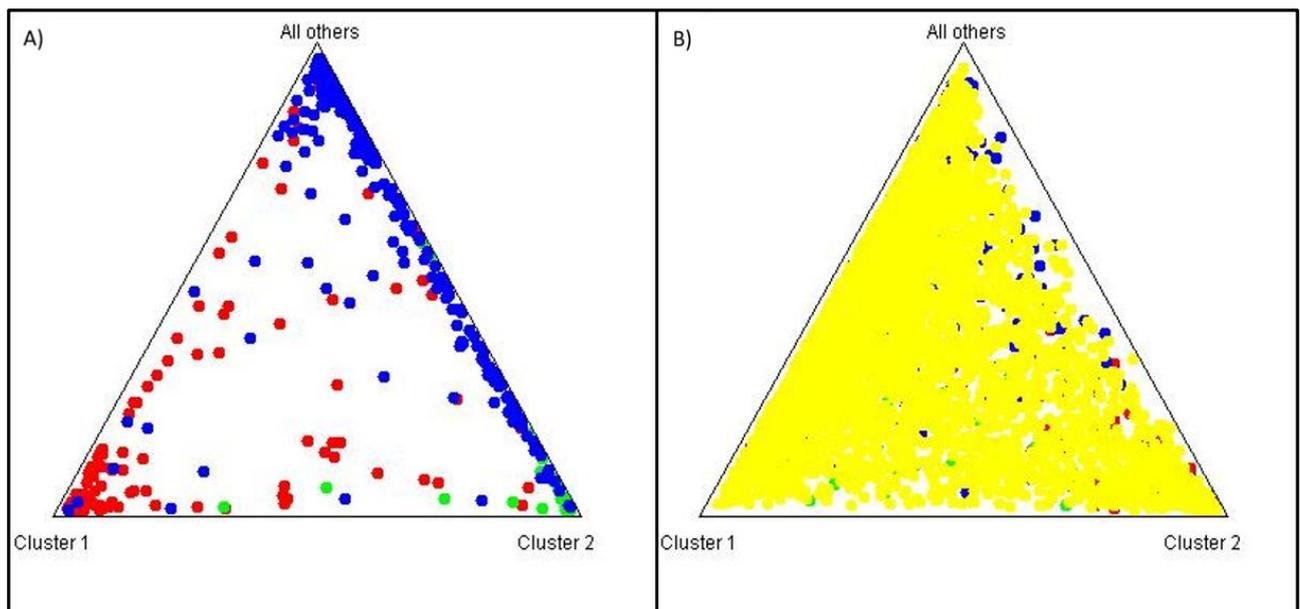


Figura 15: A) - Grafico de triangulo que representa un análisis en el Software STRUCTURE de 3 poblaciones parentales: afroamericanos (rojo), asiáticos (verde), caucásicos (azul). Los datos comprendían información para 23 marcadores STR autosómicos, en un total de 800 individuos. **B)** - Grafico de triangulo que representa un análisis en el Software STRUCTURE de 4 poblaciones: afroamericanos (rojo), asiáticos (verde), caucásicos (azul) y Uruguay (amarillo). Los datos comprendían información para 23 marcadores STR autosómicos, en un total de 4800 individuos.

En la Figura 16, se ilustran los resultados obtenidos del Software STRUCTURE, realizando corridas desde K=2 hasta K=5, donde se observa que en todos los casos se obtuvo una superposición de las diferentes muestras analizadas para cada valor de K.

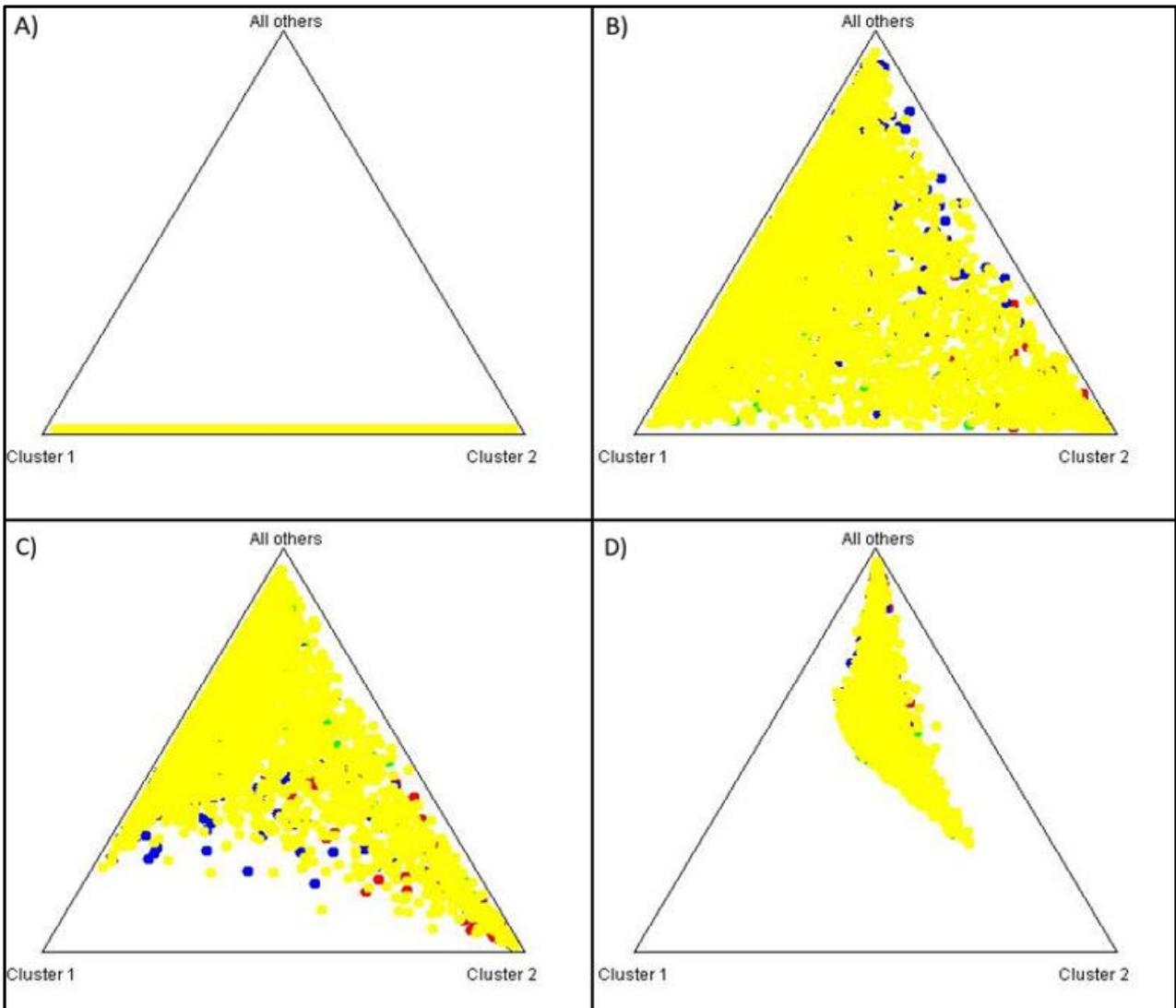


Figura 16: Resultados obtenidos del análisis del Software STRUCTURE de 4 poblaciones afroamericanos (rojo), asiáticos (verde), caucásicos (azul) y Uruguay (amarillo). **A)** K=2; **B)** K=3; **C)** K=4; **D)** K=5.

4.6 Análisis de Componentes Principales (PCA)

Se realizó un gráfico PCA (Análisis de Componentes Principales), basado en los datos de los alelos de cada uno de 23 marcadores STRs autosómicos, presentes en las muestras poblacionales disponibles en CODIS (afroamericanos, hispanos, caucásicos y asiáticos) y en la muestra de la población uruguaya.

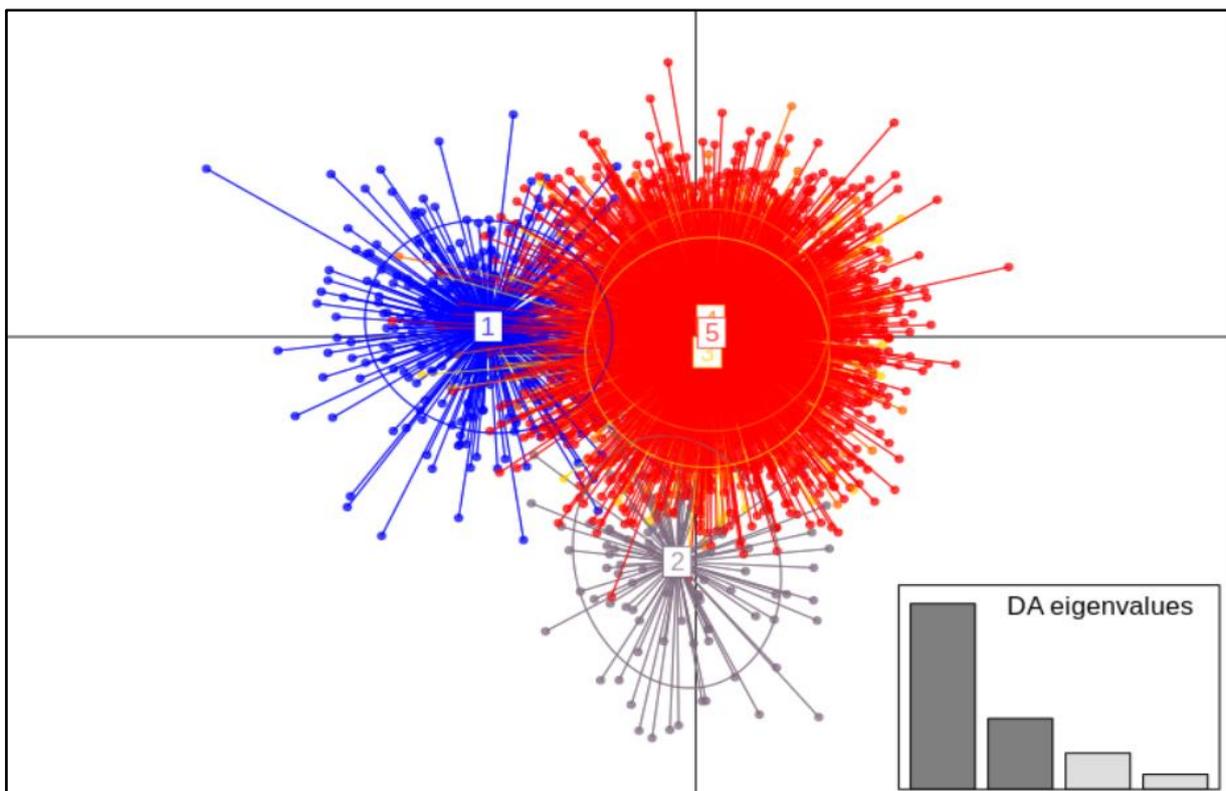


Figura 17: Ilustración de los 5 *clústers* en los que el PCA agrupó a los individuos de la población uruguaya (agrupación roja) y aquellos individuos de las poblaciones empleadas por CODIS (agrupación azul para afroamericanos, gris para asiáticos, amarilla para hispanos y naranja para caucásicos).

Se observa que las poblaciones o grupos se encuentran concentradas en 3 núcleos. Un primer grupo, donde se observa la población afroamericana, un segundo núcleo donde se puede observar a la población asiática y un tercer grupo, donde se observan superpuestas las poblaciones hispana, caucásica y la población uruguaya.

4.7 Resultados obtenidos con el método de identidad génica (ADMIX 95)

Los datos mostraron que en la muestra poblacional uruguaya, existe una alta contribución genética parental proveniente de europeos (0,6711) con un número menor de aporte africano y nativo americano, donde el aporte nativo americano (0,1621) presentó una mínima diferencia con la contribución parental de africanos (0,1668).

