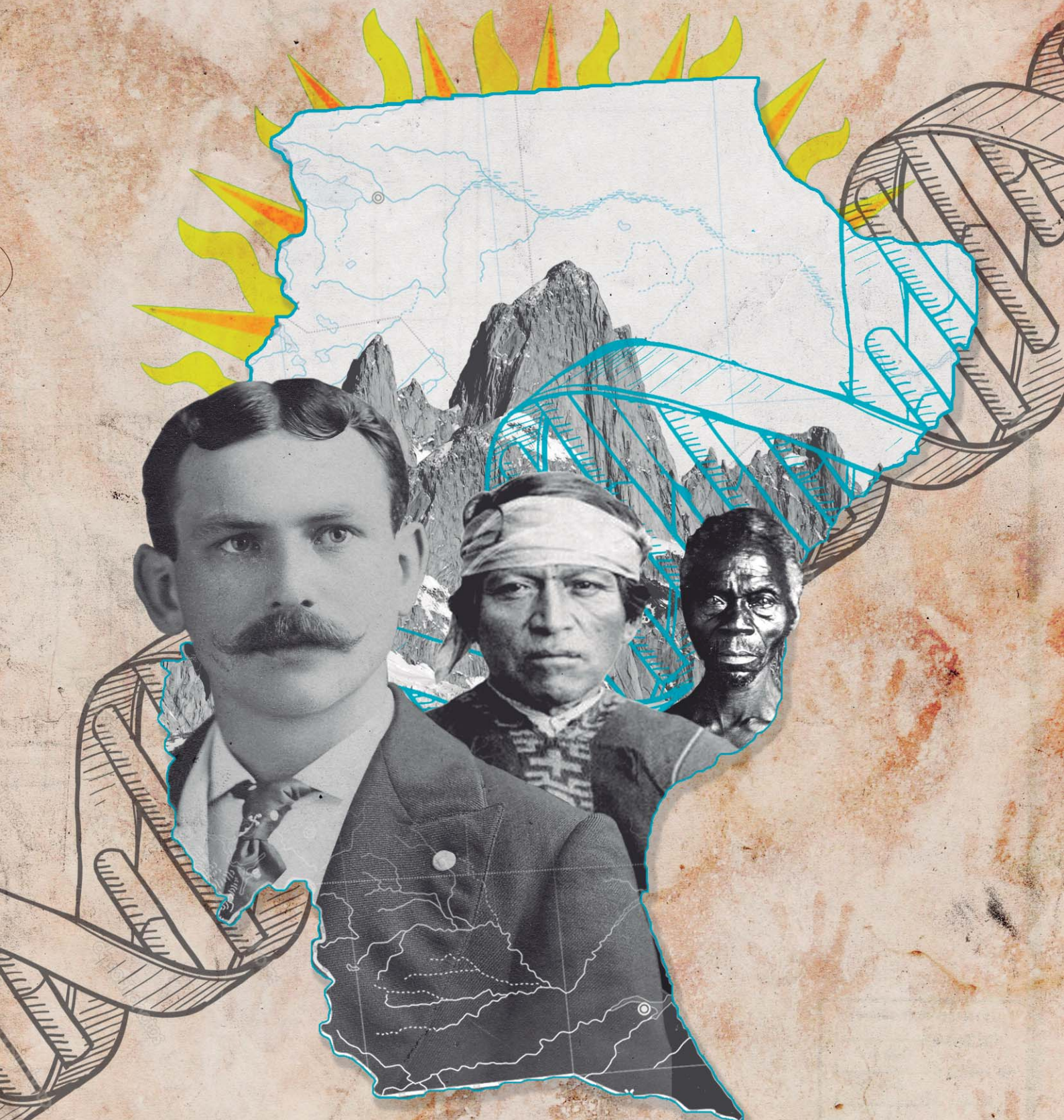


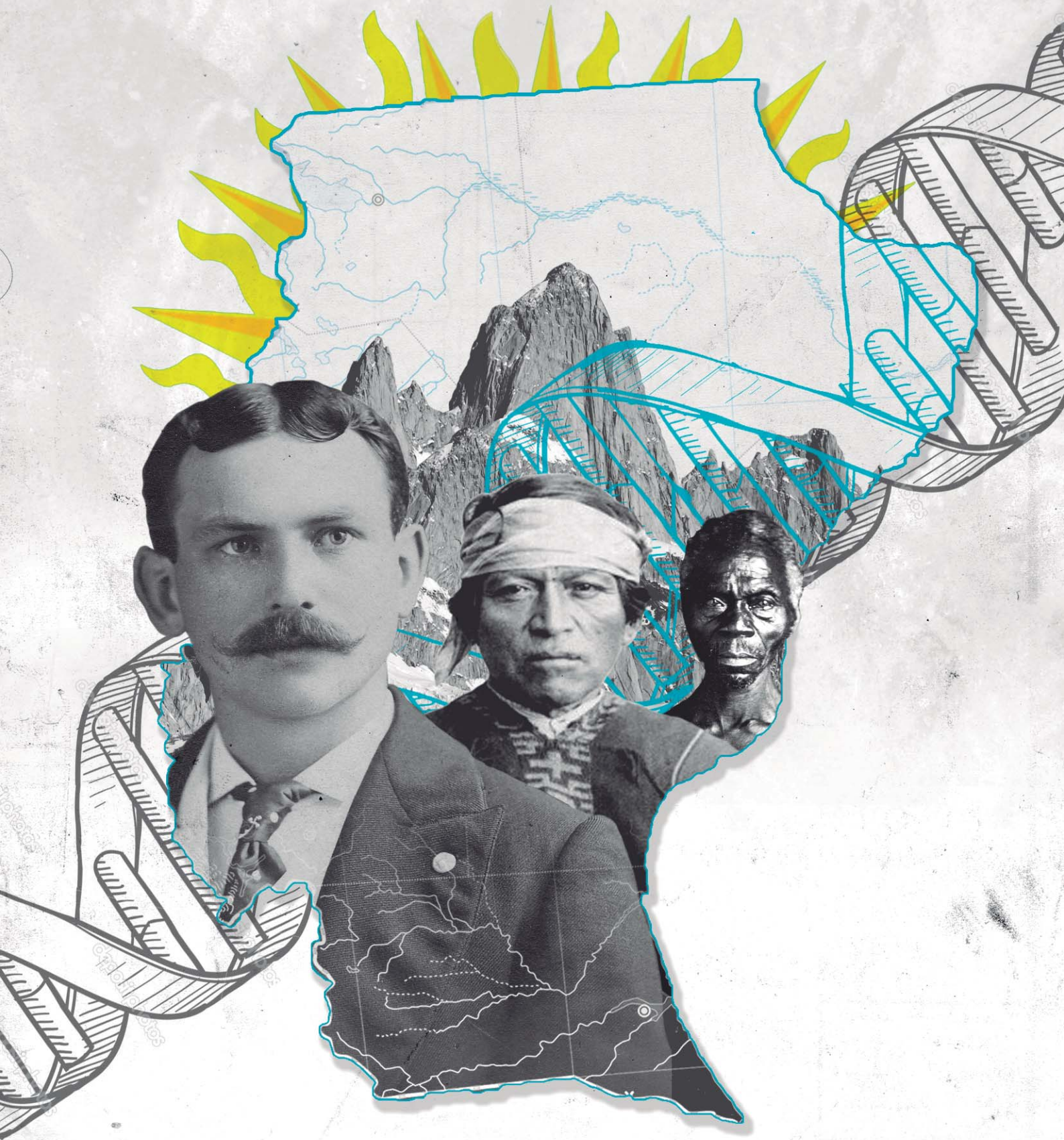
TESIS DE MAESTRIA 2015

# ANÁLISIS DE LA ESTRUCTURA GENÉTICA DE LA POBLACIÓN HUMANA DE LA PROVINCIA DE SANTA CRUZ

BQCA. HORTENSIA DEL VALLE CANO Y DIRECTORA DRA. ADRIANA ANDREA SALA









**UNSAM**  
UNIVERSIDAD  
NACIONAL DE  
**SAN MARTÍN**

**UNIVERSIDAD NACIONAL DE SAN MARTIN**  
**ANLIS "CARLOS G. MALBRAN"**

**MAESTRIA EN MICROBIOLOGIA MOLECULAR**

**“Análisis de la estructura genética de la población  
humana de la Provincia de Santa Cruz”**

**Bqca. Hortensia del Valle Cano**  
**Tesis de Maestría**

**Directora: Dra. Adriana Andrea Sala**

**Servicio de Huellas Digitales Genéticas (SHDG)**

**Facultad de Farmacia y Bioquímica - Universidad de Buenos Aires**

**Lugar de realización de la Tesis**

**Laboratorio Regional de Investigación Forense**

**Poder Judicial de la Provincia de Santa Cruz**

**2015**

*“No basta dar pasos que un día puedan conducir hasta la meta, sino que cada paso ha de ser una meta sin dejar de ser un paso”*

**Johann P. Eckermann**

Este Trabajo de Investigación, si bien ha requerido de un gran esfuerzo y dedicación por mi parte y sobre todo de la Dra. Andrea Sala, mi Directora de Tesis, no se habría podido realizar sin la valiosa y generosa colaboración de muchas personas a quienes quiero expresar mi más profundo agradecimiento.

Primero y antes que nada, dar gracias a Dios, por estar conmigo en todo momento, por iluminar mi mente y por haber puesto oportunamente en mi camino a aquellas personas que han sido mi apoyo y compañía durante todo este tiempo.

Como no agradecer a mi familia, a mi esposo Manuel, a mis hijos Juan Manuel, Agustina y María Fernanda por su inmenso amor, apoyo incondicional, sus oportunos consejos y sobre todo su constante estímulo para continuar hasta el final. Sin ustedes esto no habría sido posible. A mis hermanas y hermano Gracias por confiar en mí e impulsarme para lograr mis metas. ¡Gracias, los Amo Profundamente!

Quiero agradecer a la Universidad Nacional de San Martín, al Instituto ANLIS “Dr. C. Malbran” y especialmente a la Dra. Mirta Carlomagno quien con su gestión y acompañamiento, permitió que profesionales del interior del país tuviéramos la posibilidad de acceder a este nivel de capacitación. A la Lic. Natalia Mendonça por responder a nuestros requerimientos y facilitarnos cada trámite requerido, Muchas Gracias!

Agradezco al Dr. Daniel Corach por compartir tantos conocimientos en forma desinteresada, brindando su tiempo, compromiso y dedicación.

A la Dra. Mariela Caputo por su disposición y voluntad para enseñarme y a todo el personal del Servicio de Huellas Digitales Genéticas de la Facultad de Farmacia y Bioquímica de la UBA, por hacerme sentir durante el tiempo de pasantía como una más entre ustedes. Muchas gracias por todo !!!

También es mi deber y deseo, agradecer de manera muy especial y sincera a la Dra. Andrea Sala, por aceptarme para realizar esta tesis bajo su dirección. Su confianza en mí, infinita paciencia, apoyo, tiempo dedicado y la gran capacidad para guiar mis ideas con sus tan acertados comentarios, han sido un aporte inmensurable en el desarrollo de este trabajo y en especial, en mi formación profesional. Todo mi respeto y admiración para Usted, Gracias !!

Gracias también a mis ex compañeros del Servicio de Microbiología del Hospital Regional Río Gallegos y a mis compañeros de Laboratorio Regional de Investigación Forense, lugar donde inicié mis pasos en Genética Forense y en el que sentí su compañía incondicional en todo momento.

Gracias !! Liliana, María Cristina, Adrián, Fernando, Adriana y Horacio.

Debo agradecer también a las autoridades del Tribunal Superior de Justicia de la Provincia de Santa Cruz y a todo el personal del Poder Judicial que de una u otra manera me acompañaron proporcionándome todo lo necesario durante el desarrollo del mismo.

Quiero agradecer a Braian Paul Cortez, estudiante de Diseño Gráfico en La Universidad de La Plata y nacido en Río Gallegos por ayudarme con la compaginación y diseño de este trabajo.

Finalmente, a todas aquellas personas que se interesaron y colaboraron con este trabajo con su apoyo, aliento y orientación.

*¡A todos, Dios los Bendiga y Gracias!*

***A mis Padres Humberto Cano y Olga Rodriguez***

*Quiero Honrar a quienes me dieron la vida, el amor y los valores;  
pilares fundamentales para enfrentar y superar obstáculos en pos de lograr  
cada sueño.*

# INDICE



# INDICE

Índice de Figuras.....	1
Índice de Tablas.....	3
<b>I-INTRODUCCION</b>	
I-1.Marco histórico-demográfico de la provincia de Santa Cruz.....	5
I-1.1. Los primeros pobladores.....	6
I-1.2.La provincia de Santa Cruz, población y los procesos migratorios.....	7
I-2.Marco Teórico.....	9
I-2.1.Análisis de ADN e identificación humana.....	9
I-2.2.Evolución histórica de los sistemas identificación humana.....	9
I-2.3.Microsatélites(STRs).....	11
I-2.4.SNPs.....	12
I-2.5.InDel.....	13
I-2.6.Marcadores de herencia uniparental.....	14
I-2.6.1.Cromosoma Y.....	14
I-2.6.2.ADN mitocondrial.....	16
I-2.7.Bases de datos de referencia e Inteligentes.....	17
<b>II-OBJETIVOS E HIPOTESIS</b>	
II-1.Objetivos.....	19
II-2.Hipotesis.....	19
<b>III- MATERIALES Y METODOS</b>	
III-1.Objeto de estudio.....	21
III-2.Selección de las muestras.....	21
III-3. Muestras analizadas.....	22
III-4. Consideraciones Éticas.....	22
III-5.Método de Extracción de ADN.....	22
III-6. Cuantificación de ADN.....	22
III-7. Amplificación por PCR multiplex.....	23
III-7.1.Microsatélites o STRs autosómicos.....	23
III-7.2.Microsatelites o STRs de Cromosoma Y.....	23
III-8. SNPs de Cromosoma Y.....	23
III-9.Análisis Estadísticos.....	23
III-9.1. Análisis de STRs autosómicos.....	23
III-9.2.Comparaciones interpoblacionales.....	24
III-9.3. Análisis de polimorfismos de Cromosoma Y.....	24
Anexo I.....	25
<b>IV-RESULTADOS</b>	
IV-1. Resultados de STRsautosómicos.....	29
IV-1.1.Análisis de parámetros estadísticos de interés forense.....	40
IV-2. Comparaciones interpoblacionales.....	40
IV-3.Análisis de estructura.....	44
IV-4. Resultados de STRs de Cromosoma Y.....	45
<b>V-DISCUSIÓN</b> .....	51
<b>VI-CONCLUSIONES</b> .....	54
VI-1.Perspectivas futuras.....	55
<b>Difusión de resultados</b> .....	57
<b>Lista de Abreviaturas y Siglas</b> .....	58
<b>VII-BIBLIOGRAFIA</b> .....	62
<b>Resumen</b> .....	70
<b>Abstract</b> .....	71