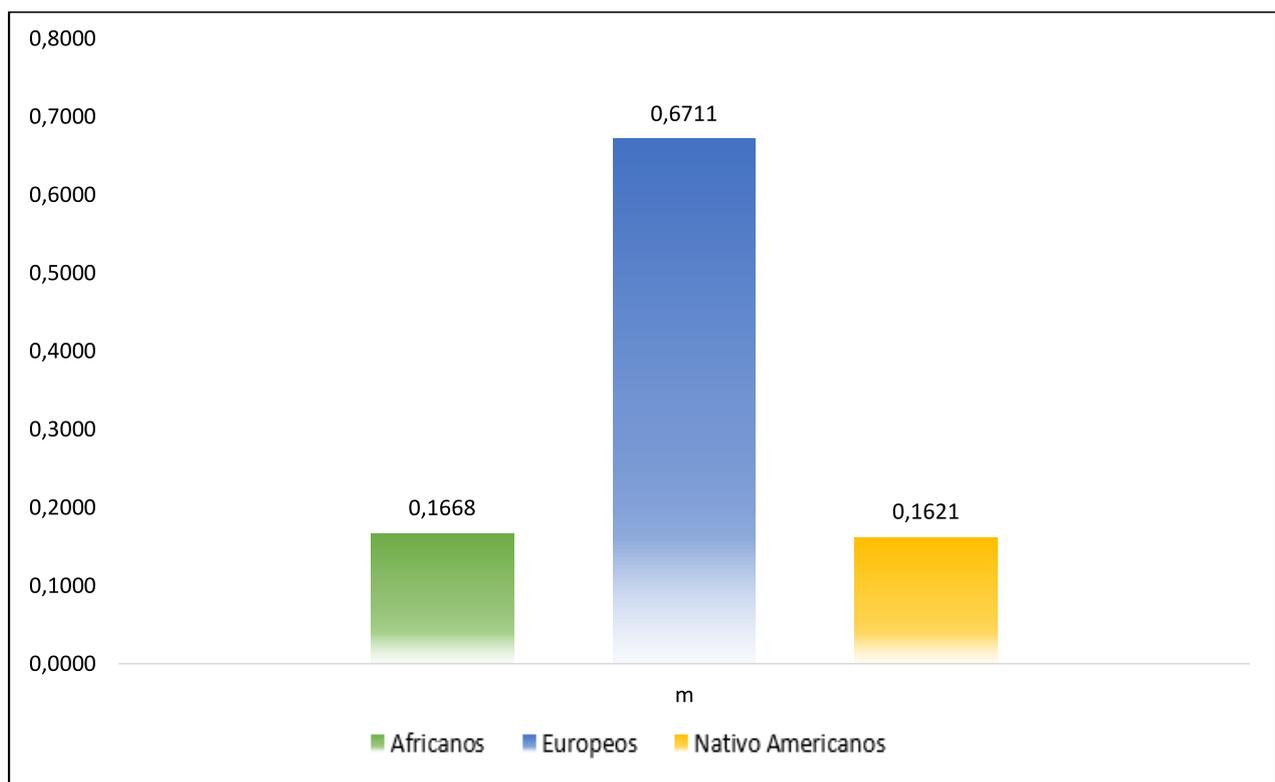


Figura 18: Representación gráfica de ADMIX 95. Gráficos de barras que representan las contribuciones parentales correspondientes al método ADMIX 95.

4.8 Resultados obtenidos con el programa VARMIX 3.0

Comparando los resultados con los obtenidos en el programa ADMIX 95, se puede decir que disminuye la contribución genética parental proveniente de europeos, siendo esta la mayor contribución a la muestra estudiada (0,6330), se observa un crecimiento en el aporte de africanos (0,2660), mientras que el aporte de nativo americanos disminuyó a 0,1010.

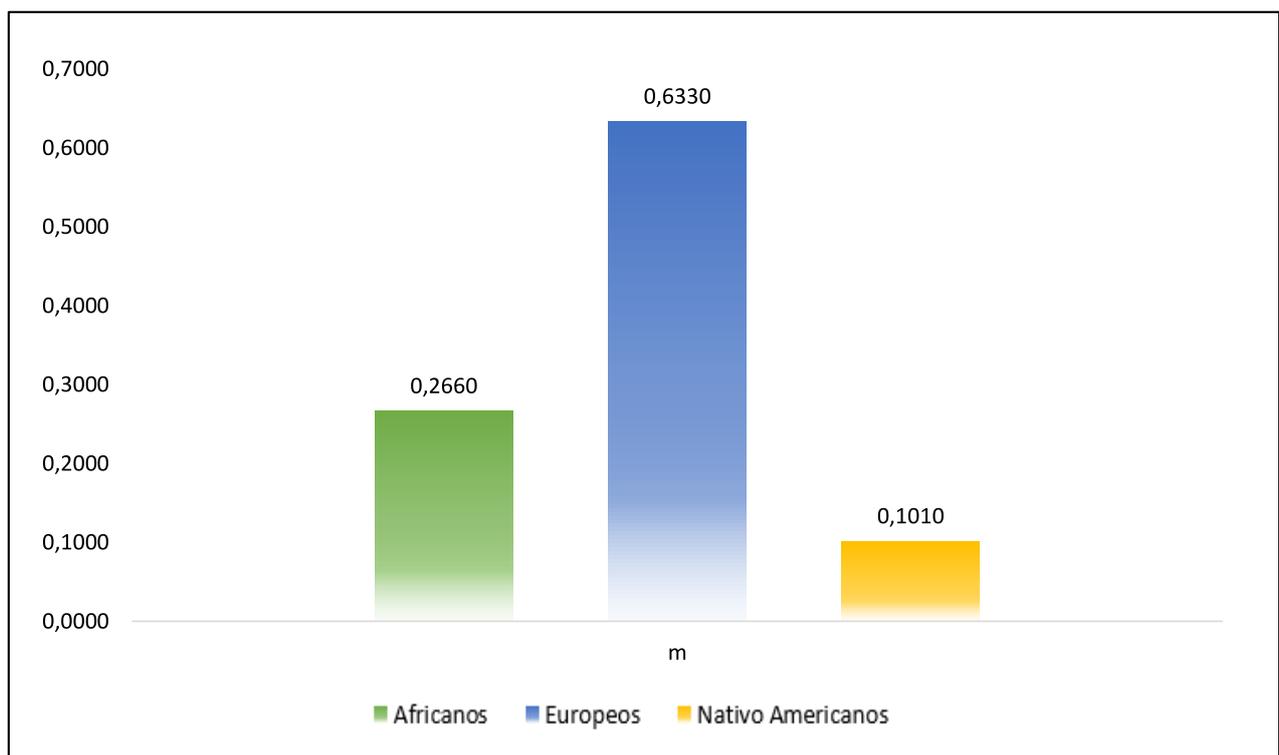


Figura 19: Representación gráfica de VARMIX 3.0. Gráficos de barras que representan las contribuciones parentales correspondientes al método VARMIX 3.0.

Si bien, se observan diferencias entre los resultados obtenidos por ambos software, se mantiene la tendencia en la contribución parental a la muestra analizada.

5. DISCUSIÓN

El análisis de los marcadores STRs autosómicos en grupos poblacionales ha permitido caracterizarlos y aportar científicamente en la práctica forense para la identificación genética de individuos (Butler, 2005). En los últimos años se han observado grandes avances de la genética forense, debido a que conjuntamente con el avance de la tecnología las técnicas se han vuelto más robustas, permitiendo amplificar perfiles genéticos a partir de evidencias con una cantidad mínima de ADN (Crespillo & Barrio, 2019). Los STRs se convirtieron en una herramienta básica para la genética forense, estos marcadores son altamente polimórficos y funcionan en PCR multiplex, aumentando el poder discriminativo entre las muestras si aumentamos el número de marcadores polimórficos (Lorente et al., 2007).

La implementación de perfiles genéticos en bases de datos cada vez son más estrictas y las comunidades científicas brindan determinados criterios para seleccionar los marcadores genéticos de acuerdo a los grupos poblaciones. A partir de enero de 2017, EE.UU amplió a 20 el número mínimo de marcadores genéticos, además de la Amelogenina que tipifica el sexo, empleados en el Software CODIS, y que deben incluirse en las bases de todo país, con el fin de reducir la probabilidad de coincidencias aleatorias (Moretti et al., 2016).

Existen otras comunidades científicas como por ejemplo: *La European Standard Set* donde los marcadores genéticos requeridos son: D22S1045, D21S11, D18S51, D12S391, D10S1248, D8S1179, D3S1358, D2S441, D1S1656, FGA, TH01, VWA, D2S1338, D16S539, D19S433, SEE33 y Amelogenina, el *UK Core Loci* (Reino Unido):

D10S1248, D12S391, D16S539, D18S51, D19S433, D1S1656, D21S11, D22S1045, D2S1338, D2S441, D3S1358, D8S1179, FGS, SEE33, TH01, VWA, HUMTH01, ACTBP2, D6S502, FIBRA, AMEL, AM y Amelogenina, *el German Core Loci*: FGA, TH01, SE33, VWA, D3S1358, D8S1179, D18S51, D21S11, Amelogenina y el *Interpol Standard Set* para el cual los marcadores requeridos son: FGA, TH01, VWA, D3S1358, D8S1179, D18S51, D21S11; y la Amelogenina (opcional) (<https://strbase.nist.gov/coreSTRs.htm>). Lo que se busca es reducir la probabilidad de coincidencias accidentales en la identificación de personas desaparecidas o inculpidas. Asimismo, se espera que en el futuro se amplíe el intercambio de perfiles genéticos entre países con el fin de identificar genéticamente a individuos que migran para cometer crímenes o desaparecen.

En este estudio, se analizaron los 20 marcadores genéticos requeridos para la utilización del Software CODIS, adicionando a su vez el marcador Penta D, Penta E y el SEE33, incluidos en el kit comercial de amplificación PowerPlex® Fusion 6C System (Promega Corporation) lo que hace un total de 23 marcadores STRs.

Algunos marcadores presentaron un número mayor de alelos, como fue el caso del marcador SEE33, el cual presentó 58 alelos, en tanto que el marcador TPOX presentó el menor número de alelos (9 alelos). Se demostró que los 23 STRs estudiados son buenos marcadores para caracterizar a la población uruguaya debido a que presentan un alto grado de polimorfismo: el marcador SE33 presentó el mayor índice de contenido de información polimórfica (PIC) con un 93,91%, seguido por el marcador Penta E con un 89,60%, el D18S51 (86,29%), el D1S1656 (86,12%), el D2S1338 (85,95%), el FGA (85,94%) y el D12S391 (85,77%), mientras que el marcador que presentó el menor índice

de información polimórfica es el TPOX con un 62%. Si consideramos que al sobrepasar el 50% un marcador genético es muy informativo para casos de identificación humana (Botstein et al., 1980), entonces los 23 marcadores moleculares estudiados se consideran útiles para su aplicación en genética forense por su alto contenido de información polimórfica y por ser fáciles de amplificar, siendo buenos discriminantes en la población uruguaya, ayudando así en la identificación genética de individuos y en casos de filiación.

El poder de discriminación (PD), se estudió para verificar la probabilidad de diferenciar una muestra entre las demás, se obtuvo un PD total mayor a 99,999999%, lo que nos indica que los marcadores moleculares estudiados son polimórficos y buenos discriminantes entre la población uruguaya para identificar a sus individuos. De igual manera, se estudió el poder de exclusión (PE), se obtuvo un PE total mayor a 99,999999%, este parámetro permite determinar los individuos que falsamente estén implicados en un peritaje forense. El marcador SEE33 presentó el PD y PE más alto (99,44% y 88,93% respectivamente), mientras que el TPOX con un 84,16% y 38,12% respectivamente, fue el marcador con el PD más bajo. Los 23 marcadores cumplen con los criterios para identificar genéticamente a personas involucradas en un proceso judicial, discriminando entre los implicados y proporcionando resultados confiables, teniendo la convicción que el identificado es quien aportó la muestra tomada de la escena del crimen.

En cuanto al índice de heterocigosidad, el marcador que exhibe el mayor índice de H_o para la población URU_4000 es el SE33 con un 94,55% y una H_e de 94,58%, mientras en la población CODIS_361 se observó una H_o de 95,01%. El valor promedio de H_o es de 80,93% y de H_e 81,16% en la población uruguaya, mientras que la H_o promedio para la muestras CODIS_361 fue de 80,60%. Los altos niveles de heterocigosidad indican que

en los STRs analizados existe una alta diversidad alélica, lo que implica un mayor número de genotipos posibles, aumentando así el poder de discriminación y disminuyendo la probabilidad de coincidencia de la muestra.

La heterocigosidad observada y esperada, el poder de discriminación y exclusión, son datos necesarios para evaluar el equilibrio de EHW, siendo uno de los objetivos de este estudio establecer frecuencias alélicas para la población uruguaya que puedan ser utilizadas en la práctica forense. De acuerdo con la ISFG, para que un marcador STR sea empleado en la práctica forense, debe encontrarse en equilibrio de EHW (Szibor et al., 2003). Al analizar el test de EHW se obtuvieron valores de probabilidad mayores a 0,05, observándose una desviación del equilibrio en tres marcadores genéticos (D8S1179, D12S391 y D13S317), sin embargo, para dos de estos marcadores (D12S391 y D13S317) esta desviación resultó no significativa tras aplicar la corrección de Bonferroni, obteniéndose valores de probabilidad mayores a 0,02 al dividir 0,05/23. Se determinó así que 22 de los 23 marcadores se encuentran en equilibrio de EHW y las frecuencias alélicas obtenidas son de gran utilidad en la práctica forense. Por otro lado, en la población CODIS_361 se había observado que 22 de los 23 marcadores analizados en el presente, se encuentran en EHW, excepto el marcador FGA que presentó una desviación del mismo.

Por lo tanto, los STRs estudiados en conjunto en la muestra poblacional URU_4000 resultaron en EHW excepto el marcador D8S1179. En el año 2000, Budowle y cols., detectaron una desviación significativa del EHW para este marcador (Budowle et al., 2000). Esta desviaciones del EHW, se puede atribuir a que este marcador presentó un H_o (0,8003) menor con respecto a la H_e (0,8141), lo que muestra una tendencia al déficit de heterocigotos observados (Chakraborty et al., 1992) en comparación con los demás

marcadores analizados. El déficit de heterocigotos observados podría ser producto de la presencia de alelos nulos (cuando uno de los alelos no son amplificados en la PCR debido a una desestabilización del cebador o *primers* durante la hibridación, es decir, un genotipo heterocigoto puede ser registrado erróneamente como un individuo homocigoto al no amplificar uno de sus dos alelos) (Pemberton et al., 1995). Leibelt y cols., (2003) demostraron que el kit de amplificación Power Plex 16 (Promega Corporation) producía perfiles D8S1179 discordantes en comparación con otro kit o juego de cebadores. Para recuperar esta pérdida alélica, emplearon cebadores inversos a los utilizados anteriormente, y se demostró que se recuperó completamente el alelo nulo (Leibelt et al., 2003). En el análisis forense, las muestras de referencia y evidencia (se origina a partir del mismo individuo que la muestra de referencia) normalmente se amplifican con el mismo juego de cebadores, pero con la utilización de conjuntos de cebadores diferentes los resultados podrían mostrar discordancias. En casos de filiación, la no amplificación de un alelo podría producir una falsa exclusión por lo que se debe tener precaución en la interpretación de un desajuste de un solo locus porque se sabe que pueden ocurrir mutaciones y abandono de alelos (Leibelt et al., 2003). El D8S1179 es uno de los 20 marcadores STRs requeridos por CODIS, por lo cual se encuentra incluido en los kits comerciales de amplificación, los cuales a su vez deben cumplir con estrictos procedimientos de validación antes de su empleo en genética forense.

Se confirmó un alto polimorfismo, alto poder de discriminación y exclusión (>99,999999%) y alta heterocigosidad, resultando en buenos marcadores para la identificación genética de personas a través del análisis de vestigios biológicos recogidos de la escena del crimen, restos óseos o cadavéricos, investigaciones de paternidad y en estudios de variabilidad genética poblacional.

En Latinoamérica, varios países se encuentran utilizando el Software CODIS, como es el caso de Brasil, donde su implementación se inició en 2010 (Grazinoli & Leal, 2015). En el año 2013 se informaron las frecuencias alélicas de 15 loci STR que incluían los 13 loci centrales requeridos por el FBI para la utilización del Software CODIS, en una muestra de más de 100.000 individuos brasileños (Aguiar et al., 2012), posteriormente Aguiar y cols., (2014), informaron las frecuencias alélicas de 20 loci STR (D10S1248, D12S391, D16S1656, D18S51, D19S433, D1S1656, D21S11, D22S1045, D2S1338, D2S441, D3S1358, D8S1179, FGA, TH01, VWA, CSF1P0, D13S317, D5S818, D7S820 y TPOX), estimadas en una muestra de 96.400 individuos (BR_96400). Al observar los parámetros estadísticos de cada marcador y realizar una comparación con los parámetros obtenidos para URU_4000 se observó que en la población BR_96400, el marcador que presentó el mayor índice de contenido de información polimórfica fue el D1S1656 con un 97,87% (86,12% para URU_4000), mientras en la muestra uruguaya fue el SEE33 (este marcador no fue incluido por Aguiar y cols., (2014) ya que no se encuentra dentro de los marcadores requeridos para CODIS). El PD y PE total de la muestra BR_96400 fue mayor a 99,999999% al igual que para la muestra URU_4000. La Ho y He más alta dentro de la muestra BR_96400 se observó en el marcador D1S1656 con 88,89% y 89,35% respectivamente, mientras que en la población uruguaya se obtuvo una Ho de 90,06% y una He de 89,50% para el marcador D1S1656.

En Uruguay el Sistema CODIS se implementó en el año 2011 (Ley N°. 18.849), el cual es un programa informático que contiene un banco de datos nacional de perfiles de ADN de personas condenadas por la justicia competente, perfiles de ADN de muestras halladas en el lugar de los hechos, perfiles correspondientes a restos cadavéricos sin identificar y perfiles de ADN de familiares de personas ausentes. Debido a que en el

presente trabajo se obtuvieron suficientes datos y patrones estadísticos deseables (PIC, PD, PE, Ho, He, etc.) y comparables con otras bases de datos conocidas, podríamos emplear la base de datos de la población uruguaya al igual que Brasil emplea su base de datos (BR_96400) en el esclarecimiento de casos forenses.

Las frecuencias alélicas obtenidas (URU_4000), fueron comparadas con las bases de datos que utilizan al momento los laboratorios forenses (URU_200 y CODIS_361), a fin de detectar la existencia de diferencias estadísticas significativas entre las poblaciones para cada marcador analizado. Se observó que al ampliar considerablemente el número de individuos empleados; 4000 para la muestra URU_4000, contra 200 individuos en la muestra URU_200 o 361 en la muestra CODIS_361; la muestra URU_4000 mostró una mayor variabilidad alélica, presentando alelos que en las otras poblaciones no se observaron y viceversa. A su vez, esta comparación de muestras, permitió observar que existen diferencias de frecuencias alélicas entre las muestras, pero al observar a modo ilustrativo si las muestras presentan igualdad de varianzas, mediante un test de Bartlett se obtuvo un $p > 0,05$ en todos los marcadores lo que nos indicó que estas diferencias no son estadísticamente significativas. Es importante destacar, que este test se realizó con la finalidad de obtener un patrón de comportamiento de las muestras, ya que las mismas no se ajustan a una distribución normal.

Como se observó anteriormente, existen diferencias significativas entre grupos poblacionales, debido a que se puede observar que ciertos alelos son más frecuentes o ausentes en ciertas poblaciones; por ello es importante que en cada país o región de la población se tengan estudios de la estructura genética poblacional. En los últimos años se ha observado un aumento en los números de migrantes de diferentes procedencias

que ingresan al país, con sus familias o reproduciéndose en la población uruguaya (a la que migraron), contribuyendo así a modificar la estructura genética de la población.

En base a esto, quisimos tener una perspectiva del efecto de la migración sobre la estructura genética de la población uruguaya a futuro; por lo que se procedió a consultar los datos estadísticos con los que cuenta la Dirección Nacional de Migración – Ministerio del Interior, correspondientes a las Residencias Definitivas Concedidas por año (Ley N°. 18.250) entre el año 2013 y 2019, los mismos son de acceso público y se encuentran disponibles en la página de la Dirección Nacional de Migración (Figura 20) (<https://migracion.minterior.gub.uy/index.php/estadisticas>).

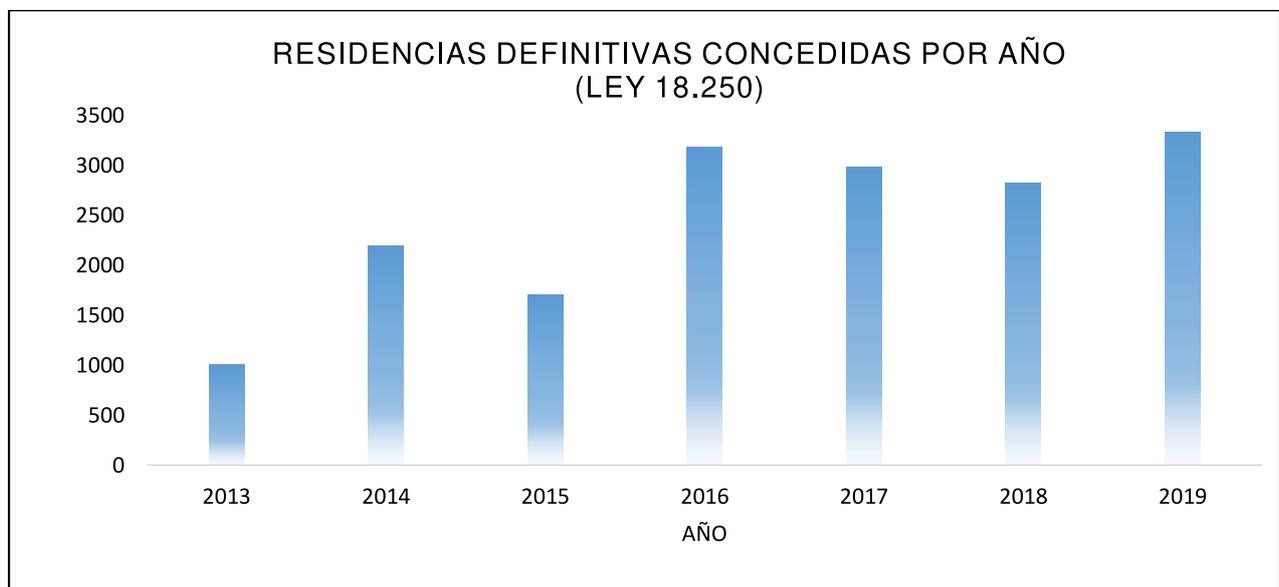


Figura 20: Número de Residencias Definitivas otorgadas por la Dirección Nacional de Inmigración entre el año 2013 y 2019.

Posteriormente, se realizó un análisis más profundo, a fin de evidenciar aquellas nacionalidades que presentaban mayor número de Residencias Definitivas Concedidas por año (Figura 21). Esta apreciación nos permite tener un panorama más real de la procedencia de los individuos migrantes.

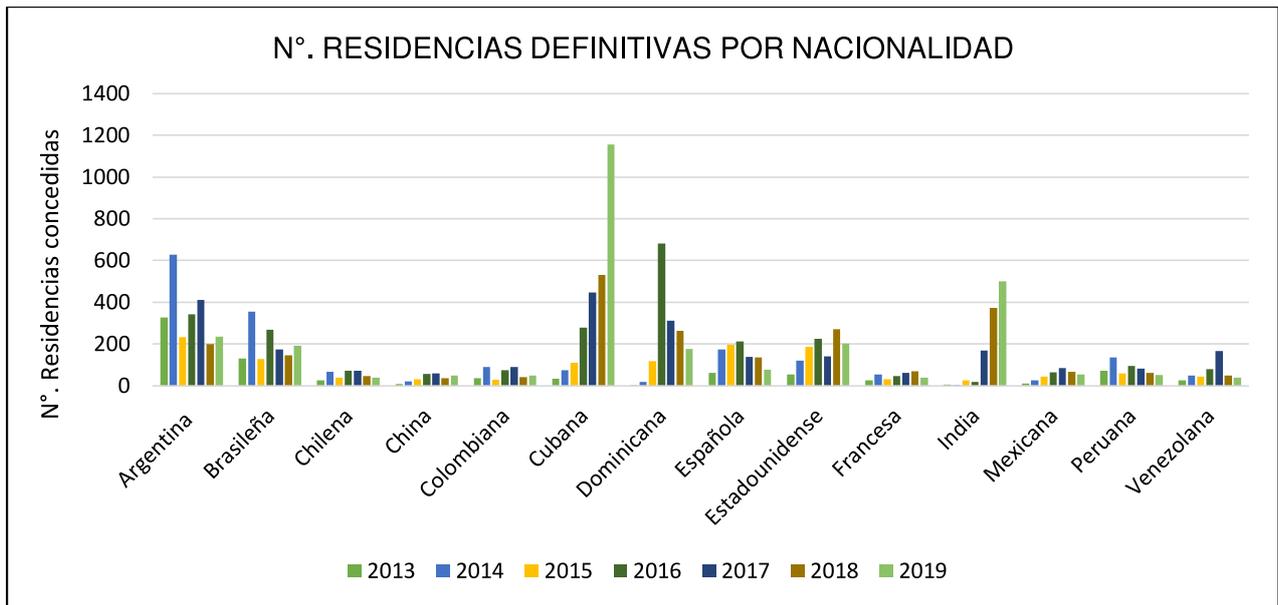


Figura 21: Numero de Residencias Definitivas por Nacionalidad otorgadas por la Dirección Nacional de Inmigración entre el año 2013 y 2019.

Se observó una alta migración de individuos procedentes de Argentina, Brasil, Cuba, República Dominicana, España, Estados Unidos e India entre los años 2013 y 2019, lo que implica con el transcurso de los años una modificación de la estructura genética poblacional uruguaya, debido a la contribución alélica e intercambio genético entre individuos extranjeros e individuos de la población uruguaya, lo que generara en la población a futuro, una variabilidad alélica diferente a la observada en el presente trabajo. Tal vez, lo más importante para la identificación genética de personas, es que los cálculos estadísticos podrían brindar resultados diferentes a los esperados si se considera y estudia a un individuo extranjero como si fuera uruguayo, en comparación a que si para dicho análisis se utilizara la base de datos del país o región de procedencia de dicho individuo.

Se estimó el grado de diferenciación genética entre la población uruguaya (URU_4000), poblaciones regionales (Argentina, Bolivia, Brasil, Chile y Paraguay), una población parental africana y una europea a partir del Índice de Fijación de Wright (F_{ST}) y

el Coeficiente de diferenciación genética (G_{ST}), estos estadísticos representan la diferenciación genética existente entre diferentes poblaciones.

Para el índice de fijación de Wright (F_{ST}), se obtuvieron valores muy cercanos a 0, lo que indica que existe muy poca diferenciación genética entre las poblaciones analizadas ($F_{ST} \leq 0,05$) (Wright, 1978). El mayor valor de F_{ST} se observó entre la población de Guinea Ecuatorial y Bolivia (0,0283). Por otro lado, la muestra poblacional uruguaya presentó mayor valor de F_{ST} (0,0050) con la población de Chile, evidenciándose así una mayor diferenciación genética entre estas dos poblaciones con respecto a las demás.