## INDICE DE FIGURAS

<b>Figura 1.</b> Mapa Demográfico de la Provincia de Santa Cruz.....	5
<b>Figura 2.</b> Evolución histórica de la Provincia de Santa Cruz.....	9
<b>Figura 3.</b> Estructura de un STR.....	11
<b>Figura 4.</b> Core de 13 STRs del CODIS.....	12
<b>Figura 5.</b> Representación esquemática de un SNP.....	13
<b>Figura 6.</b> Representación esquemática de polimorfismo Indel.....	14
<b>Figura 7.</b> Herencia del cromosoma Y.....	15
<b>Figura 8.</b> Herencia del ADNmit.....	16
<b>Figura 9.</b> Mapa de la Primera y Segunda Circunscripción Judicial.....	21
<b>Figura 10.</b> Distribución de las Frecuencias alélicas para el marcador D8S1179.....	31
<b>Figura 11.</b> Distribución de las Frecuencias alélicas para el marcador D21S11.....	31
<b>Figura 12.</b> Distribución de las Frecuencias alélicas para el marcador D7S820.....	32
<b>Figura 13.</b> Distribución de las Frecuencias alélicas para el marcador CSF1PO.....	32
<b>Figura 14.</b> Distribución de las Frecuencias alélicas para el marcador D3S1358.....	33
<b>Figura 15.</b> Distribución de las Frecuencias alélicas para el marcador THO1.....	33
<b>Figura 16.</b> Distribución de las Frecuencias alélicas para el marcador D13S317.....	34
<b>Figura 17.</b> Distribución de las Frecuencias alélicas para el marcador D16S539.....	34
<b>Figura 18.</b> Distribución de las Frecuencias alélicas para el marcador D2S1338.....	35
<b>Figura 19.</b> Distribución de las Frecuencias alélicas para el marcador D19S433.....	35
<b>Figura 20.</b> Distribución de las Frecuencias alélicas para el marcador vWA.....	36
<b>Figura 21.</b> Distribución de las Frecuencias alélicas para el marcador TPOX.....	36
<b>Figura 22.</b> Distribución de las Frecuencias alélicas para el marcador D18S51.....	37
<b>Figura 23.</b> Distribución de las Frecuencias alélicas para el marcador D5S818.....	37
<b>Figura 24.</b> Distribución de las Frecuencias alélicas para el marcador FGA.....	38
<b>Figura 25.</b> Distribución de frecuencias alélicas considerando únicamente los alelos más frecuentes para cada loci investigado.....	38
<b>Figura 26.</b> Distribución comparativa de frecuencias para el marcador D16S539.....	39
<b>Figura 27.</b> Número promedio de diferencias entre pares dentro y entre poblaciones.....	42
<b>Figura 28.</b> Representación MDS (multi dimensional scaling) de las distancias genéticas entre las poblaciones de Buenos Aires, Chubut; Río Negro y Santa Cruz.....	42
<b>Figura 29.</b> Dendograma de representación de las distancias genéticas entre las poblaciones de Buenos Aires ; Chubut ; Río Negro y Santa Cruz.....	43
<b>Figura 30.</b> Bar Plot obtenido a partir del análisis de estructura.....	44
<b>Figura 31.</b> Frecuencia de Haplogrupos en población estudiada.....	47
<b>Figura 32.</b> Representación MDS de las distancias genéticas (Fst) basadas en el haplotipo mínimo de cromosoma Y.....	47
<b>Figura 33.</b> Dendograma de representación de las distancias genéticas entre las poblaciones de Buenos Aires; Chubut ; Río Negro y Santa Cruz , haplotipominimo de cromosoma Y.....	48

## INDICE DE TABLAS

<b>Tabla 1.</b> Índices moleculares en los tres grupos regionales de la provincia de Santa Cruz.....	29
<b>Tabla 2.</b> Distribución de frecuencias alélicas y parámetros genéticos poblacionales y forenses en la población de Santa Cruz.....	30
<b>Tabla 3.</b> Distancias genéticas (Fst).....	42
<b>Tabla 4.</b> Matriz de significancia.....	42
<b>Tabla 5.</b> Resultados de AMOVA.....	43
<b>Tabla 6.</b> Composición Genética de las poblaciones.....	44
<b>Tabla 7.</b> Haplotipos de cromosoma Y en la Población de Santa Cruz.....	46

# **I - INTRODUCCION**

## I - INTRODUCCION

### I-1. *Marco histórico-demográfico de la provincia de Santa Cruz*

Desde el punto de vista geográfico, Santa Cruz es la provincia más austral de la Patagonia continental argentina. Se encuentra ubicada entre los 46°00' y 52°23' de Latitud Sur y 65°43' y 73°35' de Longitud Oeste. Limita al norte con la provincia de Chubut, al Sur con la Provincia de Tierra del Fuego y la XIIª Región de la República de Chile, al Este con el Mar Argentino, y al Oeste con las XIª y XIIª Regiones de la República de Chile. Con sus 243 945 km<sup>2</sup> de extensión, es la segunda provincia más extensa del país, solo superada por la provincia de Buenos Aires, con 307 571 km<sup>2</sup>. Según el Censo Nacional del 2010 (INDEC), la población asciende a 273.964 habitantes, siendo la segunda provincia menos densamente poblada (1,1hab/km<sup>2</sup>), después de Tierra del Fuego, Antártida e Islas del Atlántico Sur (Figura 1).

La división política administrativa del territorio provincial se divide en siete departamentos: Corpen Aike, Deseado, Güer Aike, Lago Argentino, Lago Buenos Aires, Magallanes y Río Chico. Su capital es la ciudad de Río Gallegos, cabecera del departamento de Güer Aike. Esta ciudad es sede de los tres poderes públicos, alberga 103.941 habitantes y es la mayor concentración poblacional de la provincia, siendo Caleta Olivia la segunda con una población de 63.974 habitantes.

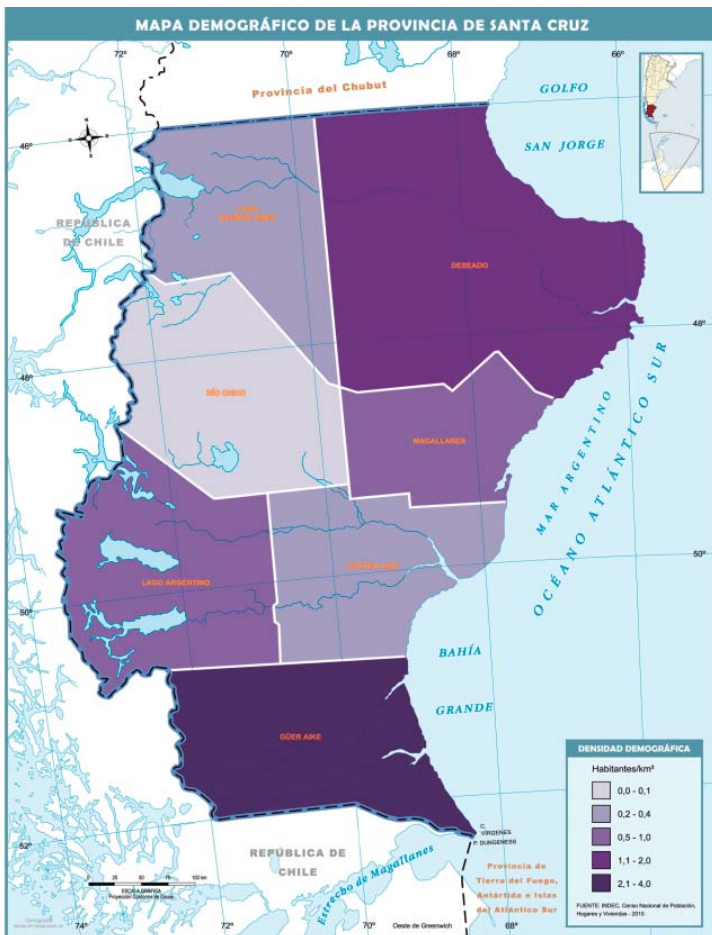


Figura 1 Mapa Demográfico de la Provincia de Santa Cruz-Mapoteca M. de Educación de la Nación.

### I-1.1. Los primeros pobladores

Se estima que los primeros pobladores, los Aónikenk (“gente del sur”) llegaron a la zona magallánica alrededor de 12.000 años a.c.(Tarragó, 2000).

Los Aónikenk(conocidos posteriormente como Patagones o Tehuelches) fueron tribus nómades pedestres, que se desplazaban por las estepas de la Patagonia en busca de animales para cazar y alimentarse. La introducción europea del caballo transformó a estos pobladores de cazadores-recolectores a nómades-ecuestres.

El primer contacto entre un europeo y un aborígen austral se produjo en la primera expedición de Hernando de Magallanes, en 1520, cuando reciben el nombre de Patagones, en alusión a sus características físicas (Pigafetta, 1899).

Varios autores reconocen dos grupos bien diferenciados desde el punto de vista cultural y lingüístico: Tehuelches septentrionales, distribuido aproximadamente entre el Río Colorado en la Provincia de Río Negro y el Río Santa Cruz, en la provincia del mismo nombre; y Tehuelches meridionales, habitantes del área al sur del Río Santa Cruz hasta la costa norte del Estrecho de Magallanes (Casamiquela, 1965; Escalada, 1949).

Durante los siglos XVII y XVIII comenzó además la penetración de los Mapuches (“gente de la tierra”) o araucanos, al este de los Andes, con fines de comercio y alianzas, que con llevó una gran influencia cultural sobre los tehuelches y otros pueblos, en un proceso denominado “araucanización” de las Pampas y la Patagonia. Los tehuelches adoptaron muchas de las costumbres y lengua mapuches, mientras los mapuches adoptaban parte del modo de vida tehuelche, constituyendo una matriz cultural mapuche-tehuelche.

Los aónikenk o tehuelches y mapuches también sufrieron el impacto de la cultura de los europeos. Sin embargo, aunque incorporaron más fácilmente las nuevas costumbres, no se libraron de las enfermedades contagiosas, como la viruela, sífilis, tuberculosis, que diezmaron estas poblaciones. Al crearse las fronteras entre Chile y Argentina y comenzar la colonización perdieron sus tierras y no pudieron seguir desplazándose libremente por las pampas magallánicas.

Entre los años 1821 y 1899 más de 10600 aborígenes de la Patagonia fueron asesinados durante las campañas militares denominadas Conquista del Desierto (Martínez Sarazola, 2005).

La Encuesta Complementaria de Pueblos Indígenas (ECPI) 2004-2005, dio como resultado que se reconocen y/o descienden en primera generación del pueblo tehuelche 4351 personas, en las provincias del Chubut y Santa Cruz y 78534 Mapuches, distribuidos en las provincias de Chubut, Neuquén, Río Negro, Santa Cruz y Tierra del Fuego.

## I-1.2. La provincia de Santa Cruz, población y los procesos migratorios

Las migraciones poblacionales continúan teniendo un importante rol en la composición demográfica y la configuración económica de la mayoría de los países del mundo. En nuestro país, tanto la inmigración masiva de extranjeros de mediados del siglo XIX, como las migraciones internas entre provincias y centros urbanos de las últimas décadas, son fenómenos que explican los patrones genéticos poblacionales existentes y que fueron demostrados previamente mediante estudios serológicos y demográficos (Palatnik Marcos 1968; Quiroga Mucho et al., 1988) y luego confirmados por estudios de identificación molecular. (Corach et al., 2006; Corach et al., 2010; Marino et al., 2006a,b,c; Marino et al., 2007; 2008; Sala et al., 1998; 1999). La observación de las tasas de migración internacional e interna en Argentina en el tiempo permite dar cuenta de la relevancia de estos hechos, ubicando al país entre los de mayor movilidad poblacional. Si bien las tasas de migración han variado en respuesta a distintos fenómenos históricos según las épocas, el país se caracteriza por experimentar altas tasas de migración y redistribución espacial de su población.

Entre las variables más utilizadas para explicar la decisión de migrar a otro país o localidad se pueden mencionar la búsqueda de mejores oportunidades laborales o bienestar económico, la formación de capital humano, factores relacionados con el ciclo de vida (v.g. para formar una familia, por jubilación o retiro); preferencias de tipo cultural, características climáticas, etc. (Pizzolitto, 2006).

Para analizar el poblamiento de Santa Cruz en los últimos cien años, debemos referirnos en primer lugar a su conformación político institucional consignando que fue un Territorio Nacional hasta su provincialización en los años 1956 – 1957 (Barbería, 1995).

Esta situación traía aparejada la pérdida de la autonomía política y la sensación permanente por parte de la sociedad civil - que concordaba con la realidad - de la ausencia y abandono por parte del Estado nacional.

En segundo lugar, debemos referirnos a la conformación de una región autárquica que abarcaba la zona sur de Chile, el territorio de Santa Cruz y la isla de Tierra del Fuego con una gran conexión con las islas Malvinas. Esta región se caracterizaba por una muy débil relación institucional con las respectivas metrópolis de Buenos Aires y Santiago de Chile, teniendo como centro a la ciudad magallánica de Punta Arenas. La misma no era sólo el centro político sino también el económico - financiero, cultural - artístico, religioso y sanitario. Punta Arenas fue el lugar por donde ingresaron la mayor cantidad de contingentes inmigratorios europeos hacia Santa Cruz, siendo también un importante centro portuario ya que toda la navegación transoceánica se realizaba por el estrecho de Magallanes y su flota cubría también el servicio de cabotaje en los puertos atlánticos santacruceños.

La otra vía de acceso al Territorio de Santa Cruz fue Buenos Aires actuando como puerta de ingreso al inmigrante europeo.