Al realizar el análisis de diferenciación genética de Nei o G_{ST} , también se obtuvieron valores cercanos a cero, lo que significa que la diversidad genética total (H_T) es igual a la heterocigosidad esperada (H_S) en cada población, es decir; las frecuencias alélicas son similares. El mayor valor de G_{ST} se observó entre Bolivia y Guinea Ecuatorial (0,0295). Con respecto a la población uruguaya, se pudo apreciar que el menor valor de G_{ST} lo presenta con la población de Mendoza y Santa Catarina (0,00202), seguido por San Pablo (0,0003) y Buenos Aires (0,0005), lo que indica que con estas poblaciones la población en estudio presenta mayor semejanza que con las demás. Mientras que el mayor valor de G_{ST} se da con la población de Bolivia (0,0168), evidenciando así que la población en análisis se diferencia mayormente de ésta.

Es importante resaltar, que tanto para F_{ST} como para G_{ST} , los valores obtenidos son cercanos a 0, y no se observaron valores tendientes a 1, por lo que las poblaciones incluidas en el estudio, no presentan una diferenciación poblacional significativa entre sí.

Existe un estimador de la diferenciación entre poblaciones, denominado R_{ST} , que resulta útil cuando las mutaciones han contribuido de manera sustancial a las diferencias alélicas que hay entre poblaciones, pero el análisis de este estadístico, no se incluyó en el presente, debido a que por su alta sensibilidad se desconoce si ofrece ventajas sobre F_{ST} (Hincapié et al., 2009). A su vez, el objetivo era observar el grado de diferenciación genética que presenta la población uruguaya con respecto a otras poblaciones regionales y dos poblaciones ancestrales, hecho que no descarta a futuro la utilización y análisis de este estimador.

A partir de los datos obtenidos para F_{ST} y G_{ST} , utilizando el algoritmo de Neighbor-Joining (NJ), se generaron árboles de distancia genética, a fin de observar las relaciones entre las muestras; observar que tan parecida es una población con las demás.

En el árbol de distancias genéticas obtenido a partir de los datos de F_{ST} , la distribución observada no se ajusta a lo que se conoce históricamente para las poblaciones de Latinoamérica (ver Figura 22), ya que lo esperado es que la población uruguaya se asemeje a poblaciones como la de Buenos Aires (Hidalgo et al., 2005).

Por otro lado, en el árbol de distancias genéticas obtenido a partir de los datos de G_{ST} , se pueden apreciar tres nodos; en el primero de ellos se observa a la población uruguaya la cual presentó una relación de similitud más cercana con las poblaciones de Santa Catarina, Buenos Aires y Mendoza, donde dicha similitud entre las mismas puede explicarse por la proporción de ancestros europeos que conforman las mismas. Esta similitud entre la muestra de Uruguay y Buenos Aires, ya había sido observada por Hidalgo y cols. (2005), lo cual concuerda con la historia demográfica de ambas regiones. A su vez

la población de Mendoza registró una estrecha relación con las poblaciones de Chile, Bolivia, Paraguay y Santa Fe, donde puede inferirse que tal comportamiento se debe a una contribución parental mayor de origen nativo americano de acuerdo a análisis previos (ver Figura 22). El segundo nodo, está conformado por las poblaciones de San Pablo, España, Rio de Janeiro y Guinea Ecuatorial, esta distribución puede explicarse porque Brasil es el país que registra los índices más altos de ancestralidad europea (78,6%) y afroamericana (9,6%) (ver Figura 22), en comparación con los demás países de la región. Por último, en un tercer nodo se observa a la población de Espirito Santo, la cual no presentó relación de ancestralidad con las demás poblaciones analizadas.

F_{ST} y G_{ST} se basan en las frecuencias alélicas, pero se puede deducir que uno funciona mejor que otro para detectar la diferenciación entre las poblaciones, siendo G_{ST} quien brinda resultados más coherentes de acuerdo con la historia microevolutiva latinoamericana.

El propósito de este análisis fue observar el comportamiento de la muestra poblacional uruguaya con respecto a las poblaciones consideradas, observar sus semejanzas, afinidades y diferencias a partir de las distancias genéticas. Por tal motivo no se realizaron otros análisis, pero sería interesante proponer otros modelos de agrupamientos y probarlos.

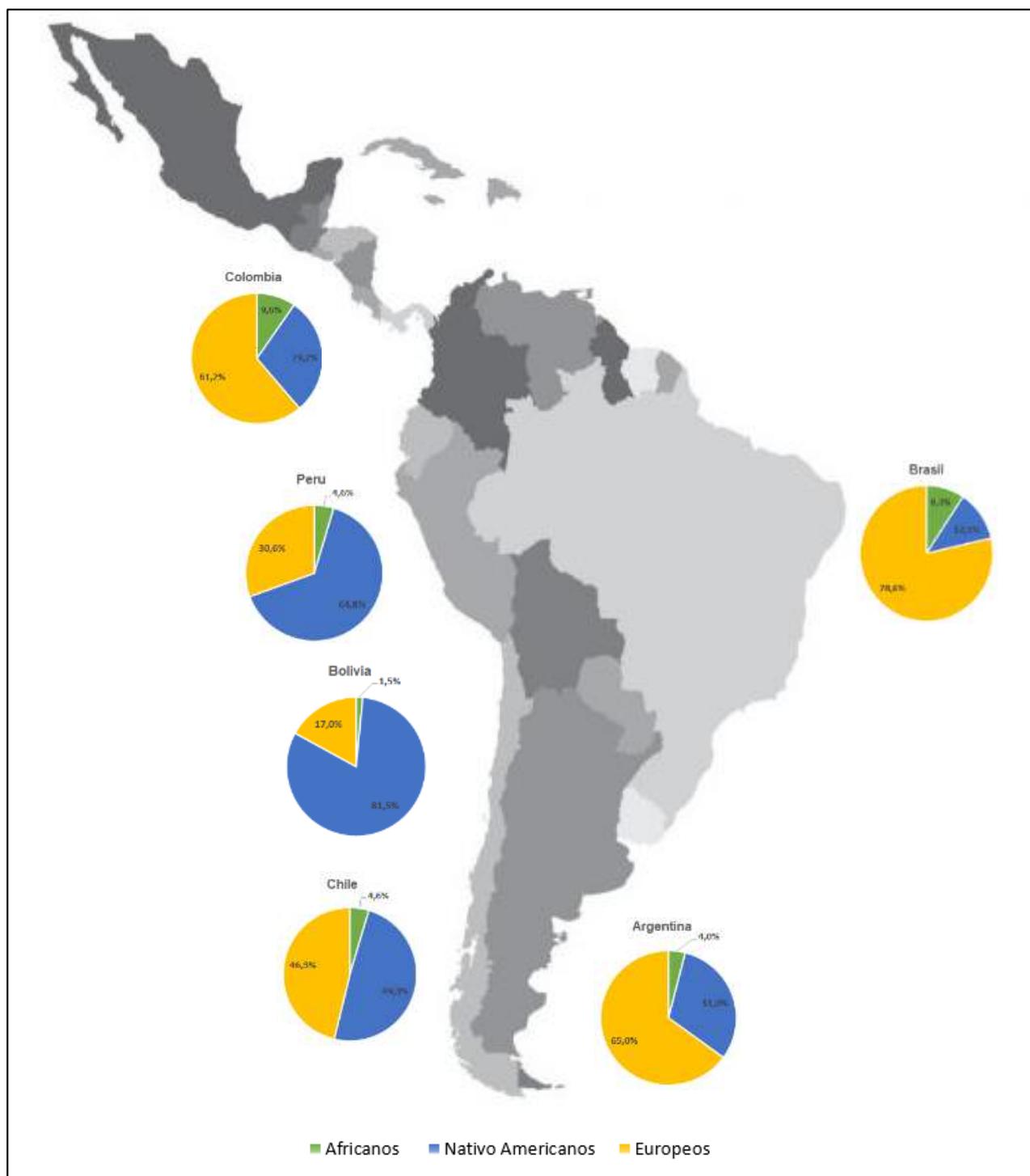


Figura 22: Proporción de ascendencia afroamericana, nativo americana y europea estimada para 6 poblaciones latinoamericanas (Brasil, Chile, Colombia y Perú (modificado de Adhikari et al, 2016), Bolivia (modificado de Heinz et al, 2013), Argentina (modificado de Avena et al, 2012)).

Los resultados obtenidos en el software STRUCTURE, fueron analizados en STRUCTURE HARVESTER a fin de determinar los valores óptimos de K definitivos, se observó que el número más probable de *cluster* obtenido fue de dos (K=2), pero debido a que STRUCTURE HARVESTER usa el estadístico ΔK de Evanno, y éste no muestra las estimaciones para la primera (K=1) y la última K (K=6) que se consideraron, otra opción a modo de evaluar el valor de K consiste en observar los valores de probabilidad obtenidos en STRUCTURE (Pritchard et al., 2000), donde el valor más alto se registró para K=1 ($\ln P(D)=-341776,2$), por lo tanto es correcto afirmar que no se observa subestructuración en la muestra poblacional uruguaya. El mismo análisis fue realizado pero incluyendo los datos poblacionales de las poblaciones parentales disponibles en CODIS (afroamericanos, asiáticos, caucásicos). Luego del análisis en STRUCTURE HARVESTER, se observó que el número más probable de clúster es de dos (K=2), y el valor más alto de probabilidad obtenido a través del Software STRUCTURE fue para K=2 ($\ln P(D)=-412139,4$) observándose diferencias de pertenencia marcadas para la población afroamericana, donde la mayoría de los individuos presentan el mayor porcentaje de pertenencia con el clúster rojo. En la población asiática, caucásica y uruguaya el porcentaje de pertenencia al clúster rojo es similar al porcentaje de pertenencia del clúster verde.

Se conoce que la población uruguaya tiene una composición histórica y genéticamente trihíbrida (Sans et al., 1997; Hidalgo et al., 2005), con contribución parental europea, nativo americana y afroamericana. Esta contribución parental no fue posible observarla a partir del análisis de STRs autosómicos, ya que no se detectó la subestructuración esperada en la muestra, pero si nos permitió deducir que la muestra se encuentra mestizada, observándose la contribución de dos subgrupos cuando se incluyen

datos de poblaciones parentales. Una de las explicaciones para esto es la imposibilidad de contar con las parentales adecuadas y el empleo de microsátélites.

El segundo número más probable de clúster o grupos génicos obtenido a partir STRUCTURE HARVESTER fue de tres ($K=3$), observación que coincide con la historia demográfica de la población uruguaya. Esto nos permitió graficar los vectores de ascendencia en un triángulo, donde cada individuo está representado por un punto de color y cada población parental está representada por un color; para un punto dado, cada uno de los tres componentes viene dado por la distancia a un borde del triángulo; los individuos que se encuentran en una de las esquinas se asignan completamente a ese clúster. En la Figura 15A, se muestra una representación del análisis para $K=3$, de las poblaciones parentales más cercanas disponibles en CODIS. Se apreció que la población afroamericana, asiática y caucásica, están conformadas por individuos que pertenecen únicamente a cada una de esas poblaciones (individuos en los vértices del triángulo), y por individuos que dependiendo de su distancia a cada uno de los bordes, presentan en mayor o menor medida contribución de otras poblaciones. En la Figura 15B, se presenta el mismo análisis pero incorporando la población uruguaya, lo que permitió observar que la población uruguaya presenta una distribución prácticamente homogénea, con individuos que se asemejan más a un clúster e individuos que presentan mayormente componentes de otros.

Al realizar el PCA, de las poblaciones parentales disponibles en CODIS (afroamericanos, asiáticos, hispanos y caucásicos) y la población uruguaya, se observó también una distribución homogénea de la población uruguaya con respecto a dos de sus parentales. El PCA mostró que las poblaciones se encuentran concentradas en 3 grandes

núcleos o subgrupos, por un lado se observa a la población afroamericana y la población nativo americana alejadas del centro, evidenciando así la diferencia que presentan con las demás. Por otro lado, sobre el centro del gráfico se observan prácticamente superpuestas las poblaciones de hispanos y caucásicos con la población uruguaya, lo que obedece a la gran similitud existente entre ellas.

En el PCA se observó que existen individuos de la población uruguaya que se alejan de su centro y se los puede observar por ejemplo en el nodo de la población asiática o afroamericana. Buscando una explicación a esta observación, se considera que este comportamiento puede deberse a la existencia de inmigrantes o hijos de inmigrantes en la muestra poblacional uruguaya con ascendencia asiática o africana. A su vez, puede explicarse por la contribución ancestral desigual que se observó para cada individuo a través de diferentes estudios. Debido a que se desconoce la procedencia geográfica de los individuos pertenecientes al muestreo analizado, se puede inferir que aquellos desvíos observados en el PCA, podrían tratarse de individuos que tienen contribución afroamericana o asiática lo suficientemente importante como para colocarlos en esos grupos.

Dado que es más accesible obtener datos de frecuencias alélicas de distintas poblaciones a nivel mundial antes que los genotipos individuales, se realizó un análisis de mestizaje, a fin de conocer cuan relacionada está la población uruguaya con sus parentales. El análisis se realizó, con dos métodos distintos para poder así comparar entre ellos los resultados. El método ADMIX 95, de identidad génica (Chakraborti, 1985) y el método VARMIX 3.0 de varianzas alélicas (Bertoni, 2005). A diferencia del método ADMIX 95, que considera la relación de frecuencias alélicas entre la población híbrida y las

parentales, el programa VARMIX 3.0 estima los aportes genéticos de las poblaciones parentales a partir de la información presente en la varianza alélica de los microsatélites.

Sans y cols., (1997), a partir de marcadores clásicos mostraron que la población presentó una contribución europea del 65%, 20% nativo americanos y 15% africanos. Por otro lado, Bonilla y cols. (2004), a partir del análisis de ADN mitocondrial revelaron que la principal contribución a través de la línea materna es la nativo americana con un 62%, mientras que el análisis de cromosoma Y mostró una contribución europea casi exclusiva para la línea paterna y una contribución nativo americana entre 1,6% y 8,3% (Bertoni et al., 2005). Hidalgo y cols. (2005), analizaron las frecuencias alélicas de 5 marcadores autosómicos, en una muestra poblacional de individuos de ambos sexos y nacidos en varias regiones del país, donde observaron que las contribuciones ancestrales para la población uruguaya fue de 84,1% europea, 10,4% nativo americana y 5,6% africana. Recientemente, Vega-Requena y cols. (2020), analizaron 8 inserciones *Alu* en una muestra poblacional conformada por 58 individuos descendientes de vascos residentes en Trinidad, 55 individuos autodefinidos de ascendencia vasca de la ciudad de Montevideo y por último un subconjunto de 70 individuos de ascendencia europea pero sin ascendencia vasca conocida, los autores observaron que la contribución ancestral que predomina es la europea con un 63%, seguido por la contribución africana con un 30% y la contribución ancestral más pequeña se corresponde con el aporte indígena con un 7%. Sans y cols. (2020), realizaron un análisis de la contribución parental de marcadores de ADNmt sobre la muestra poblacional utilizada por Bonilla y cols. (2015), donde se observó que en la muestra de mujeres nacidas en Montevideo, los aportes fueron de 33,9% indígena americano, 7,4% africano y 58,2% europeo/mediterráneo, mientras que en la muestra de residentes fue 36,6%, 8,3% y 55,1% respectivamente. El análisis de ancestría

biparental estimada por AIMS, mostró que las mujeres nacidas en Montevideo tuvieron, en promedio, aportes de 12,6% indígena, 78,6% europeo, y 8,8% africano, mientras que para las residentes, los valores promedio fueron similares, 13,8% de aporte indígena, 76,8% de aporte europeo y 9,4% de aporte africano. En la Figura 23, se representan de forma ilustrativa, las contribuciones ancestrales observadas en los diferentes estudios antes mencionados y los resultados observados para la población uruguaya a través del análisis de las frecuencias génicas (ADMIX 95) y del análisis de la información genética de los microsatélites (VARMIX 3.0).

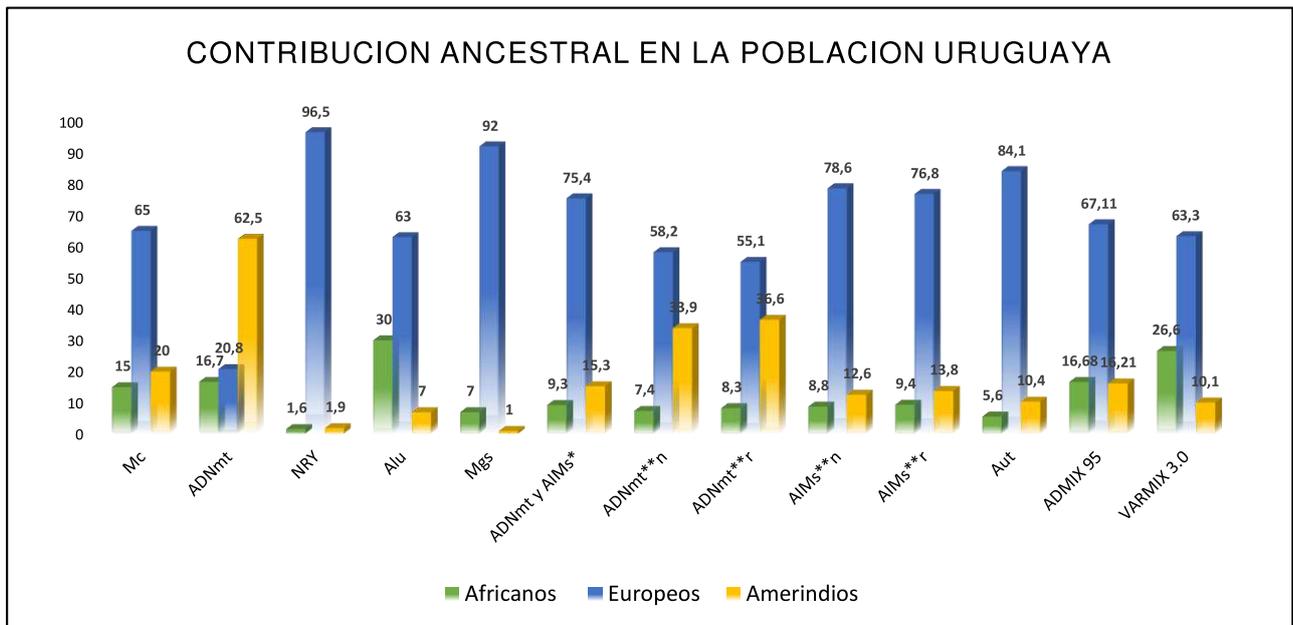


Figura 23: Representación gráfica de la contribución ancestral observada para la población uruguaya a partir del análisis de distintos marcadores (Mc: Marcadores clásicos (Sans, 1994), ADNmt: ADN mitocondrial (Bonilla et al, 2004), NRY: Región no recombinante del Cromosoma Y (Bertoni et al., 2005), Alu: Inserciones *Alu* (Vega-Requena et al, 2020), Mgs: Marcadores de grupo sanguíneo (Sans et al, 1997), ADNmt y AIMS*: ADN mitocondrial y marcadores individuales de ancestría (Bonilla et al, 2015), ADNmt**n/r: ADN mitocondrial nacidos/residentes (Sans et al, 2020), AIMS**n/r: marcadores individuales de ancestría (Sans et al, 2020), Aut: Cromosoma Autosómico (Hidalgo et al, 2005), ADMIX 95: Análisis de las frecuencias génicas de STRs autosómicos, VARMIX 3.0: Análisis de la información genética de los microsatélites).

El análisis del grado de mestizaje de la muestra poblacional uruguaya, permitió observar que los resultados obtenidos por ADMIX 95, presentan gran similitud con los observados para marcadores clásicos (Mc) con una contribución afroamericana promedio

de 15,8%, europea de 66% y nativo americana de 18%. Los resultados obtenidos a partir del método VARMIX 3.0, mostraron un alto grado de semejanza con los resultados obtenidos a partir del análisis de inserciones Alu (Vega-Requena et al, 2020), donde se observó un 63% de contribución parental europea mientras que para la población uruguaya de obtuvo un 63,3%, 7% de contribución nativo americana (10,1% para URU_4000) y un 30% de contribución afroamericana mientras que para URU_4000 esta contribución fue del 26,6%. Los resultados obtenidos para ADMIX 95 y VARMIX 3.0, presentan una diferencia significativa con respecto a la contribución afroamericana observada a partir del análisis de marcadores clásicos (Sans et al, 1997).

Excepto por los análisis realizados con marcadores de herencia uniparental; ADN mitocondrial (ADNmt) y la región no codificante del cromosoma Y (NRY), luego de comparar los resultados obtenidos con los ya publicados, se logró observar que si bien existen algunas diferencias, el análisis de marcadores STRs autosómicos también evidencia un comportamiento trihibrido de la población uruguaya. Observándose, para ambos métodos (ADMIX 95 y VARMIX 3.0), una contribución genética parental mayoritaria proveniente de europeos de 67,1% y 63,3%, con un aporte menor de africanos del 16,7% y 26,6% y de nativo americanos del 16,2% y 10,1%, respectivamente.

Los resultados obtenidos con el método de varianzas alélicas mostraron pequeñas variaciones con respecto al método de identidad génica para la población uruguaya, debido a que estos métodos utilizan aproximaciones diferentes. El método VARMIX 3.0 evidencia un aporte mayor africano, detectado a partir de la información molecular presente en la muestra. Esta observación puede explicarse porque mientras que en el

método clásico se analizan las diferencias de frecuencias alélicas, VARMIX 3.0 utiliza las varianzas alélicas de las muestras.

Los resultados del presente trabajo evidencian que la muestra poblacional analizada, podría ser utilizada en los análisis de identificación humana como población de referencia ya que poseen todas las características necesarias (un alto polimorfismo (PIC), alto poder de discriminación (PD) y exclusión (PE) (>99,999999%) y una alta heterocigosidad). La utilización de la base de datos de referencia de la población uruguaya, con fines de identificación, contribuirá a la obtención de resultados más exactos para establecer las probabilidades de identidad, parentesco y/o paternidad, poder de discriminación e índice de coincidencia de los diversos casos de índole de la Dirección Nacional de Policía Científica.

Sin embargo, resulta interesante pensar en explorar los posibles usos de estas bases de datos de referencia con técnicas modernas para poder determinar la búsqueda de parientes en genética forense y contribuir así con el esclarecimiento de casos complejos.

6. CONCLUSIONES

1. En la muestra poblacional uruguaya se observó que el marcador SEE33 es el que aporta mayor información para las pruebas de identificación genética debido a su alto índice de información polimórfica ($PIC=0,9391$), poder de discriminación ($PD=0,9944$), poder de exclusión ($PE=0,8893$), heterocigosidad ($H_o=0,9455$, $H_e=0,9458$), probabilidad de coincidencia ($PM=0,0056$) y variabilidad de alelos (58 alelos).
2. La heterocigosidad de los 23 marcadores STRs del kit PowerPlex® Fusion, fue de 81%. En combinación con su poder de discriminación (99,999999%), poder de exclusión (99,999999%) y su probabilidad de coincidencia (6,67%), los hacen buenos discriminantes por ser capaces de evitar errores de falsa inclusión, haciendo al panel una gran herramienta para el sistema de justicia.
3. Los STRs estudiados en la muestra poblacional uruguaya resultaron en EHW excepto el marcador D8S1179. Este desvió del EHW puede explicarse por un déficit de heterocigotos observados, producto de la presencia de alelos nulos. Debido a que este comportamiento fue observado en estudios anteriores, y a que actualmente este marcador es incluido en los kits comerciales de amplificación por ser uno de los marcadores requeridos por CODIS, se concluye que los 23 marcadores STRs en conjunto, son aptos para incluirse en las investigaciones forenses o de criminalidad donde estén implicados individuos de la población uruguaya.

4. Si bien se observó que la población uruguaya es semejante a la población URU_200 y CODIS_361, se evidenció la existencia de alelos propios y diferencias en las frecuencias alélicas con respecto a estas poblaciones.

5. Mediante el empleo de GST, se observó que la población uruguaya presentó poca diferenciación genética con las poblaciones regionales analizadas, observándose una mayor semejanza con las poblaciones de Santa Catarina, Buenos Aires y Mendoza, por lo que es posible emplear esta base de datos de referencia en aquellos casos donde se encuentren implicados individuos que no sean uruguayos, pero que si pertenezcan a la región.

6. No se observó subestructuración en la muestra poblacional analizada.

7. La muestra poblacional uruguaya presenta gran similitud con la muestra hispana y caucásica disponible en CODIS, por lo que el empleo de estas bases de datos, no traería diferencias significativas en los cálculos estadísticos.

8. En mestizaje calculado es concordante con lo observado en trabajos anteriores, se observó una contribución genética parental mayoritariamente proveniente de europeos (65%), con un aporte menor de africanos (21,6%) y nativo americanos (13%).

7. REFERENCIAS BIBLIOGRÁFICAS

- Adhikari, K., Mendoza-Revilla, J., Chacón-Duque, J. C., Fuentes-Guajardo, M., & Ruiz-Linares, A. (2016). Admixture in Latin America. *Current Opinion in Genetics and Development*, 41, 106–114. <https://doi.org/10.1016/j.gde.2016.09.003>
- Aguiar, V. R.C., De Vargas Wolfgramm, E., Malta, F. S. V., Bosque, A. G., Mafia, A. D. C., De Oliveira Almeida, V. C., De Andrade Caxito, F., Pardini, V. C., Ferreira, A. C. S., & Louro, I. D. (2012). Updated Brazilian STR allele frequency data using over 100,000 individuals: An analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, TPOX and vWA loci. *Forensic Science International: Genetics*, 6(4), 504–509. <https://doi.org/10.1016/j.fsigen.2011.07.005>
- Aguiar, V. R. C., De Castro, A. M., Almeida, V. C. O., Malta, F. S. V., Ferreira, A. C. S., & Louro, I. D. (2014). New CODIS core loci allele frequencies for 96,400 Brazilian individuals. *Forensic Science International: Genetics*, 13, e6–e12. <https://doi.org/10.1016/j.fsigen.2014.04.015>
- Alves, C., Gusmão, L., López-Parra, A. M., Mesa, M. S., Amorim, A., & Arroyo-Pardo, E. (2005). STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpFISTR Identifiler and Powerplex 16 kits. *Forensic Science International*, 148(2–3), 239–242. <https://doi.org/10.1016/j.forsciint.2004.05.007>

Avena, S., Via, M., Ziv, E., Pérez-Stable, E. J., Gignoux, C. R., Dejean, C., Huntsman, S., Torres-Mejía, G., Dutil, J., Matta, J. L., Beckman, K., Burchard, E. G., Parolin, M. L., Goicoechea, A., Acreche, N., Boquet, M., Ríos Part, M. D. C., Fernández, V., Rey, J., ... Fejerman, L. (2012). Heterogeneity in genetic admixture across different regions of argentina. *PLoS ONE*, 7(4). <https://doi.org/10.1371/journal.pone.0034695>

Bertoni, B. (2005). Caracterización y comportamiento de los haplotipos de cromosoma Y en las poblaciones humanas. PEDECIBA, Universidad de la Republica.