En el poblamiento del Territorio los grupos mayoritarios fueron los españoles, chilenos y británicos y los minoritarios franceses, italianos y croatas. Entre 1895 y 1914 se produce un ingreso masivo de extranjeros europeos, que accedieron a la propiedad de la tierra, y chilenos que conformaron una población muy móvil debido a las tareas estacionales que desempeñaban. Fue un proceso migratorio general de una gran masculinización y ruralización.

A mediados de la década de 1920 la apertura del Canal de Panamá provocó el cese prácticamente total de la utilización del estrecho de Magallanes como paso inter-oceánico. Se inició entonces una etapa de declinación económica y un mayor interés por parte de las respectivas metrópolis por estos territorios, lo que se tradujo en la reinstalación de la aduana argentina, la obligación de "nacionalizar" las estancias de ambos lados de la frontera y la prohibición del servicio de cabotaje de los puertos santacruceños por los barcos magallánicos.

A partir de la década de 1960 se produjeron dos hechos que dinamizaron la economía y la sociedad en Santa Cruz: la provincialización, que provocó un aumento y complejización de la estructura administrativa pública, y el comienzo de la explotación hidrocarburífera en la zona norte de la provincia y carbonífera en la de Río Turbio (con la creación de las Empresas nacionales dedicadas a la extracción: YPF, YCF y Gas del Estado). Se produce un proceso de migración interna, proveniente de otras provincias argentinas, del Norte y NO, y la continuidad de una importante migración chilena (que llegó a monopolizar la extranjera) hacia la cuenca carbonífera. En esta etapa se observó un proceso de urbanización en la zona norte de la provincia y la cuenca carbonífera, manteniendo los niveles de masculinidad.

Esta etapa de gran dinamismo económico- social, fue sucedida por la de la década de 1990 en que la privatización de las Empresas del Estado provocó reducción de personal y un cambio en las condiciones laborales, regresando muchos migrantes a sus lugares de origen (tanto a las provincias argentinas como a Puerto Natales en Chile).

En la actualidad nos encontramos en una etapa de expansión económica debido a una nueva etapa de reestatización de YPF, como así también un fuerte desarrollo turístico en algunas localidades de la provincia, en especial Calafate, lo que generará migración interna y externa.

Además de los grupos poblacionales ya señalados, existe una corriente migratoria limítrofe (en su mayoría bolivianos), muy visible en las localidades de Calafate, Caleta Olivia y Puerto Deseado.

Podríamos concluir entonces que la provincia de Santa Cruz fue siempre, históricamente, un territorio de migraciones, tanto internas como externas, que continúan aunque en menor proporción hasta la fecha (Pierini, 2010) (Figura 2).



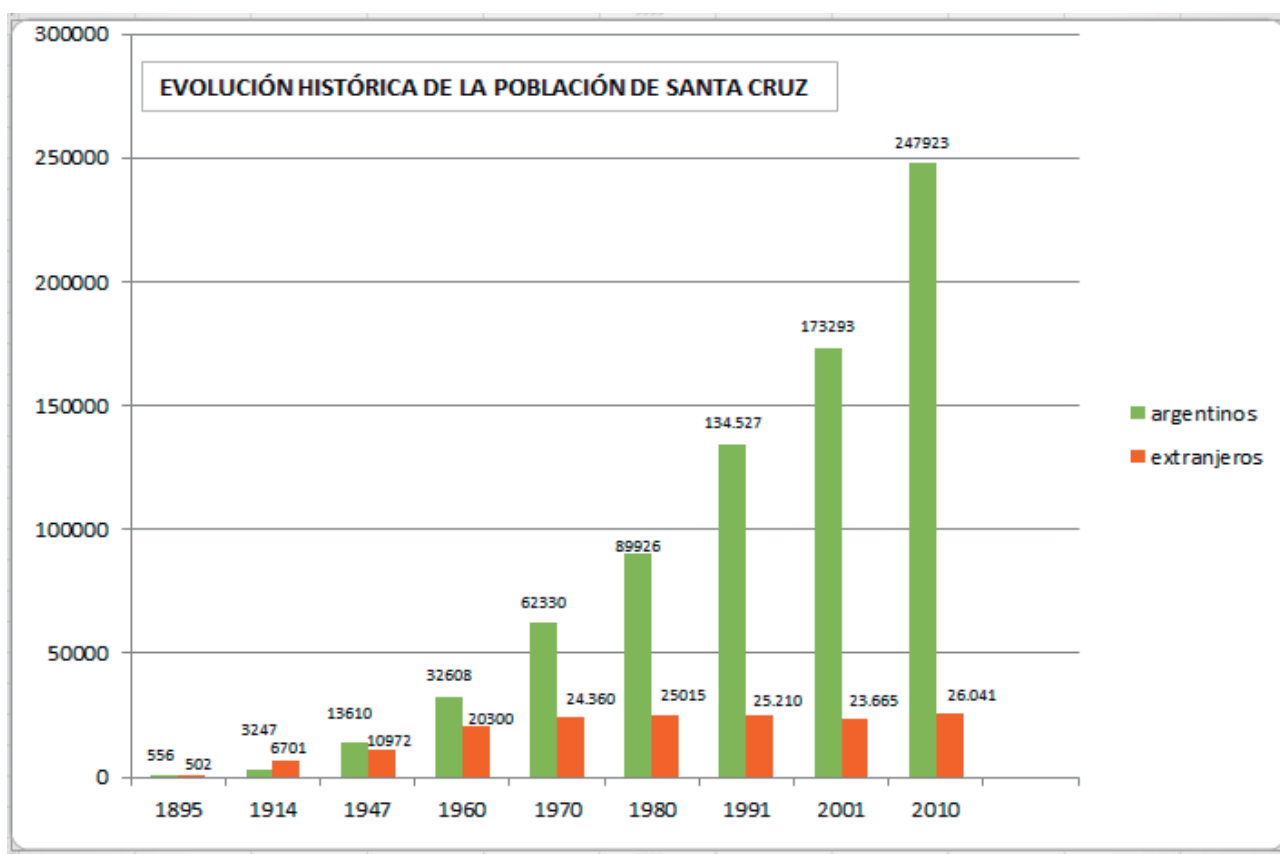


Figura 2 Evolución histórica de la Provincia de Santa Cruz Fuente: INDEC - Instituto Nacional de Estadística y Censos de la República Argentina

## I-2. Marco Teórico

### I-2.1. Análisis de ADN e identificación humana

La Genética Forense podría definirse como una rama de las ciencias forenses que emplea técnicas de biología molecular para la identificación de individuos en base al análisis de ADN (ácido desoxiribonucleico), aportando información relevante en la investigación de casos judiciales, tanto en el marco civil como penal.

El objetivo de la genética forense es identificar el origen de una muestra biológica con la mayor certeza posible. Los polimorfismos presentes en el ADN a los que es posible acceder (caracterizar) hace que ésta herramienta sea extremadamente informativa y el grado de certeza, en consecuencia, puede ser muy alto.

### I-2.2. Evolución histórica de los sistemas identificación humana

La investigación de la variación genética humana se remonta a más de un siglo atrás con Karl Landsteiner (Landsteiner et al., 1900), quien logró poner de manifiesto las variaciones asociadas a los grupos sanguíneos ABO, e inmediatamente se vislumbró su potencial en la resolución de crímenes. Un sistema genético simple, como el ABO, podía ser utilizado para mostrar de manera concluyente que una persona determinada no era el donante de una

evidencia: para probar una ‘exclusión’. Sin embargo, la situación contraria, demostrar que una muestra en realidad provenía de una persona en particular es más difícil y depende del grado de variación revelado por el sistema de tipificación.

Hasta la década de 1980, métodos serológicos y electroforesis de proteínas fueron utilizados para acceder a la diversidad en los grupos sanguíneos y polimorfismos de proteínas. Estudios poblacionales realizados mediante el análisis de estos polimorfismos permitieron caracterizar grupos étnicos particulares (Carnese et al. 1996; Goycochea et al., 1996), aportando información objetiva de gran valor tendiente a establecer la contribución genética de los diferentes grupos étnicos que componen nuestro país (Avena y col., 1999, 2001). Sin embargo, en el campo de la identificación forense mostraron relativamente baja variabilidad y capacidad informativa: la probabilidad de que dos personas no relacionadas compartan una combinación determinada mediante el análisis de 8 sistemas serológicos (probabilidad de coincidencia o “match”) era de  $\sim 0,01$  a  $0,001$ . Si se analizaba otro tipo de fluido, por ejemplo semen, en el que no todos los marcadores están presentes, la probabilidad de coincidencia era mayor. El análisis de una muestra con mezcla de fluidos (ej. mancha mixta en caso de una violación) hacía más complejo aún el análisis (donde el componente celular vaginal enmascara la contribución de los espermatozoides).

La revolución del ADN comenzó en 1984 con el descubrimiento, por Alec Jeffreys en Leicester, Reino Unido, de loci hipervariables conocidos como “Minisatélites”.

Estos eran detectados por hibridación con sondas mediante la técnica de Southern blot. La generación de complejos patrones de bandas mediante el empleo de sondas minisatélites de locus múltiple (MLP o Multi Locus Probe) originó el nombre de DNA Fingerprints o Huellas Digital Genética (Jeffreys et al., 1985).

La probabilidad de coincidencia empleando una sola sonda MLP es de  $3 \times 10^{-11}$ ; dos sondas arroja un valor menor a  $5 \times 10^{-19}$ . Esta probabilidad es tan baja que sólo dos gemelos monocigóticos podrían tener la misma huella digital genética.

Las sondas MLP fueron luego reemplazadas por sondas de locus simple (SLP o Single locus probe) (Nakamura, 1987) que presentaban la ventaja de ser reproducibles y de más fácil interpretación, utilizándose por lo general un conjunto de cuatro a seis sondas.

El primer caso penal en el que se presentó una prueba basada en ADN mediante el empleo de SLP fue en 1986, en Inglaterra, y permitió condenar al sospechoso de un crimen.

El desarrollo de la Reacción en Cadena de la Polimerasa (PCR) en 1987 (Mullis and Faloona, 1987) revolucionó el área de la genética forense ya que permitió incrementar la sensibilidad, amplificando la cantidad de ADN a partir de cantidades ínfimas e incluso de muestras con ADN degradado.

Esta nueva metodología permitió detectar otro tipo de polimorfismos: los microsatélites o STRs (Short Tandem Repeats).

### I-2.3. Microsatélites (STRs)

Los microsatélites fueron descritos inicialmente por Weber y May (1989).

Son polimorfismos de longitud y su variabilidad reside en la cantidad de veces que se repite una unidad de repetición o “core”, de 2 a 7 pb de tamaño (Figura 3). Su gran utilidad se puso rápidamente de manifiesto debido a su abundancia, naturaleza polimórfica, distribución regular por el genoma, tasa de mutación relativamente baja, ubicación cromosómica establecida y su pequeño tamaño, que los hace aptos para ser amplificados por PCR. Fueron introducidos en el campo forense a principios de la década de los noventa por Edwards et al. (1991, 1992), Polymeropoulos et al., (1992) y Tautz (1993), entre otros investigadores, y constituyen elementos extraordinariamente útiles en la identificación humana, mapeo genético y estudio de poblaciones (Chakraborty et al., 1994; Gill et al., 1995; Brinkmann et al., 1996).

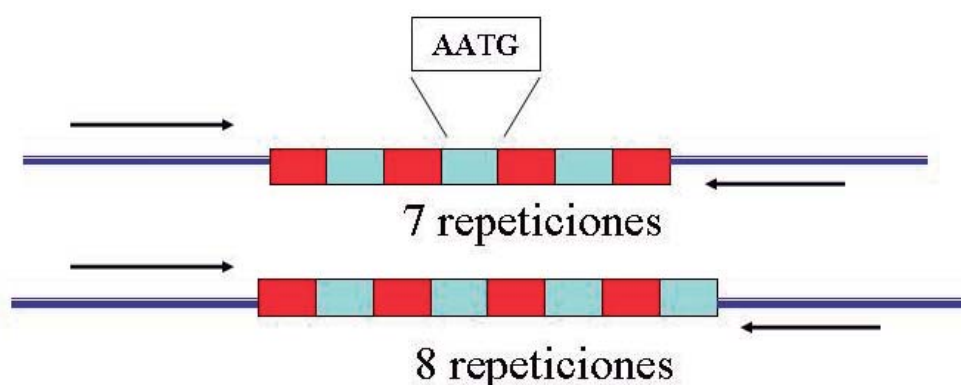


Figura3: Estructura de un STR, cuya unidad de repetición es un core de 4 bases. Se representan los alelos 7 y 8.

El polimorfismo de los microsatélites se debe a la variación en el tamaño de los alelos, pudiendo además presentar variaciones en la secuencia de la unidad de repetición (Urquhart et al., 1994; Brinkmann et al. 1998).

En el año 1999 el FBI definió un conjunto de 13 marcadores denominado CODIS (Combined DNA Index System), que han sido validados internacionalmente y constituyen en su conjunto el número mínimo de marcadores utilizados en genética forense para la identificación y establecimiento de vínculos de parentesco entre individuos (Budowle et al., 1999) (Figura4).

En el año 2006 se define el conjunto estándar Europeo de 12 de STRs (Gill et al., 2006). Dado que existe una superposición parcial con los marcadores del CODIS, constituyen en su conjunto un core de 18 marcadores STRs en total. La incorporación de estos STRs en kits comerciales ha permitido optimizar su empleo en genética forense, generando además resultados reproducibles (Nagy et al., 2005).

La probabilidad de que dos individuos tengan las mismas combinaciones genotípicas para cada uno de los 13 loci incluidos en el CODIS supera uno de cada mil millones. Si existe

un match entre el perfil genético del sospechoso y una mancha en la escena del crimen, la probabilidad de que la mancha haya sido dejada por otro individuo distinto del sospechoso es de una en 1 millón (Jobling, 2003).

En el área de la genética forense los microsatélites continúan siendo los polimorfismos de elección en la determinación de vínculos filiatorios e identificación genética. Existe en la actualidad un gran número de microsatélites descritos y caracterizados en diferentes poblaciones humanas, ubicados tanto en cromosomas autosómicos como en cromosomas sexuales (X e Y) (Gill et al., 2006; Kayser et al., 1997; STRBase : <http://www.cstl.nist.gov>).