Bertoni, B., Velazquez, T., Sans, M., & Chakraborty, R. (2012). A Molecular Information Method to Estimate Population Admixture. In *Handbook of Statistics* (pp. 339-352). (Handbook of Statistics; Vol. 28). Elsevier B.V.. <https://doi.org/10.1016/B978-0-44-451875-0.00013-0>

Bonferroni CE (1936) Teoria statistica delle classi e calcolo delle probabilit `a. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8:3–62.

Bonilla, C., Bertoni, B., González, S., Cardoso, H., Brum-Zorrilla, N., & Sans, M. (2004). Substantial Native American female contribution to the population of Tacuarembó, Uruguay, reveals past episodes of sex-biased gene flow. *American Journal of Human Biology*, 16(3), 289–297. <https://doi.org/10.1002/ajhb.20025>

- Bonilla, C., Bertoni, B., Hidalgo, P. C., Artagaveytia, N., Ackermann, E., Barreto, I. Kittles, R.A. (2015). Breast cancer risk and genetic ancestry: a case-control study in Uruguay. *BMC Women's Health*, 15, 11. <http://doi:10.1186/s12905-015-0171-8>
- Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*, 32(3), 314–331
- Budowle, B., Defenbaugh, D. A., & Keys, K. M. (2000). Genetic variation at nine short tandem repeat loci in Chamorros and Filipinos from Guam. *Legal Medicine*, 2(1), 26-30.
- Butler, J.M. (2005). *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers* (2nd ed.). New York: *Elsevier Academic Press*. p. 688. ISBN 978-0121479527
- Butler, J. M. (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. *Journal of forensic sciences*, 51(2), 253-265.
- Camacho, M. V., Benito, C., & Figueiras, A. M. (2007). Allelic frequencies of the 15 STR loci included in the AmpFISTR Identifiler PCR Amplification Kit in an autochthonous sample from Spain. *Forensic science international*, 173(2-3), 241–245. <https://doi.org/10.1016/j.forsciint.2007.02.006>

Carracedo, A. (1999). "Valoración de la prueba del ADN". En Martínez Jarreta "la prueba del ADN en medicina forense: la genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad". España. *Masson* pág. 301 a 308.

Carracedo, A. (2013). ADN: La genética forense y sus aplicaciones en investigación criminal. Instituto de ciencias forenses. U. de Santiago de Compostela. [http://sgfm.elcorteingles.es/SGFM/FRA/recursos/doc/2013/PONENCIA S/Junio/1559347945_1062013102130.pdf](http://sgfm.elcorteingles.es/SGFM/FRA/recursos/doc/2013/PONENCIA_S/Junio/1559347945_1062013102130.pdf)

Cavalli-Sforza LL, Bodmer W.F y Walter F. (1981). "Genética de las poblaciones humanas". *Ediciones Omega*.

Chakraborty, R. (1985). Gene identity in racial hybrids and estimation of admixture rates. *Genetic microdifferentiation in man and other animals*. Neel, J.V. y Ahuja, Y. Delhi, Indian: *Anthropological Association*. Delhi University, Anthropology Department: 171-180.

Chakraborty, R., De Andrade, M., Daiger, S. P., & Budowle, B. (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of human genetics*, 56(1), 45–57. <https://doi.org/10.1111/j.1469-1809.1992.tb01128.x>

Clark, L. V., & Jasieniuk, M. (2011). polysat: An R package for polyploid microsatellite analysis. *Molecular Ecology Resources*, 11(3), 562–566. <https://doi.org/10.1111/j.1755-0998.2011.02985.x>

CODIS and NDIS Fact Sheet — FBI. (n.d.). Retrieved August 9, 2021, from <https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet>.

Core STR Loci Used in Human Identity Testing. <https://strbase.nist.gov/coreSTRs.htm>
Accessed: 2021-10-27

Crespillo Márquez, M. C., & Barrio Caballero, P. (2019). Genética forense: del laboratorio a los tribunales. Ediciones Díaz de Santos.

Crossetti, S. G., Demarchi, D. A., Raimann, P. E., Sai Zano, F. M., Hutz, M. H., & Callegari-Jacques, S. M. (2008). Autosomal STR genetic variability in the Gran Chaco native population: Homogeneity or heterogeneity? *American Journal of Human Biology*, 20(6), 704–711. <https://doi.org/10.1002/ajhb.20798>

Crow, J. F. and Kimura, M. (1970). "An introduction to population genetics theory". New York, *Evanston and London: Harper & Row, Publishers:83-98.*

De Souza Góes, A. C., Da Silva, D. A., Fonseca Gil, É. H., Da Silva, M. T. D., Pereira, R. W., & De Carvalho, E. F. (2004). Allele frequencies data and statistic parameters for 16 STR loci - D19S433, D2S1338, CSF1PO, D16S539, D7S820, D21S11,

D18S51, D13S317, D5S818, FGA, Penta E, TH01, vWA, D8S1179, TPOX, D3S1358 - In the Rio de Janeiro population, Brazil. *Forensic Science International*, 140(1), 131–132. <https://doi.org/10.1016/j.forsciint.2003.11.023>

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi:10.1111/j.1365-294x.2005.02553.x>

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(3), 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>

Fernández, A., Luna, F., & Moral, P. (1999). “Variation of 17 RFLP markers on chromosome 13 in southeastern Spain”. *Anthropologischer Anzeiger, Bericht Über Die Biologisch Anthropologische Literatur*. 57(2), 97-103

Gascue, C., Mimbacas, A., Sans, M., Gallino, J.P., Bertoni, B., Hidalgo, P.C., ... Cardoso, H. (2005). Frequencies of the Four Major Amerindian mtDNA Haplogroups in the

Population of Montevideo, Uruguay. *Human Biology* 77(6), 873-878. [doi:10.1353/hub.2006.0015](https://doi.org/10.1353/hub.2006.0015).

Grazinoli Garrido, R., & Leal Rodrigues, E. (2015). O Banco de Perfis Genéticos Brasileiro Três Anos após a Lei nº 12.654. *O Banco de Perfis Genéticos Brasileiro Três Anos Após a Lei Nº 12.654*, 1(35), 94–107. <https://doi.org/10.1344/rbd2015.35.14284>

González Andrade, F. (2006). “Análisis Molecular de Variación de Polimorfismos STR autosómicos y de cromosoma Y en grupos étnicos de Ecuador con aplicación Médico Forense”. Zaragoza.

Guo, X., & Elston, R. C. (1999). Linkage information content of polymorphic genetic markers. *Human Heredity*, 49(2), 112–118. <https://doi.org/10.1159/000022855>

Hares, D. R. (2015). Selection and implementation of expanded CODIS core loci in the United States. *Forensic Science International: Genetics*, 17, 33–34. <https://doi.org/10.1016/j.fsigen.2015.03.006>

Heinz, T., Álvarez-Iglesias, V., Pardo-Seco, J., Taboada-Echalar, P., Gómez-Carballa, A., Torres-Balanza, A., Rocabado, O., Carracedo, Á., Vullo, C., & Salas, A. (2013). Ancestry analysis reveals a predominant Native American component with moderate European admixture in Bolivians. *Forensic Science International: Genetics*, 7(5), 537–542. <https://doi.org/10.1016/j.fsigen.2013.05.012>

- Hidalgo, P. C., Bengochea, M., Abilleira, D., Cabrera, A., y Alvarez, I. (2005) Genetic Admixture Estimate in the Uruguayan Population Based on the Loci LDLR, GYPA, HBGG, GC and D7S8. *International Journal of Human Genetics*, 5(3), 217-222. <https://doi.org/10.1080/09723757.2005.11885929>
- Hill, C. R., Duewer, D. L., Kline, M. C., Coble, M. D., & Butler, J. M. (2013). U.S. population data for 29 autosomal STR loci. *Forensic Science International: Genetics*, 7(3), e82–e83. <https://doi.org/10.1016/j.fsigen.2012.12.004>
- Hincapie, M.A., Gil, A.M., PICO, A. L., Gusmao, L., Rondon, F., Vargas, C., & Castillo, A. (2009). Análisis de la estructura genética en una muestra poblacional de Bucaramanga, Departamento de Santander * *Colombia Médica*. 40, 361–372.
- Jamieson A. (1994). The effectiveness of using co-dominant polymorphic allelic series for (1) checking pedigrees and (2) distinguishing full-sib pair members. *Animal genetics*, 25 Suppl 1, 37–44. <https://doi.org/10.1111/j.1365-2052.1994.tb00401.x>
- Jeffreys, A.J., Wilson, V. & Thein, S.L. (1985a). Hypervariable 'minisatellite' regions in human DNA. *Nature* 314, 67–73. <https://doi.org/10.1038/314067a0>
- Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985b). Individual-specific 'fingerprints' of human DNA. *Nature*, 316(6023), 76–79. <https://doi.org/10.1038/316076a0>
- Jobling, M. A., & Gill, P. (2004). Encoded evidence: DNA in forensic analysis. *Nature reviews. Genetics*, 5(10), 739–751. <https://doi.org/10.1038/nrg1455>

Jombart T, Ahmed I (2011). “adegenet 1.3-1: new tools for the analysis of genome-wide SNP data.” *Bioinformatics*. <https://doi/10.1093/bioinformatics/btr521>

Jones, D. A. (1972). Blood Samples: Probability of Discrimination. *Journal of the Forensic Science Society*, 12(2), 355–359. [https://doi.org/10.1016/S0015-7368\(72\)70695-7](https://doi.org/10.1016/S0015-7368(72)70695-7)

Krenke B, Tereba A, Anderson S, Buel E, Culhane S, Finis C, Christine S. Tomsey, Zachetti JM, Masibay A, Rabbach DR, Amriott EA, Sprecher CJ, (2002) Validation of a 16-Locus Fluorescent Multiplex System. *J Forensic Sci* 47: 773-785

Lagos, M., Helena Poggi, M., & Cecilia Mellado, S. (2011). Conceptos básicos sobre el estudio de paternidad. *Revista Medica de Chile*, 139(4), 542–547. <https://doi.org/10.4067/S0034-98872011000400019>

Lander ES, et al (2001) “Initial sequencing and analysis of the human genome”. *Nature* 409: 860-921.

Landsteiner, K. (1990) “Zur Kenntnis der antifermentativen, lytischen und agglutinoierenden Wirkungen des Blutserums und der Lymphe”. *Zentralblatt Bakteriologie* 27: 357-362.

Lee, H. S., Lee, J. W., Han, G. R., & Hwang, J. J. (1998). Statistical Parameters for Evaluating the Identification Power of STR Loci. *Korean Journal of Legal Medicine*, 22(2), 13–19.

- Leibelt, C., Budowle, B., Collins, P., Daoudi, Y., Moretti, T., Nunn, G., Reeder, D., & Roby, R. (2003). Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. *Forensic science international*, 133(3), 220–227. [https://doi.org/10.1016/s0379-0738\(03\)00035-5](https://doi.org/10.1016/s0379-0738(03)00035-5)
- Ley N°. 18.250. Centro de Información Oficial, 06 de enero de 2008. <https://www.impo.com.uy/bases/leyes/18250-2008/76>
- Ley N°. 18.849. Centro de Información Oficial, 02 de diciembre de 2011. <http://www.impo.com.uy/bases/leyes/18849-2011>
- Lins, A. M., Sprecher, C. J., Puers, C., & Schumm, J. W. (1996). Multiplex sets for the amplification of polymorphic short tandem repeat loci - Silver stain and fluorescence detection. *BioTechniques*, 20(5), 882–889. <https://doi.org/10.2144/96205rr01>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299. <https://doi.org/10.1093/bioinformatics/btr642>.
- Lorente, J.A., & Lorente, A.M. (1995). “El ADN en la identificación criminal y en la paternidad biológica”. España, Comares.
- Lorente, J.A., Budowle, B., Gangitano, D., Figueiredo, M., Jorquera, H., Melendez, E., Pagano, S., Entrala, C., Alvarez, J.C., Lorente, M., & Villanueva, E. (2000).

“Análisis de la distribución alélica de 15 STR en la población Uruguaya” III Simposio Latino-Americano de Identificación Humana, Curitiba, Brasil.

Lorente, J. A. (2004). Un detective llamado ADN: tras las huellas de criminales desaparecidos y personajes históricos. Ediciones Temas de Hoy.

Lorente, J.A., Navarrete, I., Rosas, G., & Vega, M. (2007). “Genética forense, la Ciencia al Servicio de la Justicia”, disponible en: [www .criminalistica .net /forense/hemeroteca/adn/genetica-forense-al-servicio-de-la-justicia.html](http://www.criminalistica.net/forense/hemeroteca/adn/genetica-forense-al-servicio-de-la-justicia.html), http://www.cstl.nist.gov/strbase/str_fact.htm.