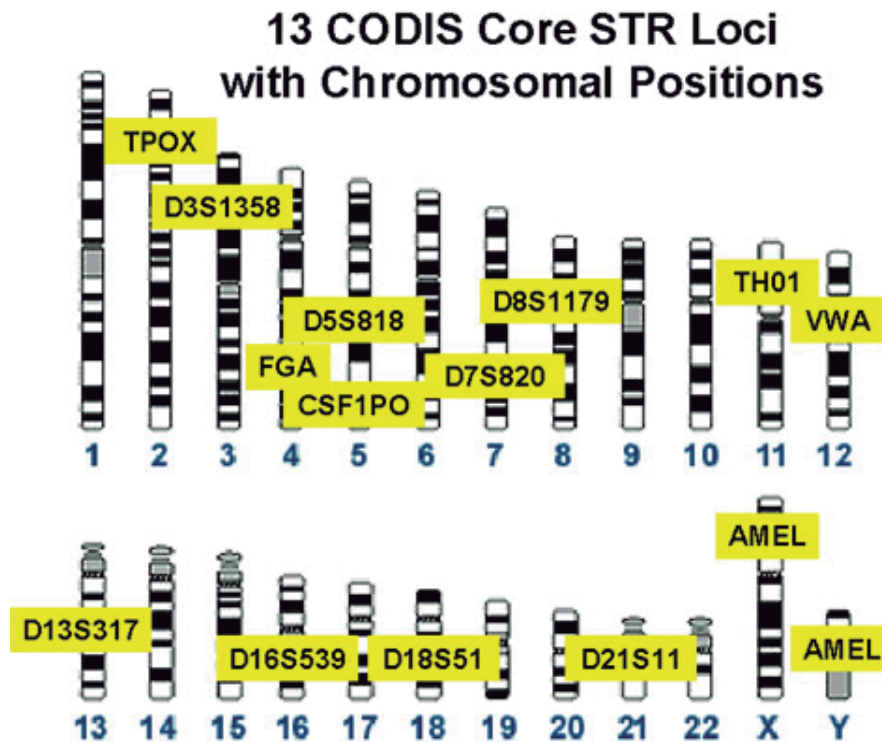


Figura4: Core de 13 STRs del CODIS, de acuerdo a su ubicación cromosómica

#### I-2.4.SNPs

Los SNPs (Single Nucleotide Polymorphism) son marcadores binarios en los que el polimorfismo reside en el cambio de un único nucleótido (Figura 5). Constituyen la fuente más común de variación genética y representan la clase más abundante de polimorfismos humanos (Venter et al., 2001; Budowle et al., 2004; Hammer et al., 2005). En relación a los STRs, los SNPs tienen heterocigosidades mucho más bajas (con un máximo de 0.5 por SNP) y así se requiere tipificar aproximadamente 50 SNPs para reemplazar la capacidad discriminativa de 25 STRs.

La ventaja práctica de un SNP es que puede ser tipificado a partir de amplicones de pequeño tamaño, haciendo de estos marcadores una herramienta útil en material severamente degradado. Presentan, sin embargo, una limitación en el análisis de perfiles mezclados, dada la complejidad en la interpretación de resultados.

Este tipo de marcadores es especialmente útil en genética de poblaciones y estudio de

ancestralidad (Nassir et al., 2009; Li et al., 2008; Lao et al., 2006; 2010).

El SNP Consortium ha creado una base de datos que contiene alrededor de 1,8 millones de SNPs (The SNP Consortium, 2001, 2002 y 2005, disponible en: <http://snp.cshl.org/>), por su parte el National Center for Biotechnology Information (NCBI) mantiene también una base de datos con más de 4 millones de SNPs (<http://www.ncbi.nlm.nih.gov>).

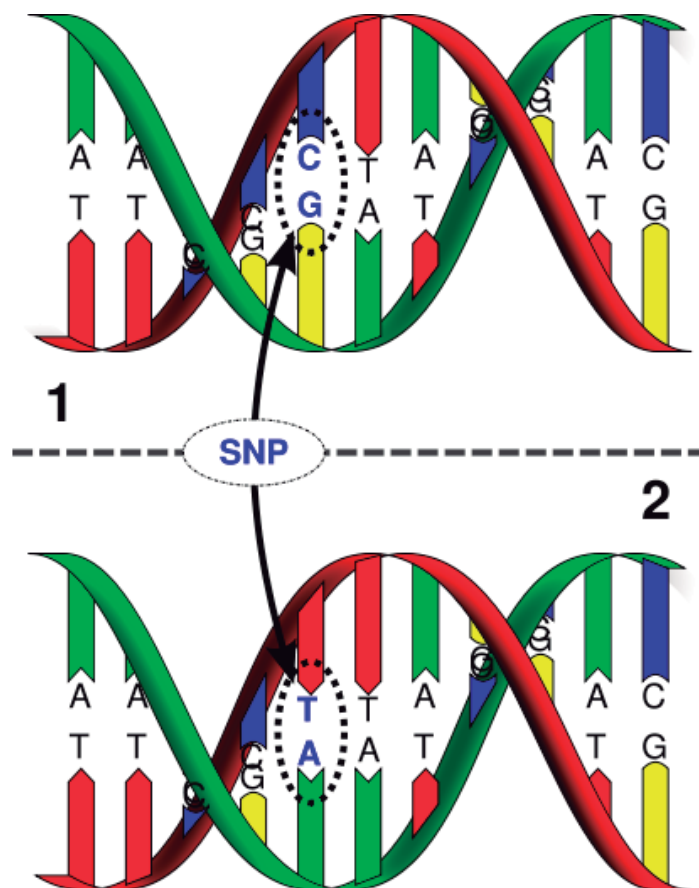


Figura 5: Representación esquemática de un SNP: la molécula de ADN 1 difiere de la 2 en un único par de bases (polimorfismo C/T)

## I-2.5. InDel

Los polimorfismos de Inserción/delección (InDels) son generados por la inserción o delección de uno o más nucleótidos en un determinado locus del genoma (Figura 6). En 2002, Weber et al. identificaron y caracterizaron 2000 polimorfismos de InDels bi-alélica en humanos, poniendo de relieve su utilidad para estudios genéticos, con referencia a su abundancia y facilidad de análisis. Desde entonces el análisis de estos polimorfismos se ha utilizado con distintos objetivos como análisis de filiación (Yang et al., 2005), e investigación en poblaciones humanas como marcadores étnico-específicos así como en identificación forense (Vali et al., 2008; Santos et al., 2010; Da Costa Francezetal., 2012). Existen más de 415.000 polimorfismos únicos InDels en el genoma humano, con una densidad media de uno por cada 7,2 kb. (Mills et al., 2006). El 41% de los InDels corresponden a

polimorfismos cuya longitud promedio es de unas 100 pares de bases, lo que permite resolverlos por amplificación por PCR y electroforesis.

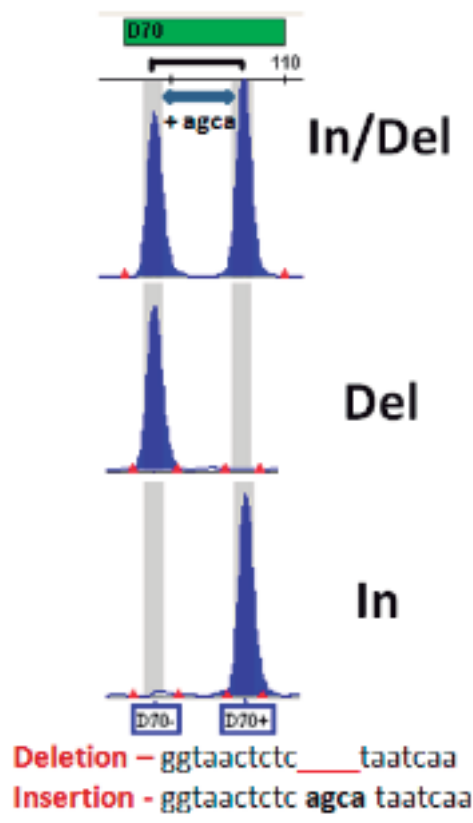


Figura6: Representación esquemática de polimorfismo InDel

## I-2.6. Marcadores de herencia uniparental

Los Marcadores de herencia uniparental han permitido avanzar en la investigación de las poblaciones humanas ya que permiten rastrear evolutivamente el origen de líneas de herencia.

La ausencia de recombinación del ADN mitocondrial (DNAmt) y la región no-recombinante del cromosoma Y (NRY) y su herencia uniparental hacen posible reconstruir la genealogía de cada uno con una precisión relativamente alta.

La comparación de los patrones de diversidad encontrados en estos dos cromosomas permiten inferir el origen étnico/geográfico del linaje al que pertenece un individuo (Vigilant et al., 1991; Van Oven y Kayser, 1999).

### I-2.6.1. Cromosoma Y

El cromosoma Y se hereda de padres a hijos varones sin sufrir recombinación, excepto en las regiones denominadas PAR 1 y 2 ubicadas en los extremos del cromosoma. En la región no recombinante, también llamada NRY (Non Recombinant Region), se localizan secuencias polimórficas, mini y microsatélites y polimorfismos de nucleótido único o SNPs (Roewer et

al., 1996 y 2000; Kayser et al., 1997; Jobling et al., 2001). La caracterización molecular de un conjunto de polimorfismos ubicados en el cromosoma Y constituye un haplotipo.

El análisis de polimorfismos de cromosoma Y tiene crucial importancia en la identificación forense, donde los asaltos sexuales así como la mayoría de los delitos graves son cometidos por hombres. En el caso de asaltos sexuales la lisis diferencial permite recuperar una fracción de ADN enriquecido en material masculino (Gill et al, 1985).

Existen más de 200 STRs conocidos en el cromosoma Y. Actualmente en genética forense se emplean un conjunto mínimo de 9 marcadores, existiendo Kits comerciales que posibilitan co-amplificar hasta 25 Y-STRs. Para haplotipos de cromosoma Y existe una base de datos a nivel mundial (ver Bases de Datos).

Excluyendo la mutación, todos los individuos varones de una misma patrilinea (hermanos, padre, hijos, tíos paternos, abuelo paterno, etc.) de un sospechoso compartirán su haplotipo, hecho que debe tenerse en cuenta al evaluar el peso de esta evidencia (Figura 7).

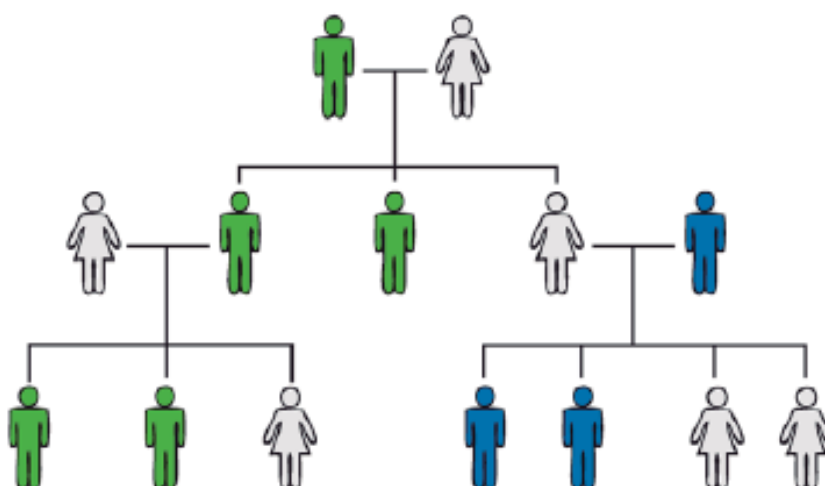


Figura 7 Herencia del cromosoma Y

El análisis de otro tipo de polimorfismos en cromosoma Y, los SNPs, permiten además asignar un origen étnico/geográfico al portador de un haplotipo determinado, herramienta que es de gran utilidad en el estudio de poblaciones humanas y movimientos migratorios.

En 1996, Underhill y cols. describen una variante derivada del tipo ancestral (mutación C/T) del locus denominado inicialmente DYS199 (hoy denominado M3-Q3), que fue encontrada en todos los grupos nativos del norte, centro y sur de América, estando ausente en África, Europa y "Australasia". Este SNP define el haplogrupo Nativo Americano Q1a3a. En nuestro país, este haplogrupo se ha hallado en más del 64% de los cromosomas Y Amerindios analizados (Corach et al., 2001, 2010; Sala y Corach, 2014).

Más recientemente y a partir de 245 marcadores bialélicos, el Consorcio del Cromosoma Y (YCC), reconstruyó la filogenia del hombre en 18 haplogrupos del cromosoma Y, denominados con las letras de A a R (YCC, 2002).

### I-2.6.2. ADN mitocondrial

El genoma mitocondrial no sufre recombinación y es transmitido a los descendientes exclusivamente por vía materna.

La variabilidad presente en el ADN mitocondrial (ADNmt) está determinada por la presencia de SNPs y en menor medida variantes de longitud constituidas por “strech” de Citosinas o presencia de dinucleótidos.

La ventaja del análisis de ADNmt en forense reside en el alto número de copias por célula, entre ~ 200 y 1.700 (Giles et al., 1980; Malyarchuk et al., 2002), haciendo factible su recuperación aún en muestras degradadas o antiguas.

Su análisis en forense se limita a casos en los que se trabaja con cantidad de ADN limitada (ej. pelo, material antiguo o degradado) y en los que es posible contar con individuos de línea materna para su identificación (Gill et al, 1994).

La región no codificante, conocida como región control o D-Loop, de 1122 pb se caracteriza por su elevada tasa de mutación, aproximadamente 10 veces mayor que en las regiones codificantes (Greenberg et al., 1983; Vigilant et al., 1991) y por su elevada variabilidad intrapoblacional (Coskun et al., 2003). La secuenciación de esta región permite definir el haplotipo mitocondrial de un individuo en comparación con la secuencia de referencia (Andrews et al, 1999).

Todos los miembros de una misma línea materna presentan la misma información genética en su genoma mitocondrial, excepto en el caso de mutaciones.

La heteroplasmía mitocondrial (existencia de dos poblaciones de genomas mitocondriales) puede dar lugar a diferentes secuencias encontradas en diferentes pelos o tejidos en un mismo individuo (Figura 8).

El estudio de linajes mitocondriales ha permitido establecer la existencia de haplogrupos Amerindio-específicos, que reafirmaron las teorías migratorias del poblamiento de América (Vigilant, 1991; Torroni et al., 1992).