López-Flores, I., & Garrido-Ramos, M. A. (2012). The repetitive DNA content of eukaryotic genomes. *Genome dynamics*, 7, 1–28. <https://doi.org/10.1159/000337118>

Marino, M., Sala, A., & Corach, D. (2006). Population genetic analysis of 15 autosomal STRs loci in the central region of Argentina. *Forensic Science International*, 161(1), 72–77. <https://doi.org/10.1016/j.forsciint.2005.08.014>.

Martins, J. A., De Freitas Figueiredo, R., Yoshizaki, C. S., Paneto, G. G., & Barretto Cicarelli, R. M. (2011). Genetic data of 15 autosomal STR loci: An analysis of the Araraquara population colonization (São Paulo, Brazil). *Molecular Biology Reports*, 38(8), 5397–5403. <https://doi.org/10.1007/s11033-011-0693-6>.

Moretti, T. R., Moreno, L. I., Smerick, J. B., Pignone, M. L., Hizon, R., Buckleton, J. S., Bright, J. A., & Onorato, A. J. (2016). Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States. *Forensic Science International: Genetics*, 25, 175–181. <https://doi.org/10.1016/j.fsigen.2016.07.022>

Morling, N. (1998). Amplification of short tandem repeat loci using PCR. *Methods in Molecular Biology* (Clifton, N.J.), 98, 173–180. <https://doi.org/10.1385/0-89603-443-7:173>

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology*, 51 Pt 1, 263–273. <https://doi.org/10.1101/sqb.1986.051.01.032>

Nei M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70(12), 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321>

Nei, M. (1977). “F-Statistics and Gene Diversity in Subdivided Populations.” *Ann. Human Genetics* 41:221-233: 225–33.

Non-Coding DNA. (n.d.). Retrieved August 9, 2021, from <https://www.genome.gov/genetics-glossary/Non-Coding-DNA>

- Ocampos, M., Fernandes, R. C., Latorre, A. F. S., da Silva, C. M. D., Korndorfer, F. P., de Carvalho Giamarusti, A., & Menezes, M. E. (2009). 15 STR loci frequencies in the population from Santa Catarina, Southern Brazil. *Forensic Science International: Genetics*, 3(4), 129–131. <https://doi.org/10.1016/j.fsigen.2008.10.010>
- Pagano, S., Alvarez, J. C., Entrala, C., Lorente, J. A., Lorente, M., Budowle, B., & Villanueva, E. (2001). Uruguayan population data for eight STR loci (using the PowerPlex 1.2 kit). *Journal of forensic sciences*, 46(1), 178.
- Paredes, M., Crespillo, M., Luque, J. A., & Valverde, J. L. (2003). STR frequencies for the PowerPlex® 16 System Kit in a population from Northeast Spain. *Forensic Science International*, 135(1), 75–78. [https://doi.org/10.1016/S0379-0738\(03\)00172-5](https://doi.org/10.1016/S0379-0738(03)00172-5)
- Pellegrino, J. A., & Crespillo-Márquez, M. (2021). El Genoma Humano y el desarrollo de la Genética Forense. *Rev. Asoc. Méd. Argent*, 134(2), 21-5.4
- Pemberton, J. M., Slate, J., Bancroft, D. R., & Barrett, J. A. (1995). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular ecology*, 4(2), 249-252.
- Peña, J.A., Alfonso-Sánchez, M.A., (2016). ForensicMarkers: un programa para el cálculo de parámetros forenses. *Antropo*, 35, 79-82. www.didac.ehu.es/antropo

PowerPlex® Fusion 6C System. (n.d.). Retrieved August 9, 2021, from <https://worldwide.promega.com/products/forensic-dna-analysis-ce/str-amplification/powerplex-fusion-6c-system/?catNum=DC2705>

Porrás-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., & Lareu, M. V. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics*, 4(MAY), 1–13. <https://doi.org/10.3389/fgene.2013.00098>

Promega Corporation. (2018). *PowerPlex® Fusion 6C System Technical Manual (TMD045 Rev. 12/18)*. <https://worldwide.promega.com/-/media/files/resources/protocols/technical-manuals/tmd/tmd053-powerplex-fusion-6c-for-use-on-the-spectrum-compact-ce-system-technical-manual.pdf>

Quintero Ferrer, J. M., Pardo Govea, T. C., & Borjas Fuentes, L. B. (2020). Parámetros forenses de diez microsatélites del Cromosoma X en una muestra de la población del Estado Zulia, Venezuela. *Revista de Ciencias Forenses de Honduras*, 6(1), 3–13. <https://doi.org/10.5377/rcfh.v6i1.9936>

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org/>.

Rangel, H. (2010). La prueba de paternidad con ADN: una descripción para abogados, laboratorios clínicos y público en general. *Revista NOTICONAQUIC*. 18(49):40-51.

Rosado FF (2008) “Análisis de 15 loci tipo short tandem repeats (STR) en la población de Paraguay para su uso en identificación humana”. Tesis Doctoral, Universidad de Granada.

Quispe SE (2018) “Análisis de polimorfismos genéticos en la población de Bolivia, con fines de identificación y filiación forense” Universidad Mayor De San Andrés Universidad de Barcelona.

Saitou, N., & Nei, M. (1987) In *Molecular biology and evolution* (Vol. 4, Issue 4, pp. 406–425). <https://pubmed.ncbi.nlm.nih.gov/3447015/>

Sala, A., Argüelles, C. F., Marino, M. E., Bobillo, C., Fenocchio, A., & Corach, D. (2010). Genetic analysis of six communities of Mbyá-Guaraní inhabiting Northeastern Argentina by means of nuclear and mitochondrial polymorphic markers. *Human Biology*, 82(4), 433–456. <https://doi.org/10.3378/027.082.0406>

Sans, M. (1994). Estudio genético e histórico de la población del Departamento de Tacuarembó, Uruguay. (Tesis de Doctorado). PEDECIBA-Facultad de Ciencias, Montevideo.

Sans, M. (2000). Aplicaciones de las nuevas tecnologías al estudio de la población uruguaya. En: *Tecnología Genética*,:45-49, Instituto Goethe (ed.), Montevideo, Ediciones Trilce.

Sans, M. (2009). "Raza", adscripción étnica y genética en Uruguay. *RUNA, Archivo Para Las Ciencias Del Hombre*, 30(2), 163-174. <https://doi.org/10.34096/runa.v30i2.771>

Sans, M., Salzano, F. M., y Chakraborty, R. (1997). Historical genetics in Uruguay: estimates of biological origins and their problems. *Human Biology*, 69, 161-170.

Sans, M., Figueiro, G., Bonilla, C., Bertoni, B., Cappetta, M., Artagaveytia, N., Ackermann, E., Mut, P., & Hidalgo, P. C. (2020). Genetic ancestry and social stratification in montevideo, Uruguay. *Revista Argentina de Antropología Biológica*, 23(1), 1–19. <https://doi.org/10.24215/18536387E029>

Shriver, M. D., Smith, M. W., Jin, L., Marcini, A., Akey, J. M., Deka, R., & Ferrell, R. E. (1997). Ethnic-affiliation estimation by use of population-specific DNA markers. *American journal of human genetics*, 60(4), 957.

Smalldon, K.W., & Moffat, A. C. (1973). The Calculation of Discriminating Power for a Series of Correlated Attributes. *Journal of the Forensic Science Society*, 13(4), 291-295. [http://doi:10.1016/S0015-7368\(73\)70828-8](http://doi:10.1016/S0015-7368(73)70828-8)

- Silva, B. C., De Vargas Wolfgramm, E., Da Costa Aguiar, V. R., Varela Malta, F. S., De Castro, A. M., De Souza Ferreira, A. C., De Paula, F., & Louro, I. D. (2011). Genetic diversity and statistical parameters of 15 autosomal STR loci in the Pomeranian Subpopulation of Espirito Santo State, Brazil. *Molecular Biology Reports*, 38(5), 3013–3016. <https://doi.org/10.1007/s11033-010-9966-8>
- Stark, Alan E. (2005). The Hardy-Weinberg principle. *Genetics and Molecular Biology*, 28(3), 485. <https://doi.org/10.1590/S1415-47572005000300027>
- Szibor, R., Krawczak, M., Hering, S., Edelmann, J., Kuhlisch, E., & Krause, D. (2003). Use of X-linked markers for forensic purposes. *International Journal of Legal Medicine*, 117(2), 67–74. <https://doi.org/10.1007/s00414-002-0352-5>
- Toscanini, U., Moreno, F., Pantoja-Astudillo, J. A., Morales, E. A., Bustos, P., & Salas, A. (2015). A reference frequency database of 15 autosomal STRs in Chile. *Forensic science international. Genetics*, 19, 35–36. <https://doi.org/10.1016/j.fsigen.2015.05.016>
- Tillmar, A. (2010). Populations and Statistics in Forensic Genetics. In *Medicine* (Issue 1175).
- Trujillo Alvarez, Y., Arce Bustabad, S., Viguera, R., Motas, I. M., & Mediaceja, V. W. (2018). El complejo mayor de histocompatibilidad. Organización genética, estructura, localización y función. *Panorama Cuba y Salud, ISSN-e 1995-6797, Vol. 13, Nº. 1, 2018, Págs. 69-73, 13(1), 69–73.*

<https://dialnet.unirioja.es/servlet/articulo?codigo=7318076&info=resumen&idioma=SPA%0Ahttps://dialnet.unirioja.es/servlet/articulo?codigo=7318076>

Vega-Requena, Y. V., Hidalgo, P. C., Ackermann, E., Flores-Gutierrez, S., & Sans, M. (2020). Genetic Admixture Analysis in the Population of Tacuarembó-Uruguay Using Alu Insertions. *Human biology*, 91(4), 249–256.

<https://doi.org/10.13110/humanbiology.91.4.03>

Vergara, I. A., Villouta, P., Herrera, S., & Melo, F. (2012). Autosomal STR allele frequencies for the CODIS system from a large random population sample in Chile. *Forensic Science International: Genetics*, 6(3).

<https://doi.org/10.1016/j.fsigen.2011.07.002>

Verzeletti, A., Cortellini, V., Cerri, N., Bosetti, A., Pretto, S., & De Ferrari, F. (2013). PowerPlex® Fusion kit: A 23 plex autosomal STR kit for human identity testing. *Forensic Science International: Genetics Supplement Series*, 4(1), e344–e345.

<https://doi.org/10.1016/j.fsigss.2013.10.175>

Vullo, C., Rotondo, M., Daguerre, N., Quiroz, A., Velázquez, V., Machado, P., Serra, H., & Borosky, A. (2016). Genetic polymorphism of 22 autosomal STR markers in Paraguay. *Forensic Science International: Genetics*, 25, e16–e17.

<https://doi.org/10.1016/j.fsigen.2016.08.002>

Weber, J.L., & Wong, C. (1993). Mutation of human short tandem repeats. *Hum Mol Genet*; 2:1123-1128

Weir, B. S. and Clark Cockerham C. (2012). Estimating F-Statistics for the Analysis of Population Structure. *Society for the Study of Evolution*, 38(6), 1358–1370.
<http://www.jstor.org/stable/2408641>

Wright, S. (1951) The Genetical Structure of Populations. *Annals of Eugenics*, 15, 323-354. <http://dx.doi.org/10.1111/j.1469-1809.1949.tb02451.x>

Wright, S. (1965). The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating Author (s): Sewall Wright Published by: Society for the Study of Evolution Stable. *Journal of Evolution*, 19(3), 395–420.
<http://www.jstor.org/stable/2406450>

Wright, S. 1978. *Evolution and the Genetics of Natural Populations, Vol. 4: Variability within and among Natural Populations*. Chicago, USA: University of Chicago Press.

8. ANEXOS

Anexo 1: *Protocolo de Amplificación directa de ADN presente en tarjetas de almacenamiento.*

En placa óptica de reacción de 96 pocillos MicroAmp® (Applied Biosystems) se corta un disco de 1,2 mm aproximadamente de las muestras bucales presentes en los dispositivos de almacenamiento Bode Buccal DNA Collector™ (Bode Technology) mediante el empleo del Sistemas de Perforación BSD600 Duet puncher (BSD Robotics).

Posteriormente se tratan las muestras con 10 µL Punch Solution™ Kit (Promega Corporation) con el objetivo de lisar las muestras antes de agregar la mezcla de amplificación PCR.

Se determina el número de reacciones que se configurarán, incluyendo las reacciones de control positivo y negativo. Se realiza la mezcla de amplificación de la PCR agregando un volumen extra de 1 o 2 reacciones a fin de compensar el error de pipeteado, asegurando así contar con el volumen suficiente de mezcla y que cada reacción contenga la misma mezcla de amplificación de la PCR. En la tabla 6 se detallan los volúmenes requeridos de reactivos para cada reacción.

Tabla 6: Mezcla de amplificación de la PCR para amplificación directa de ADN de las tarjetas perforadas de almacenamiento.

Componente de la mezcla de amplificación de la PCR	Volumen por reacción	x	Número de reacciones	=	Volumen final
Agua, grado de amplificación	7,5 µl	x		=	
PowerPlex® Fusion 5 Master Mix	2,5 µl	x		=	
PowerPlex® Fusion 5 Primer Pair Mix	2,5 µl	x		=	
Volumen total de la reacción	12,5 µl				

Para el control positivo de la amplificación, se agregó 0,5 µl de ADN de control 2800M (10 ng) a un pocillo de reacción con 12,5 µl de mezcla de amplificación de la PCR. Además se reservó un pocillo como control negativo, el cual solo contiene la mezcla de amplificación de la PCR.