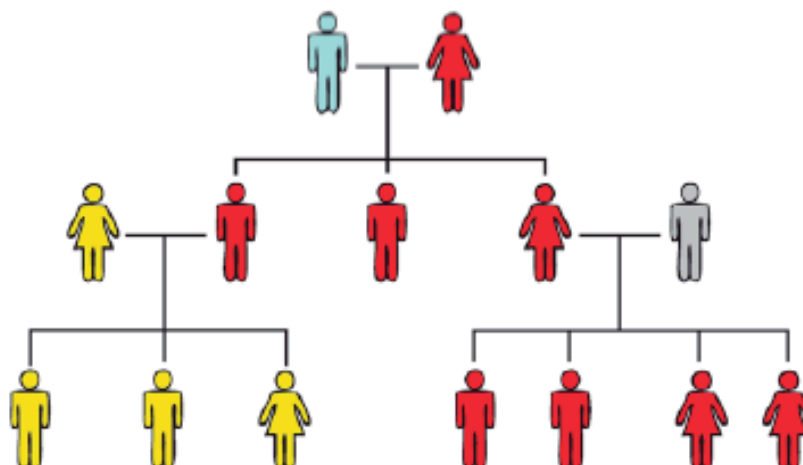


Figura 8 Herencia del ADNmit



## I-2.7. Bases de datos de referencia e Inteligentes

La singularidad de ADN de cada persona (con la excepción de gemelos monocigóticos) y su mera codificación numérica o alfanumérica ha llevado a la creación de bases de datos de ADN inteligentes en los países desarrollados en todo el mundo, la primera en 1995 en el Reino Unido (Martin, 2001).

Las bases de datos inteligentes almacenan la información genética de evidencias y muestras de referencias relacionadas a hechos delictivos, permitiendo realizar búsquedas para hallar coincidencias o “match” que conduzcan a la identificación del perpetrador de un crimen.

Estas bases de datos inteligentes presentan diferentes características, según la legislación de cada país, entre las que se destacan: tipo de perfiles incorporados, tipo de muestras, permanencia en la base de datos, etc. El porcentaje de personas incorporadas en las Bases de datos va en aumento en todos los países, aunque varían por país: mientras que en el Reino Unido el 10% de la población está en la base de datos nacional de ADN, el porcentaje en Alemania y en los Países Bajos es de sólo 0,9% y 0,8%, respectivamente (ENFSI 2011).

Las bases de datos de referencia, por otro lado, son aquellas que registran frecuencias alélicas de cada uno de los marcadores que han sido caracterizados en la población de interés. La determinación de estas frecuencias es necesaria para que, mediante el empleo de algoritmos matemático-estadísticos, permita al perito establecer el nivel de certeza de un estudio de ADN.

Actualmente existen dos bases de datos de referencia a nivel mundial para marcadores de herencia uniparental: de haplotipos de cromosoma Y forense ([www.yhrd.org](http://www.yhrd.org)), organizada en el Instituto de Medicina Legal y Ciencias Forenses en Berlín, Alemania, con más de 143.044 haplotipos (para el haplotipo mínimo, Realease 49) muestreados en 850 poblaciones (Willuweit, 2015) y para datos haplotípicos de ADN mitocondrial: base de datos del EMPOP ([www.empop.org](http://www.empop.org)) alojada en el Instituto de Medicina Legal en Innsbruck, Austria, con más de 34.000 haplotipos muestreados en 63 países (Parson, 2007). Estas bases permiten determinar la frecuencia de aparición de un determinado haplotipo en una base de datos mundial o en la base de datos nacional, contando además con herramientas que permiten asignar haplogrupos.

Las bases de datos de referencia para microsatélites autosómicos son generadas a nivel local, existiendo en la actualidad bases de datos de referencia de un amplio número de poblaciones a nivel mundial.

# **II - OBJETIVOS E HIPOTESIS**

## II-1. OBJETIVOS

II-1. Considerando que la población de la provincia de Santa Cruz (SC) constituye un modelo de análisis de gran interés dado la complejidad con la que se ha ido delineando la arquitectura de su población, el objetivo general del presente trabajo es caracterizar la composición genética de esta población mediante el análisis de polimorfismos de herencia mendeliana y de herencia uniparental (cromosoma Y).

Los **objetivos específicos** son:

- Establecer si existen diferencias significativas, desde el punto de vista genético poblacional, en las diferentes regiones geográficas de la provincia.
- Analizar parámetros genético poblacionales en la población en estudio (SC), determinando frecuencias alélicas, genotípicas, índices de diversidad, equilibrio de Hardy-Weinberg y otros parámetros de interés forense (poder discriminativo y de exclusión, Índice de paternidad típico, probabilidad de coincidencia).
- Comparar las frecuencias alélicas y genotípicas de marcadores estudiados en la población santacruceña con los reportados en provincias patagónicas que presentan características similares de mestizaje (Chubut, Río Negro) y con la población de Buenos Aires, dada la historia demográfica de la provincia.
- Determinar parámetros genético-poblacionales de interés: diversidad génica, distancia genética de la población santacruceña respecto de otras poblaciones de nuestro país.
- Realizar un análisis de estructura comparando la población en estudio con poblaciones parentales (v.g Europea, Nativo-americana y Africana).
- Correlacionar los datos genéticos con la información demográfica disponible.

Los resultados obtenidos a través de la realización de este trabajo tiene importancia, por un lado, desde el punto de la investigación básica por su aporte a la genética de poblaciones, y por otro, en el área de la Genética Forense, aportando a la caracterización de poblaciones y generación de base de datos de referencia.

## II-2. HIPOTESIS

La hipótesis del trabajo es que:

*“El análisis de los microsatélites autosómicos propuestos para el presente plan, podrá ser una herramienta útil en la investigación de la estructura genética de la población actual de la provincia, permitiendo correlacionar la información genética con la información histórica –demográfica disponible”.*

# **III - MATERIALES Y METODOS**

### III -MATERIALES Y MÉTODOS

#### III-1. Objeto de estudio: Población de Santa Cruz

#### III-2. Selección de las muestras:

Teniendo en cuenta la extensión territorial de la provincia y la baja densidad poblacional, se decidió dividir la población de Santa Cruz en tres grupos de acuerdo a la organización judicial (Primera y Segunda Circunscripción) y localización geográfica. De este modo quedaron identificadas de la siguiente manera:

- **C1E:** Población perteneciente a la Primera Circunscripción Judicial, localizada geográficamente al Este de la provincia (Río Gallegos y Puerto San Julián)
- **C1O:** Población perteneciente a la Primera Circunscripción Judicial, al Oeste de la provincia (Río Turbio y El Calafate)
- **C2:** Población perteneciente a la Segunda Circunscripción (Caleta Olivia, Las Heras, Puerto Deseado y Pico Truncado)



Figura 9 Mapa de la Primera y Segunda Circunscripción Judicial

### **III-3. Muestras analizadas**

El trabajo se llevó a cabo con 145 muestras de referencia (sangre o células de mucosa bucal), de individuos no relacionadas genéticamente, que fueron procesadas durante el año 2014 en el Laboratorio Regional de Investigación Forense (LRIF, Poder Judicial de la provincia de Santa Cruz). De éstas, 44 corresponden a individuos varones. Las muestras proceden de las localidades de Caleta Olivia, Las Heras, Puerto Deseado, Pico Truncado, Puerto San Julián, Río Turbio, El Calafate y Río Gallegos.

El tamaño muestral distribuido según las circunscripciones fue: C1E (N=71); C1O (N=21) y C2 (N=53).

### **III-4. Consideraciones Éticas**

Todas las muestras con las que se trabajó surgieron como requerimientos de causas Civiles y Penales, para establecer relaciones de vínculos biológicos o identificación en causas forenses, y como consecuencia la toma de muestra se realizó por orden judicial. Para el presente análisis, las mismas fueron codificadas a los efectos de trabajar en forma anónima.

### **III-5. Método de Extracción de ADN**

Extracción de muestras de sangre sobre papel de filtro:

- El método utilizado fue validado en el laboratorio del LRIF y se fundamenta en la liberación de hemoglobina por lisis osmótica quedando el ADN fijado al papel (Caputo M., en vías de publicación).
- Para las células de descamación de mucosa bucal, se utilizó métodos convencionales (Quinque et al., 2006), columnas Nexttec y/o Kits de purificación semiautomatizado utilizando el equipo Maxwell Forensic Promega, respetando las instrucciones del fabricante.

### **III-6. Cuantificación de ADN**

La cuantificación de las muestras de ADN utilizadas para la optimización de las reacciones de PCR multiplex se realizó mediante el uso del Kit Quantifiler Human (Applied Biosystems AB) empleando el 7500-Real Time PCR Sistem (AB), respetando los protocolos del fabricante.

## **III-7. Amplificación por PCR multiplex**

### **III-7.1. Microsatélites o STRs autosómicos**

Se amplificaron por PCR multiplex los siguientes STRs: D3S1358, THO1, D21S11, D18S51, D5S818, D16S539, D7S820, D13S317, CSF1PO, VWA, D8S1179, TPOX, FGA, D2S1338, D19S433 y Amelogenina, a través del empleo del kit AmpFISTR® Identifier Plus (Applied Biosystems) de acuerdo al protocolo del fabricante.

### **III-7.2. Microsatelites o STRs de Cromosoma Y**

Se amplificaron por PCR multiplex un conjunto de 23 Y-STRs (Power Y23, Promega Corp). Diez de las muestras fueron analizadas con 17 Y-STRs (kit AmpFISTR® STRs Y filer, AB).

Los productos amplificados, tanto de STRs autosómicos como de cromosoma Y, fueron analizados por electroforesis capilar en el Secuenciador automático ABI 3500 (AB). El tamaño de los fragmentos (en pb.) y su tipificación se determinó por comparación con los correspondientes ladders alélicos provistos por cada kit, en forma automática mediante el uso del software para análisis de fragmentos (Gene Mapper IDX, AB).

## **III-8. SNPs de Cromosoma Y**

Los SNPs de cromosoma Y: M269 (haplogrupo Europeo R1b1b2); M3-Q3\* (Nativo Americano: Q1a3a), U179 (haplogrupo Europeo I), M172 (haplogrupo Europeo: J) y M35 (haplogrupo S Europeo-NE Africano: E1b1b) se analizaron mediante desnaturalización de alta resolución empleando un PCR en Tiempo Real (Rotor Gene-6000) (Zucarelli et al, 2010).

## **III-9. Análisis Estadísticos**

### **III-9.1. Análisis de STRs autosómicos**

Se determinaron parámetros genético-poblacionales que caracterizan a un marcador dentro de una población:

- Distribución de frecuencias alélicas, heterocigosidad observada y esperada, fueron determinadas mediante el empleo del Software GenAIEx v6.501 (Peakall and Smouse, 2006; 2012)
- Frecuencia mínima: para cada locus fue determinada de acuerdo al NRC Report (1992;

1996)

- El Test de Garza Williamson:  
Mide la reducción en el tamaño efectivo de la población por un fenómeno cuello de botella reciente.  
Si el parámetro da  $M < 0.68$ , implica reducción reciente del tamaño de la población. Se debe realizar el test considerando 7 loci o más.
- Equilibrio de Hardy-Weinberg: el Test para determinar si la población se encuentra en EHW se realizó mediante el empleo del Software Arlequin v3.5.1.2 (Excoffier et al., 2010), considerando como significativo una  $p < 0.05$ .
- Los parámetros forenses: Poder discriminativo y de exclusión; PIC (Contenido de Información polimórfica); IPT (Índice de paternidad típico) fueron determinados mediante el empleo del Software PowerStats (Schneider et al., 2000)

### **III-9.2. Comparaciones interpoblacionales:**

Las distancias genéticas (Fst), test de AMOVA, test de Diferenciación Exacto, análisis de parámetros moleculares, se determinaron mediante el empleo del software Arlequin v 3.5.1.2. A partir de las distancias genéticas se realizaron estudios de Multi Dimensional Scaling (XLSTAT Addinsoft Corp, software) comparando la población en estudio con otros grupos geográficamente relacionados.

El análisis de la composición de la población en relación a poblaciones parentales se realizó con el software Structure v 2.3.4 (Pritchard et al., 2000).

Las poblaciones parentales, incluyendo 70 Africanos Sub Saharianos, 53 Europeos y 84 Nativo-americanos fueron seleccionadas del panel del HGDP-CEPH (Human Genome Diversity Cell Line Panel, Human Polymorphism Study Center, resultados suministrados por el Dr. Peter de Knijf, Leiden University).

### **III-9.3. Análisis de polimorfismos de Cromosoma Y**

La pertenencia a Haplogrupos fue determinada en base a los haplotipos de Y-STRs, empleando el Software Haplogroup Predictor (<http://www.hprg.com>) y la herramienta de la base de datos mundial YHRD ([www.yhrd.org](http://www.yhrd.org), release 49). Fueron posteriormente confirmados mediante la identificación de SNPs característicos de haplogrupo por Real Time PCR. Los análisis estadísticos (parámetros de diversidad molecular, frecuencias haplotípicas) y las comparaciones interpoblacionales se realizaron en base a las herramientas informáticas descriptas.



## Anexo I

### Parámetros estadísticos de interés en genética forense:

**a) Frecuencias alélicas y genotípicas:** se determinan por conteo directo a partir de los datos observados, según

$$\text{Frecuencias alélicas } P_i = \frac{x_i}{2N}$$

Donde:

$p_i$  : frecuencia alélica.

$X_i$ : número de veces que se ha observado el alelo "i".

$N$ : número total de individuos en la base de datos.

$$\text{Frecuencias genotípicas } P_{ij} = \frac{g_i}{N}$$

Donde:

$P_{ij}$ : frecuencia genotípica

$g_i$ : número de veces que se observó el genotipo "i"

$N$ : número de genotipos (individuos) de la base de datos.

**b) Heterocigosidad:** la Heterocigosidad observada ( $H_o$ ): se obtiene contando el número de individuos heterocigotas dividido el número total de individuos.

La Heterocigosidad esperada ( $H_e$ ): fue determinada de acuerdo a Nei(1987), como:

$$H = \frac{N}{N-1} \left[ 1 - \sum_{j=1}^k (p_j)^2 \right]$$

Donde  $j_1, j_2, \dots, j_k$  : son los alelos contados en un locus, en una muestra de  $n$  genes de una población, y  $p_j$  es la frecuencia alélica.

**c) Poder de Discriminación (PD):** la capacidad de un sistema para diferenciar entre dos individuos cualesquiera tomados al azar. El poder discriminativo está relacionado con las frecuencias genotípicas del sistema. Se usa en referencia a un marcador genético o combinación de marcadores.

$$PD = 1 - \sum P_{ij}^2$$

Donde  $P_{ij}$ = frecuencia genotípica

**d) Poder de Exclusión (PE):** es la capacidad intrínseca de un sistema de excluir la existencia de vínculo cuando dos personas no están relacionadas biológicamente.

$$PE = Ho^2(1 - 2Ho(1 - Ho))^2$$

Donde Ho es Heterocigosidad observada

**e) Probabilidad de Coincidencia (PC):** es la probabilidad de que dos individuos elegidos al azar tengan un genotipo idéntico en un determinado locus.

$$PC = \sum_{i=1}^n x_i^2$$

Xi Frecuencia genotípica observada para cada locus

**f) Índice de Paternidad Típico (IPT):** es la información del sistema que determina en promedio el IP que puede arrojar. Da una idea de la capacidad intrínseca del sistema en el establecimiento de este tipo de vínculo.

$$IP_t = \frac{1}{2h}$$

En donde IPT: Índice de paternidad típico y h: homocigosidad.

**g) Frecuencia alélica mínima:** si un determinado alelo no se ha detectado en la base de datos se recomienda utilizar la siguiente frecuencia mínima, que corresponde a 5 observaciones en la base de datos de N individuos (NRC II, 1996):

$$F = \frac{5}{2N}$$

**h) Contenido de Información polimórfica (PIC),** índice que evalúa la informatividad de un marcador en la población de acuerdo con las frecuencias de los alelos.

$$PIC = 1 - \sum (P^2 i)j$$

Donde Pi es la frecuencia alélica del i-esimo alelo, j es el locus

# **IV - RESULTADOS**

## IV-RESULTADOS

A continuación se describen los resultados obtenidos del análisis de polimorfismos genéticos realizado sobre la población de Santa Cruz.

### IV-1. Resultados de STRs autosómicos

El análisis fue realizado sobre muestras de individuos no relacionados genéticamente, que, de acuerdo a la organización judicial y localización geográfica, fueron divididos en tres grupos, denominados: C1E (Primera Circunscripción Este); C1O (Primera Circunscripción Oeste) y C2 - Población perteneciente a la Segunda Circunscripción. Estas muestras fueron tipificadas mediante la amplificación de 15 STRs autosómicos.

A continuación se describen los resultados teniendo en cuenta los objetivos planteados.

**Primer Objetivo:** Describir la composición genética de la población santacruceña en estudio, analizando los parámetros genéticos poblacionales de los tres grupos para establecer si existen diferencias significativas. Si no existen, entonces se trabajará con las tres regiones agrupadas constituyendo una única base de datos definida como “Santa Cruz” (SC).

Para cumplir el primer objetivo se analizaron y compararon la distribución de frecuencias alélicas de los STRs autosómicos en los tres grupos regionales; se determinaron parámetros moleculares (heterocigosidad observada, índice de diversidad génica) y se determinó la distancia genética entre estas poblaciones así divididas. La Tabla 1 exhibe los resultados obtenidos.

Tabla 1: Índices moleculares en los tres grupos regionales de la provincia de Santa Cruz

	Het observada promedio (s.d)	Het esperada promedio (s.d)	Diversidad génica promedio	Test de Garza Williamson promedio (s.d)
C1E	0.78122 (0.05928)	0.79288 (0.05354)	0.792881 +/- 0.399468	0.72908 (0.4128)
C1O	0.81270(0.13409)	0.78544 (0.07379)	0.785443 +/- 0.402897	0.77105 (0.36798)
C2	0.77692(0.07189)	0.79427 (0.06041)	0.794274 +/- 0.401184	0.78451 (0.37111)

*Test de Garza Williamson ( $M < 0.68$ )*

La heterocigosidad observada y esperada, diversidad génica y test de G-W son resultados promediados considerando los 15 STRs autosómicos analizados.

El test de diferenciación exacto basado en la distribución de frecuencias alélicas arrojó resultados no significativos (dato no mostrado) en la comparación de las tres subpoblaciones.

La distancia genética entre las tres subpoblaciones fue no significativa.

Conclusión: de acuerdo a los resultados obtenidos no existen diferencias significativas entre los tres grupos y como consecuencia se incluirán las tres subpoblaciones en una única

base de datos denominada “Santa Cruz” (SC).

**Segundo Objetivo:** Analizar parámetros genético poblacionales en la población en estudio (SC).

Se determinaron distribución de frecuencias alélicas, Heterocigosidad observada y esperada, test de Equilibrio de Hardy-Weinberg y los parámetros de interés forense (Poder discriminativo, Poder de Exclusión, Índice de Paternidad Típico, Contenido de Información Polimórfica y Probabilidad de Coincidencia).

Estos parámetros y la distribución de frecuencias alélicas son resumidos en la Tabla 2.

Tabla 2 Distribución de frecuencias alélicas, parámetros genéticos poblacionales y forenses en la población de Santa Cruz.

Allele/n	D8S1179	D21S11	D7S820	C SF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	VWA	TPOX	D18S51	D5S818	FGA
6						0,2655						0,0035			
7			0,0069			0,2628						0,0035		0,1069	
7.2										0,0035					
8	0,0172		0,1000	0,0172		0,0897	0,0483	0,0172				0,4690		0,0035	
9	0,0069		0,0690	0,0172		0,1069	0,2138	0,1655				0,0724	0,0035	0,0241	
9.3						0,2379									
10	0,0483		0,3000	0,2931		0,0172	0,0621	0,1241				0,0517	0,0069	0,0483	
11	0,0621		0,3069	0,2483			0,2586	0,2517		0,0069	0,0069	0,3207	0,0172	0,3897	
12	0,1379		0,1828	0,3448			0,2655	0,2552		0,0655		0,0759	0,1000	0,2862	
12.1			0,0035												
12.2										0,0069					
13	0,3379		0,0310	0,0621			0,0931	0,1414		0,2138		0,0035	0,1310	0,1414	
13.2										0,0759					
14	0,2035			0,0138	0,0690		0,0586	0,0448		0,2483	0,0414		0,2310		
14.2										0,0310					
15	0,1759			0,0035	0,3517					0,2310	0,0897		0,1000		
15.2										0,0621					
16	0,0069				0,2552				0,0276	0,0448	0,3103		0,0897		
16.2										0,0069					
17	0,0035				0,1448				0,2379		0,2897		0,1655	0,0069	
17.2										0,0035					
18					0,1690				0,0628		0,2000		0,0655		0,0172
19					0,0103				0,1931		0,0552		0,0345		0,1035
20									0,1483	0,0069			0,0345	0,0690	
21									0,0310				0,0035		0,1552
22									0,0345				0,0069		0,0966
23									0,1207				0,0035		0,1103
24									0,0724						0,1586
25									0,0310						0,1448
26		0,0035							0,0207						0,0759
27		0,0138											0,0069		0,0379
28		0,0828													0,0172
28.2		0,0035													
29		0,1828													
30		0,2586													0,0035
30.2		0,0276													
31		0,0690													
31.2		0,1414													0,0035
32		0,0103													
32.2		0,1517													
33		0,0035													
33.2		0,0483													
35.2		0,0035													
N	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145
Fm	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172	0,0172
Na	10	14	8	8	6	6	7	7	11	13	8	8	16	7	14
Ho	0,7724	0,8000	0,7724	0,7241	0,7103	0,8069	0,8138	0,8207	0,8621	0,8089	0,8345	0,6207	0,8621	0,7100	0,8552
He	0,7906	0,8446	0,7693	0,7314	0,7594	0,7759	0,8014	0,8092	0,8561	0,8251	0,7896	0,6658	0,8699	0,7345	0,8880
HWE (p)	0,7290	0,5788	0,4667	0,5332	0,2842	0,1464	0,1940	0,0263	0,3456	0,3661	0,2607	0,1419	0,4845	0,6009	0,2948
PC	0,0769	0,0455	0,0970	0,1284	0,1005	0,0972	0,0761	0,0827	0,0457	0,0572	0,1001	0,1704	0,0364	0,1070	0,0292
PD	0,9231	0,9545	0,9030	0,8716	0,8995	0,9028	0,9239	0,9173	0,9543	0,9428	0,8999	0,8296	0,9636	0,8930	0,9708
PIC	0,7590	0,8232	0,7304	0,6807	0,7185	0,7365	0,7697	0,7786	0,8370	0,7994	0,7306	0,6093	0,8533	0,6903	0,8728
PE	0,5489	0,5990	0,5489	0,4666	0,4444	0,6119	0,6249	0,6380	0,7188	0,6119	0,6645	0,3164	0,7188	0,4444	0,7031
IPT	2,1970	2,500	2,1970	1,8125	1,7262	2,5893	2,6852	2,7885	3,6250	2,5893	3,0208	1,3182	3,6250	1,7262	3,4286

Referencias: N= N° de individuos, Fm: frecuencia mínima; Na = N° de alelos diferentes; Ho = Heterocigosidad observada; He = Heterocigosidad esperada; HWE (P)-Equilibrio de Hardy-Weinberg (significancia); PC=Probabilidad de coincidencia; PD= Poder de discriminación; PIC=Contenido de información Polimórfica; PE=Probabilidad de Exclusión; IPT=Índice de Paternidad Típico

A continuación se presentan los gráficos de distribución de frecuencias alélicas para cada marcador.

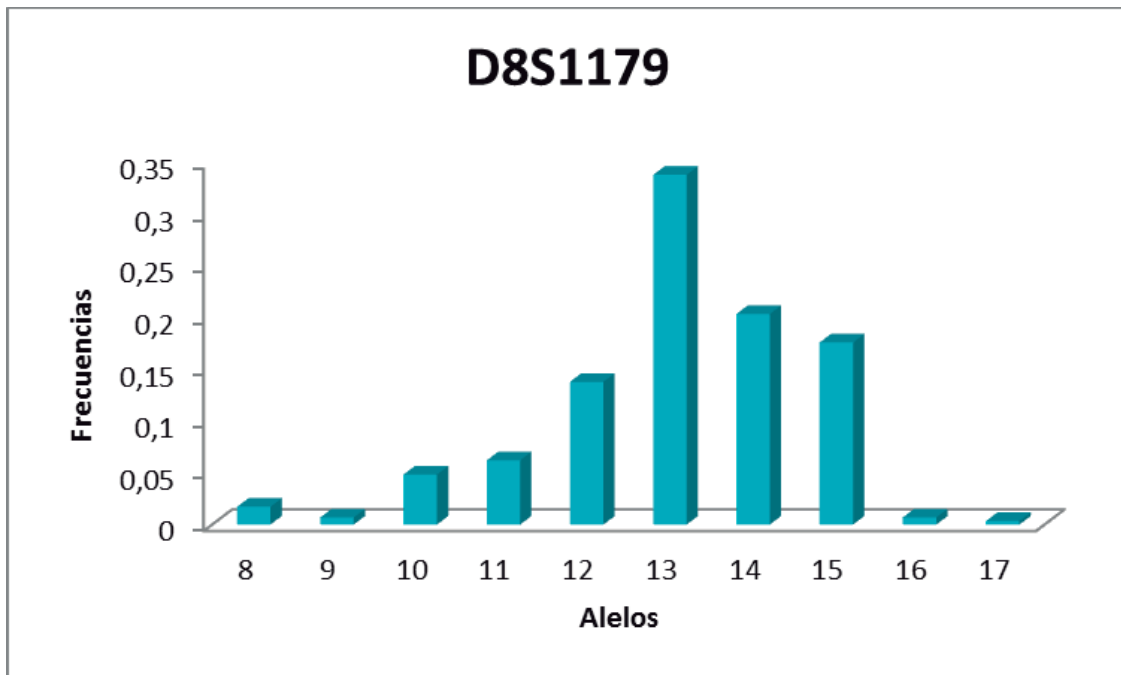


Figura 10 Distribución de las Frecuencias alélicas para el marcador D8S1179

Para el marcador D8S1179 se observó 10 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelo 17) y 0.3379 (alelo 13).

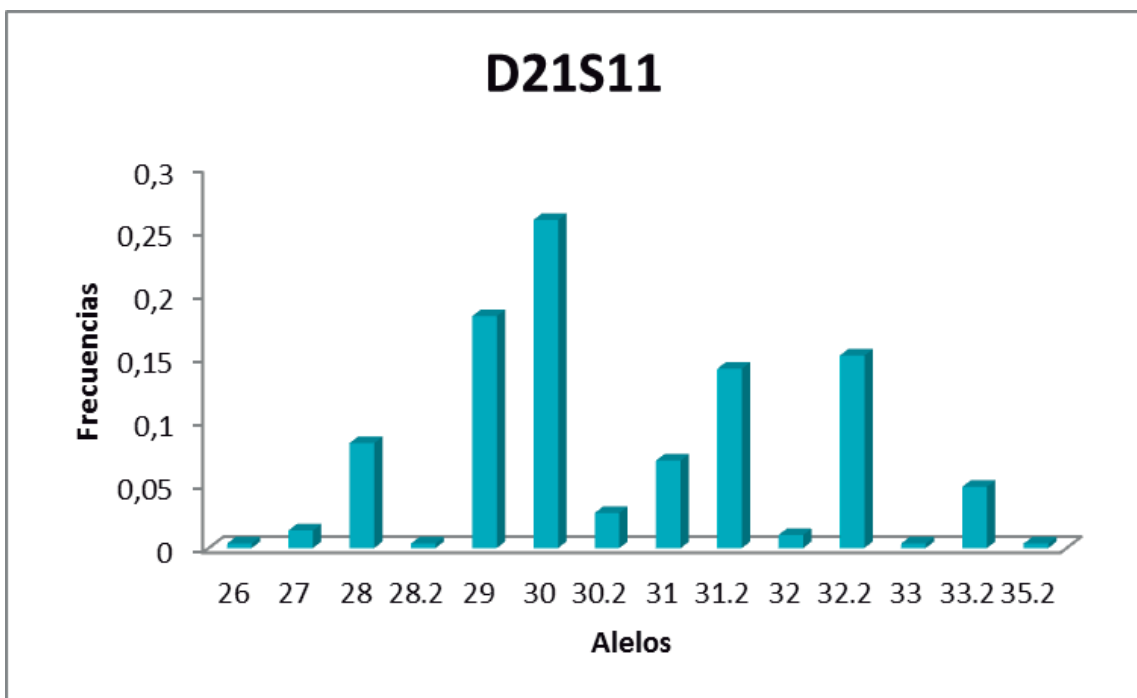


Figura Distribución de las Frecuencias alélicas para el marcador D21S11

Para el marcador D21S11 se observó 14 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelos 26, 28.2, 33 y 35.2) y 0,2586 (alelo 30).