Luego se lleva la placa a termociclador. En la tabla 7 se detalla el protocolo del ciclo térmico.

Tabla 7: Protocolo del ciclo térmico.

Protocolo del ciclo térmico
96 °C durante 1 minuto, luego:
94 °C durante 10 segundos
59 °C durante 1 minutos
para 27 ciclos, luego:
72 °C durante 30 segundos
60 °C durante 20 minutos
inmersión a 4 °C

Anexo 3: Perfil genético de un individuo (ejemplo)

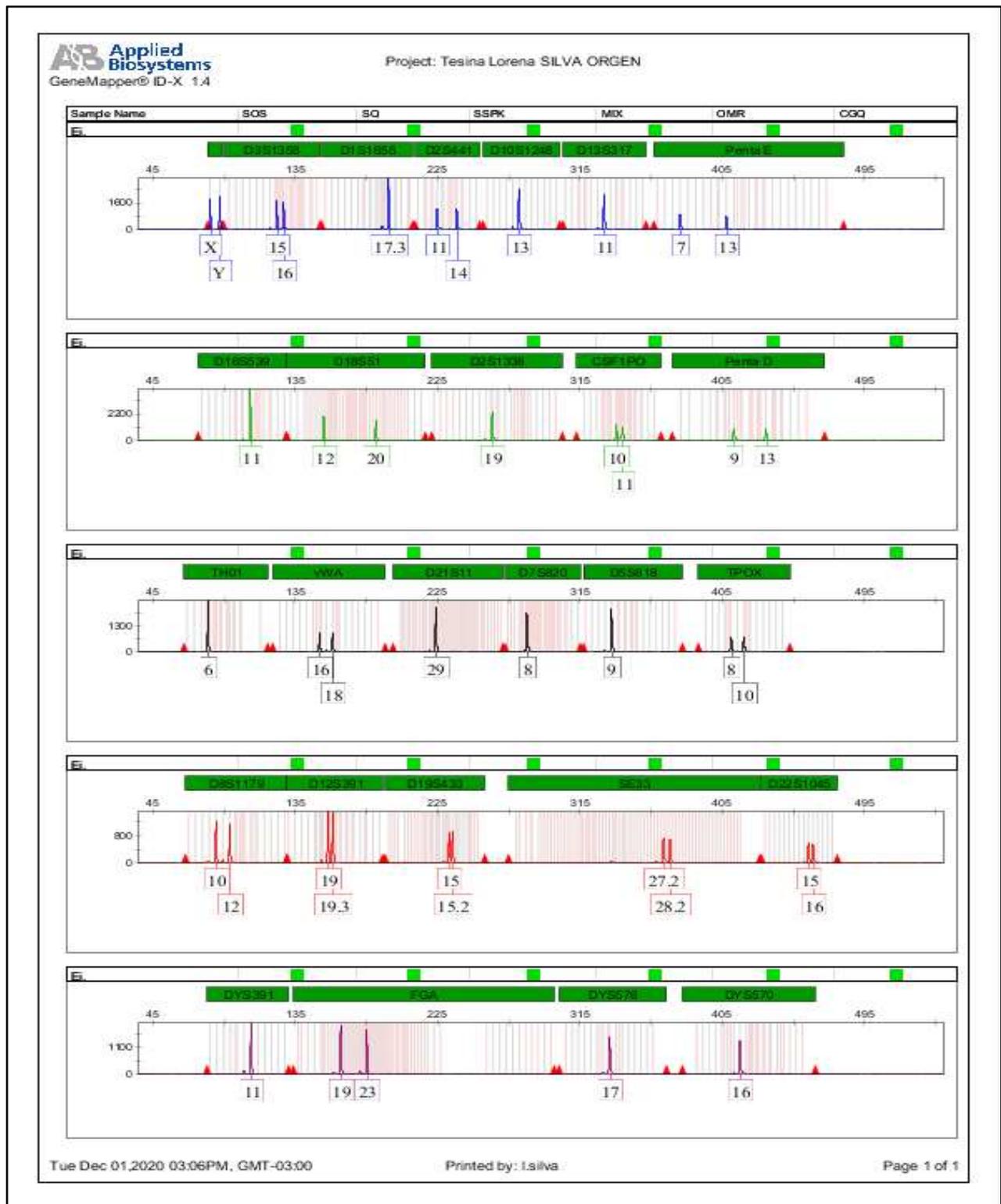


Figura 25: Perfil genético de un individuo. El tamaño (pb) de cada alelo (pico) y el STR que representa se muestra en la parte superior (gris). Debajo de cada pico se indican los alelos para cada STR. Para Amelogenina se muestra el genotipo XY (varón). La medida del pico en RFUs se muestra en el eje de las Y. Se observa la presencia de algunos stutters (pequeños picos a la izquierda de algunos alelos).

Anexo 4: Frecuencias alélicas de la población caucásica disponible en CODIS.

Tabla 8a: Frecuencias alélicas obtenidas por el FBI en una muestra poblacional caucásica de 361 individuos estadounidenses disponibles en CODIS para los alelos comprendidos entre el 2.2 y 18.3.

Alelos	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D7S820	D8S1179	FGA	Penta_D	Penta_E	SE33	TH01	TPOX	vWA
2.2																		0,0042					
3.2																							
4.0																							
4.2																							
5.0																			0,0762		0,0014	0,0014	
6.0																		0,0042			0,2355	0,0014	
6.3																							
7.0														0,0028	0,0277			0,0042	0,1690		0,1939		
8.0	0,0055			0,1205	0,0180									0,0055	0,1440	0,0139		0,0208	0,0139		0,0956	0,5249	
8.1															0,0014								
9.0	0,0139			0,0776	0,1066									0,0416	0,1676	0,0055		0,2216	0,0125		0,1191	0,1274	
9.1															0,0014								
9.3																						0,3449	
10.0	0,2202			0,0471	0,0568	0,0083	0,0014	0,0028				0,2105		0,0554	0,2562	0,1025		0,1150	0,0859		0,0083	0,0499	
10.1																							
10.2																							
10.3																							
11.0	0,3089	0,0014		0,3255	0,3144	0,0097	0,0055	0,0776		0,1399		0,3435	0,0014	0,3560	0,2050	0,0762		0,1260	0,0873	0,0014	0,0014	0,2521	
11.2																							
11.3												0,0609											
12.0	0,3601	0,0319		0,2687	0,3144	0,1136	0,0706	0,1163		0,0125		0,0471		0,3878	0,1593	0,1676		0,2327	0,1994	0,0069		0,0416	0,0014
12.2								0,0014															
12.3												0,0042											
13.0	0,0817	0,3075		0,1163	0,1634	0,1233	0,2548	0,0665		0,0069		0,0291	0,0014	0,1427	0,0346	0,3296		0,1967	0,0859	0,0166		0,0014	0,0014
13.2							0,0069																
13.3																							
13.4																							
14.0	0,0097	0,2978		0,0429	0,0263	0,1343	0,3615	0,0789		0,0568		0,2410	0,1066	0,0069	0,0042	0,1662		0,0609	0,0623	0,0249		0,0928	
14.2					0,0014	0,0235															0,0028		
14.3								0,0028															
15.0		0,1967	0,0319	0,0014		0,1704	0,1565	0,1496		0,3213	0,0014	0,0596	0,2729	0,0014		0,1039		0,0097	0,0429	0,0402		0,1053	
15.2							0,0360																
15.3								0,0582															
15.4																					0,0014		
16.0		0,1330	0,0222			0,1468	0,0568	0,1357		0,3823	0,0374	0,0014	0,2382			0,0332		0,0028	0,0512	0,0402		0,2008	
16.2						0,0014	0,0152														0,0014		
16.3								0,0609															
17.0		0,0277	0,1274			0,1385	0,0069	0,0471		0,0748	0,1856		0,2105			0,0014		0,0014	0,0485	0,0623		0,2839	
17.1																							
17.2							0,0014																
17.3			0,0208					0,1330													0,0014		
18.0		0,0014	0,1717			0,0776		0,0055		0,0055	0,0734		0,1510				0,0249		0,0332	0,0720		0,2022	
18.1																							
18.2							0,0014																
18.3			0,0249					0,0499													0,0014		

Tabla 8b: Frecuencias alélicas obtenidas por el FBI en una muestra poblacional caucásica de 361 individuos estadounidenses disponibles en CODIS para los alelos comprendidos entre el 19 y 36.

	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D7S820	D8S1179	FGA	PentaD	PentaE	SE33	TH01	TPOX	vWA
Alelos																							
19.0		0,0028	0,1247			0,0402					0,1205		0,0166				0,0499		0,0152	0,0720			0,1039
19.1																							
19.2																				0,0042			
19.3			0,0042					0,0152															
19.4																							
20.0			0,1108			0,0180					0,1565		0,0014				0,1233		0,0097	0,0582			0,0069
20.1																							
20.2																					0,0097		
20.3																							
21.0			0,1288			0,0097					0,0374						0,1787		0,0028	0,0249			0,0014
21.2																	0,0055			0,0235			
21.3																							
22.0			0,0956			0,0069					0,0346						0,2050		0,0014	0,0139			
22.2																	0,0125			0,0374			
22.3																							
23.0			0,0693								0,1053						0,1524			0,0028			
23.2																	0,0028			0,0360			
23.3																							
24.0			0,0471								0,1150						0,1343		0,0014	0,0014			
24.2																	0,0014			0,0222			
24.3																							
25.0			0,0166								0,1025						0,0789						
25.2								0,0014												0,0416			
26.0			0,0028								0,0305						0,0263						
26.2																				0,0416			
27.0			0,0014						0,0222								0,0042						
27.2																				0,0942			
27.3																							
28.0								0,1593															
28.2																				0,0762			
28.3																				0,0014			
29.0									0,2022														
29.2									0,0028												0,0554		
29.3																							
30.0									0,2825														
30.2									0,0291												0,0568		
30.3																							
31.0									0,0720												0,0028		
31.2									0,0983												0,0235		
32.0									0,0055												0,0014		
32.2									0,0900												0,0125		
33.0									0,0014												0,0014		
33.1																							
33.2									0,0263												0,0042		
34.0																					0,0083		
34.2									0,0042												0,0014		
35.0									0,0014														
36.0									0,0014														

Anexo 5: Frecuencias alélicas para 15 marcadores STRs obtenidas a partir de una muestra población de 200 individuos uruguayos.

Tabla 9: Frecuencias alélicas obtenidas para 15 marcadores STR en una muestra poblacional uruguaya de 200 individuos.

	TH01	TPOX	CSF1PO	VWA	D3S1358	FGA	D5S818	D13S317	D7S820	D8S1179	D16S539	D18S51	D21S11	PentaD	PentaE
Alelos															
2.2														0,00311	
5														0,00311	0,05901
6	0,2112	0,00311													0,01553
7	0,1957	0,00621	0,00311				0,03106		0,02174					0,01863	0,08385
8	0,1211	0,45342	0,00621				0,01242	0,09627	0,13354	0,00621	0,01553			0,02484	0,01553
9	0,1801	0,08696	0,02795				0,04037	0,09317	0,14286	0,01863	0,09938	0,00311		0,19565	0,01863
9.3	0,2795														
10	0,0124	0,07143	0,2795				0,05901	0,06832	0,30435	0,07143	0,08385	0,00311		0,16149	0,09006
11		0,32609	0,34783		0,00311		0,35714	0,28261	0,21739	0,09627	0,34161	0,01242		0,13665	0,14596
12		0,04658	0,29193				0,31988	0,25155	0,13975	0,1118	0,28571	0,16149		0,13043	0,21118
13		0,00311	0,04348		0,00311		0,17081	0,17081	0,03727	0,28261	0,15217	0,14286		0,18944	0,1087
13.2													0,00621		
14				0,07143	0,08696		0,00621	0,03416	0,00311	0,25776	0,01863	0,15217		0,10559	0,0559
15				0,13975	0,28261		0,00311	0,00311		0,13043	0,00311	0,13354		0,01863	0,0559
16				0,21739	0,29814					0,01553		0,1118		0,00311	0,01863
17				0,32919	0,1677	0,00311				0,00932		0,10559		0,00311	0,04969
18				0,15528	0,15217	0,00621						0,08075			0,02795
19				0,06522	0,00621	0,09627						0,03727			0,00932
20				0,01863		0,15528						0,03106			0,01863
21				0,00311		0,14596						0,00932			0,00311
21.2						0,00311									
22						0,19565						0,00311			0,01242
22.2						0,00311									
23						0,13043						0,00621			
23.2						0,00311									
24						0,12112									
24.2						0,00311							0,00621		
25						0,07453									
25.2						0,00311							0,00311		
26						0,03727									
27						0,01553									
28													0,04348		
28.2						0,00311							0,13043		
29													0,00621		
29.2													0,18323		
30													0,00311		
30.2													0,20497		
31													0,02795		
31.2													0,07453		
32													0,1118		
32.2													0,00621		
33.2													0,12422		
34.2													0,06832		
													0,00311		