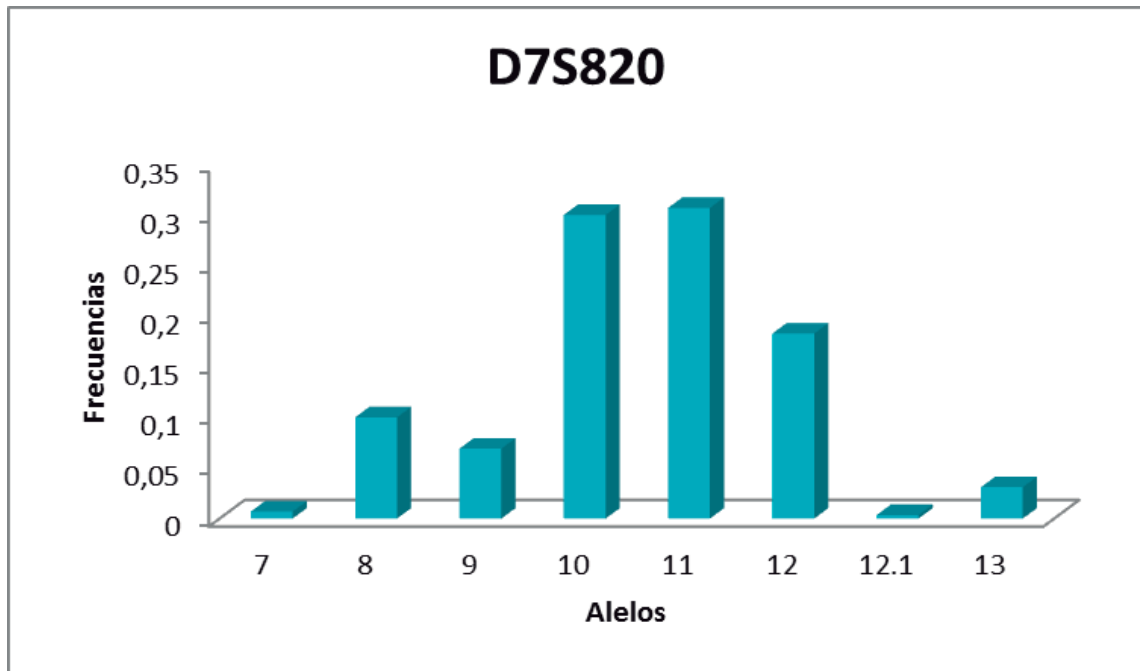


Figura 12 Distribución de las Frecuencias alélicas para el marcador D7S820

Para el marcador D7S820 se observó 8 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelo 12.1) y 0,3069 (alelo 11).

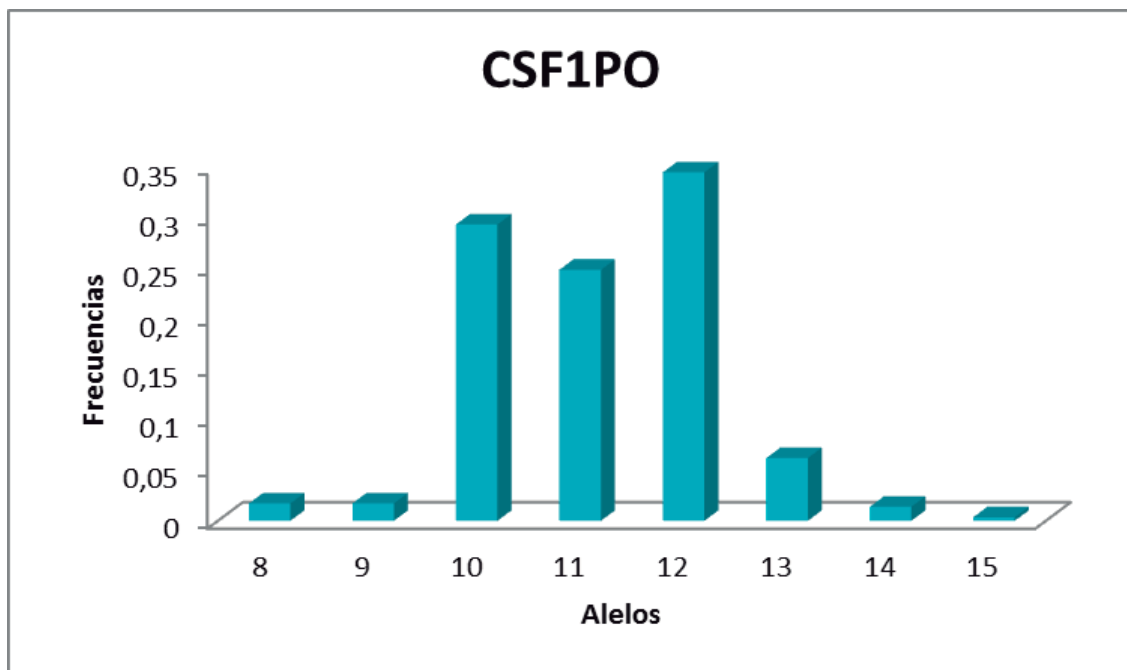


Figura 13 Distribución de las Frecuencias alélicas para el marcador CSF1PO

Para el marcador CSF1PO se observó 8 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelo 15) y 0,3448 (alelo 12).

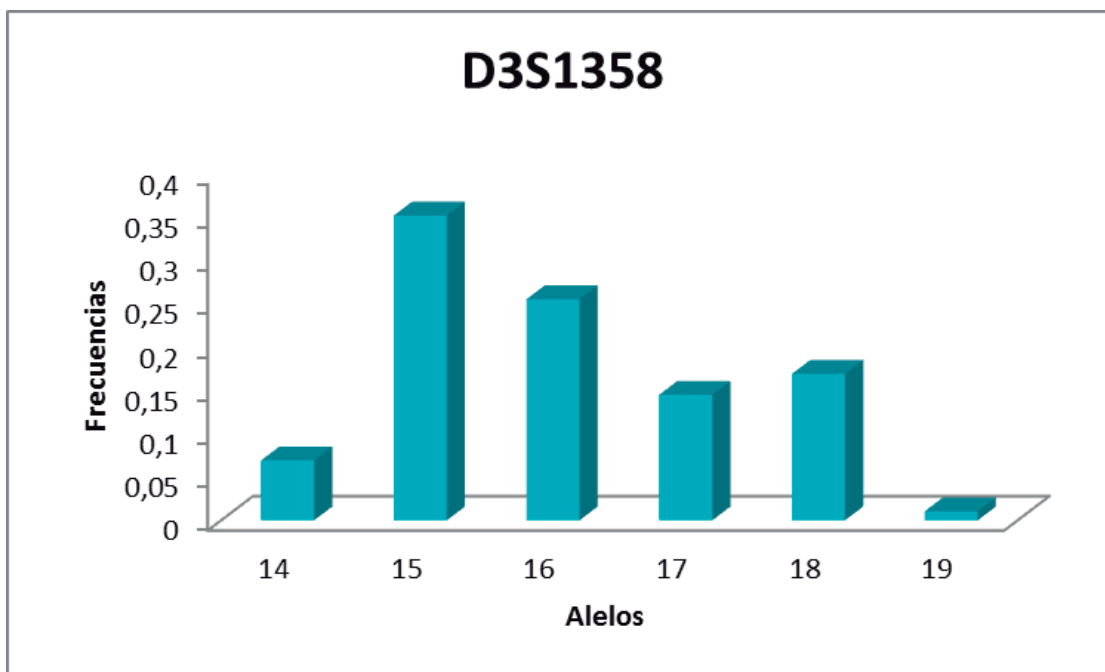


Figura 14 Distribución de las Frecuencias alélicas para el marcador D3S1358

Para el marcador D3S1358 se observó 6 alelos distintos. Las frecuencias oscilan entre 0,0103 (alelo 19) y 0,3517 (alelo 15).

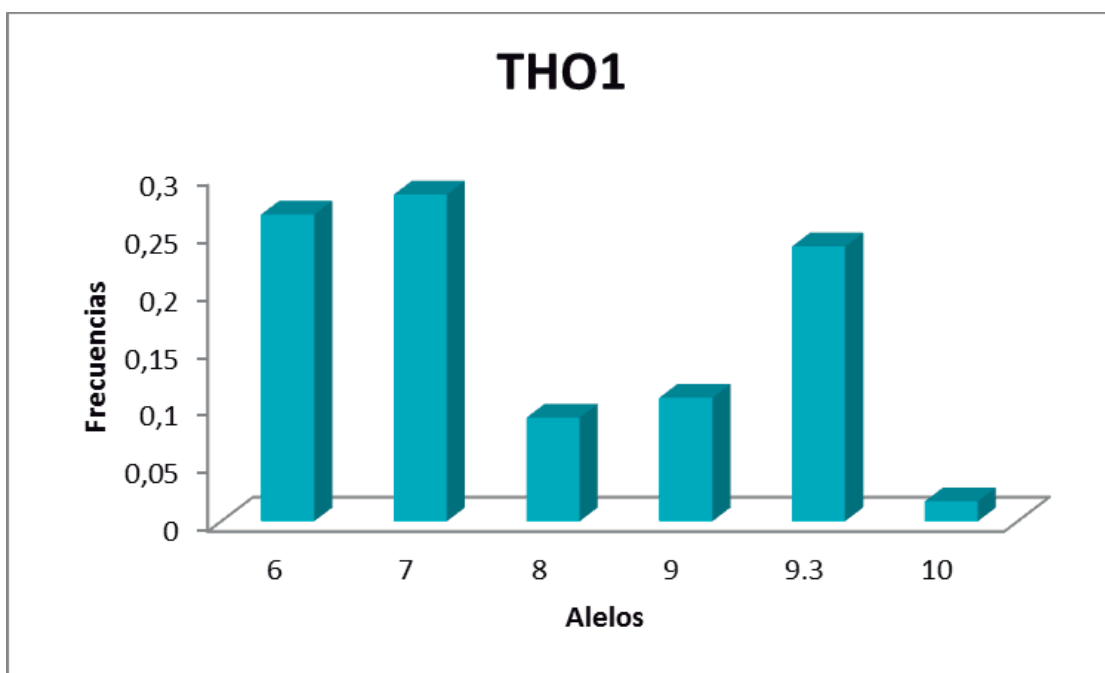


Figura 15 Distribución de las Frecuencias alélicas para el marcador THO1

Para el marcador THO1 se observó 6 alelos distintos. Las frecuencias oscilan entre 0,0172 (alelo 10) y 0,2828 (alelo 7).



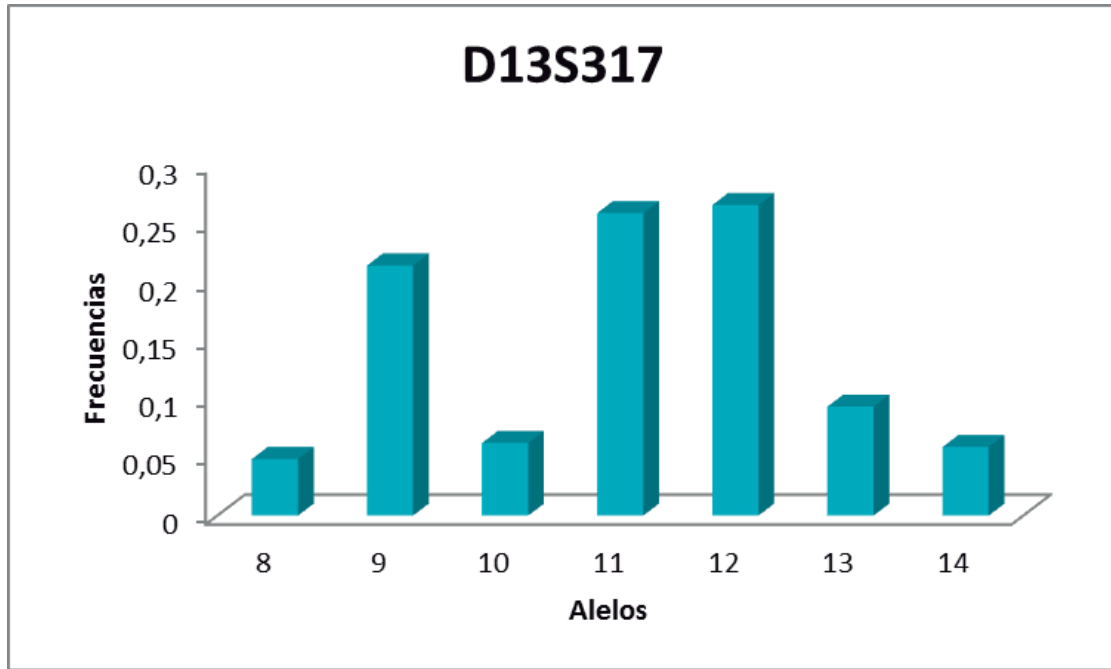


Figura 16 Distribución de las Frecuencias alélicas para el marcador D13S317

Para el marcador D13S317 se observó 7 alelos distintos. Las frecuencias oscilan entre 0,0483 (alelo 8) y 0,2655 (alelo 12).

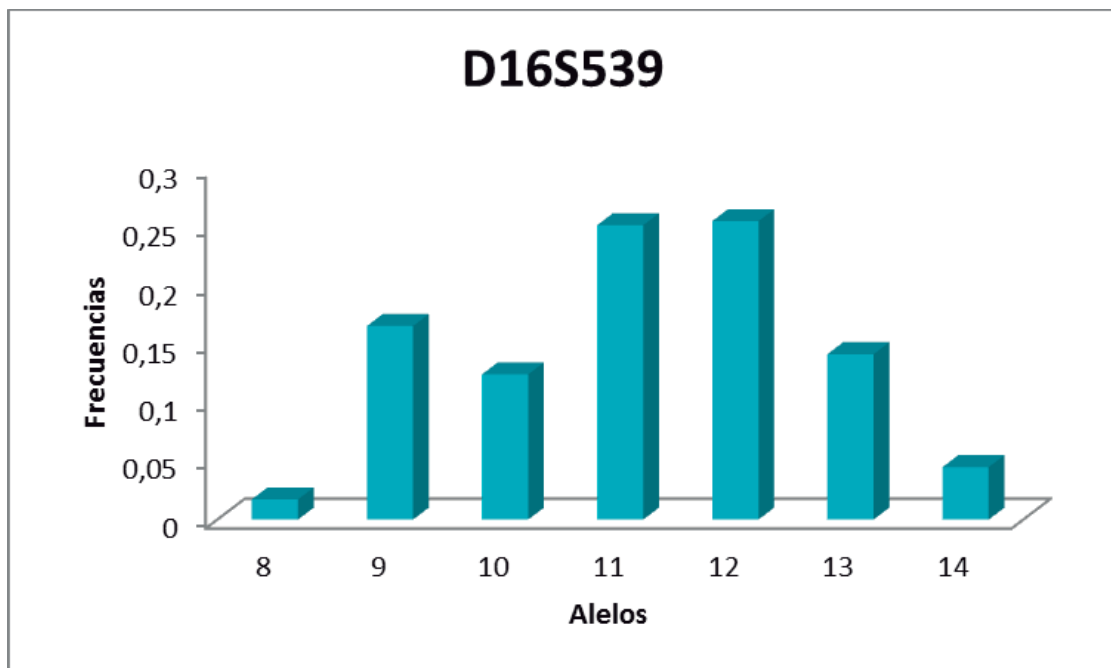


Figura 17 Distribución de las Frecuencias alélicas para el marcador D16S539

Para el marcador D16S539 se observó 7 alelos distintos. Las frecuencias oscilan entre 0,0172 (alelo 8) y 0,2552 (alelo 12).

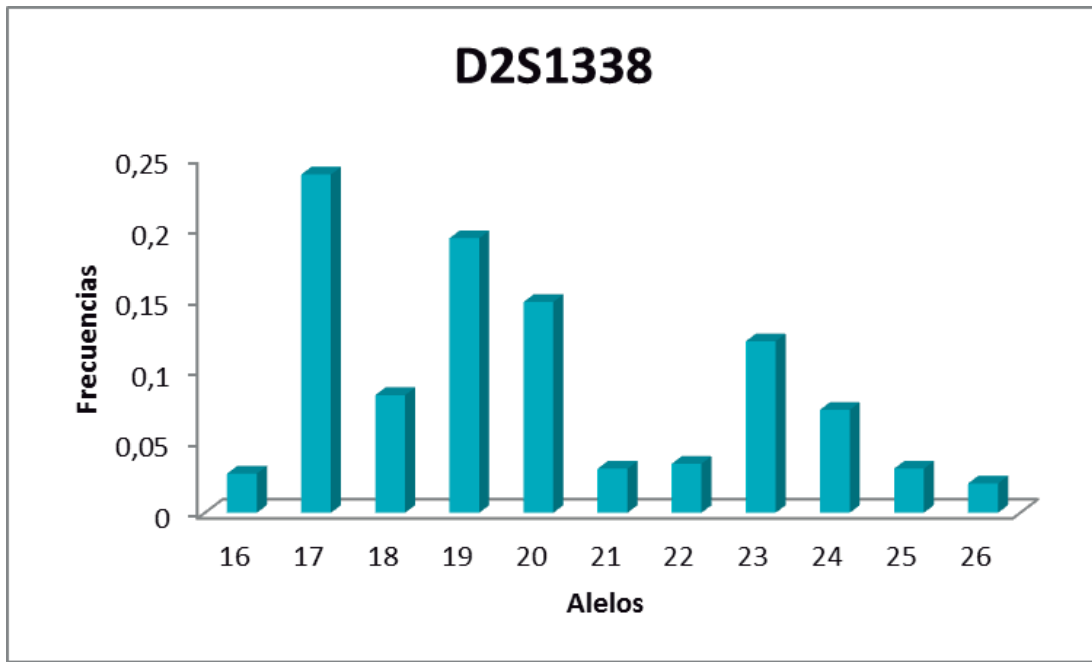


Figura 18 Distribución de las Frecuencias alélicas para el marcador D2S1338

Para el marcador D2S1338 se observó 11 alelos distintos. Las frecuencias oscilan entre 0.0207 (alelo 26) y 0,2379 (alelo 17).

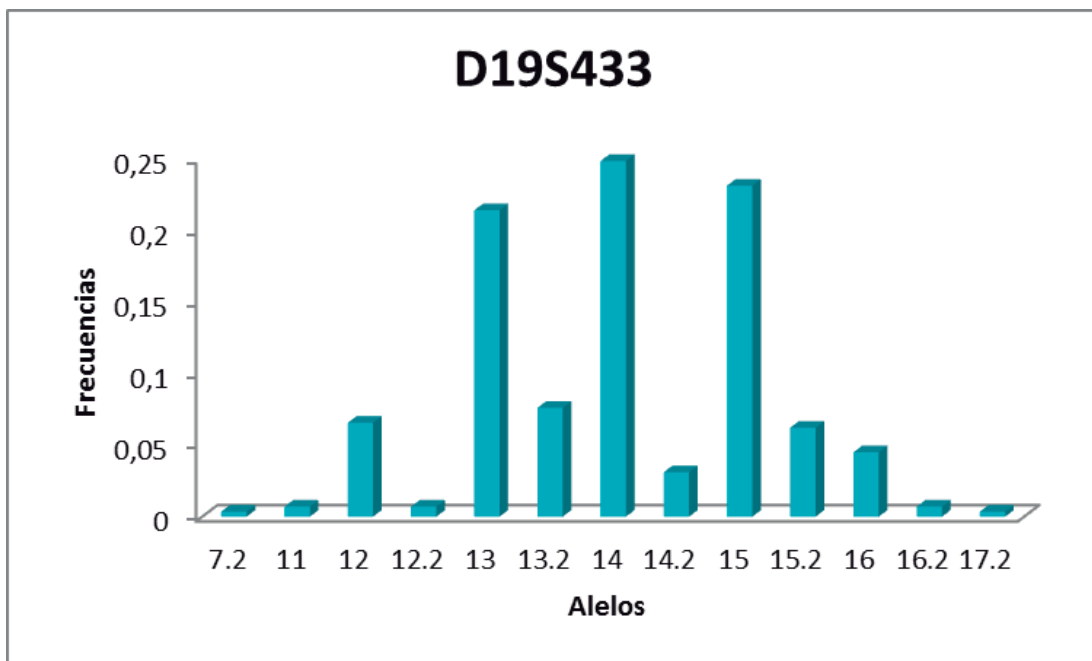


Figura 19 Distribución de las Frecuencias alélicas para el marcador D19S433

Para el marcador D19S433 se observó 13 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelos 7.2, 17.2) y 0,2483 (alelo 14).

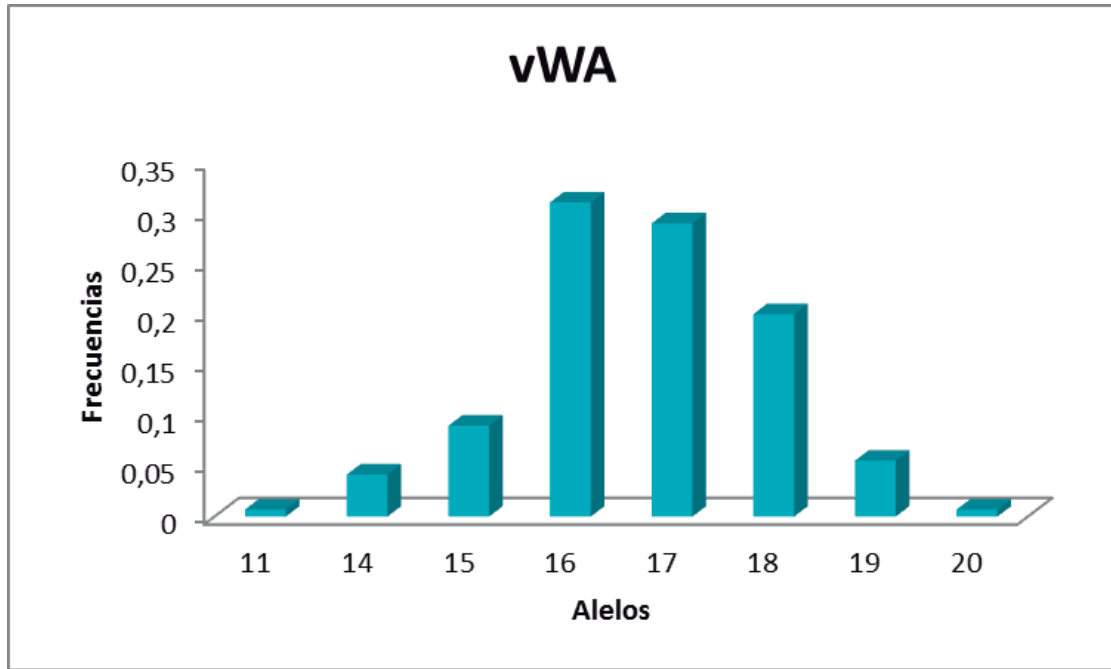


Figura 20 Distribución de las Frecuencias alélicas para el marcador vWA

Para el marcador vWA se observó 8 alelos distintos. Las frecuencias oscilan entre 0.0069 (alelos 11, 20) y 0,3103 (alelo 16).

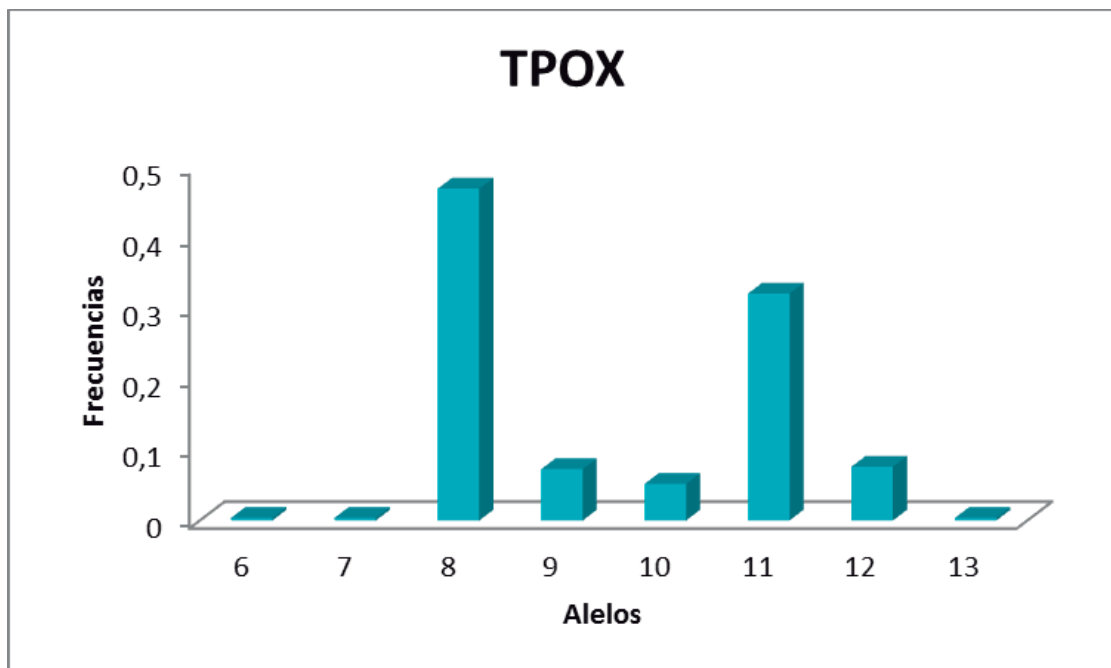


Figura 21 Distribución de las Frecuencias alélicas para el marcador TPOX

Para el marcador TPOX se observó 8 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelos 6, 7, 13) y 0,4690 (alelo 8).

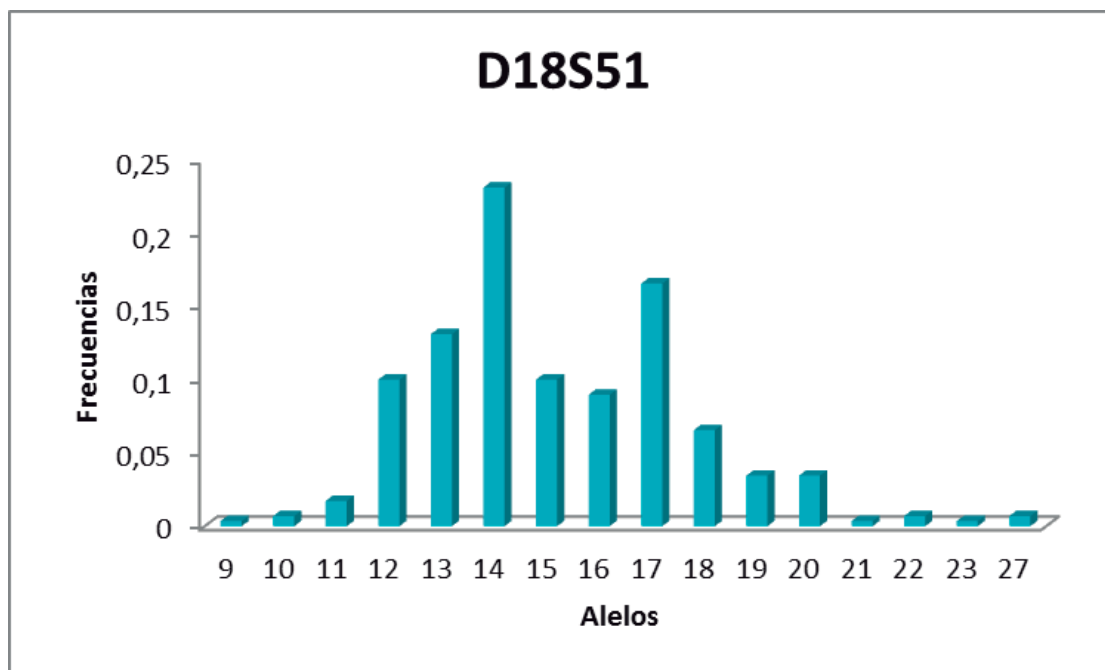


Figura 22 Distribución de las Frecuencias alélicas para el marcador D18S51

Para el marcador D18S51 se observó 16 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelos 9, 21, 23) y 0,2310 (alelo 14).

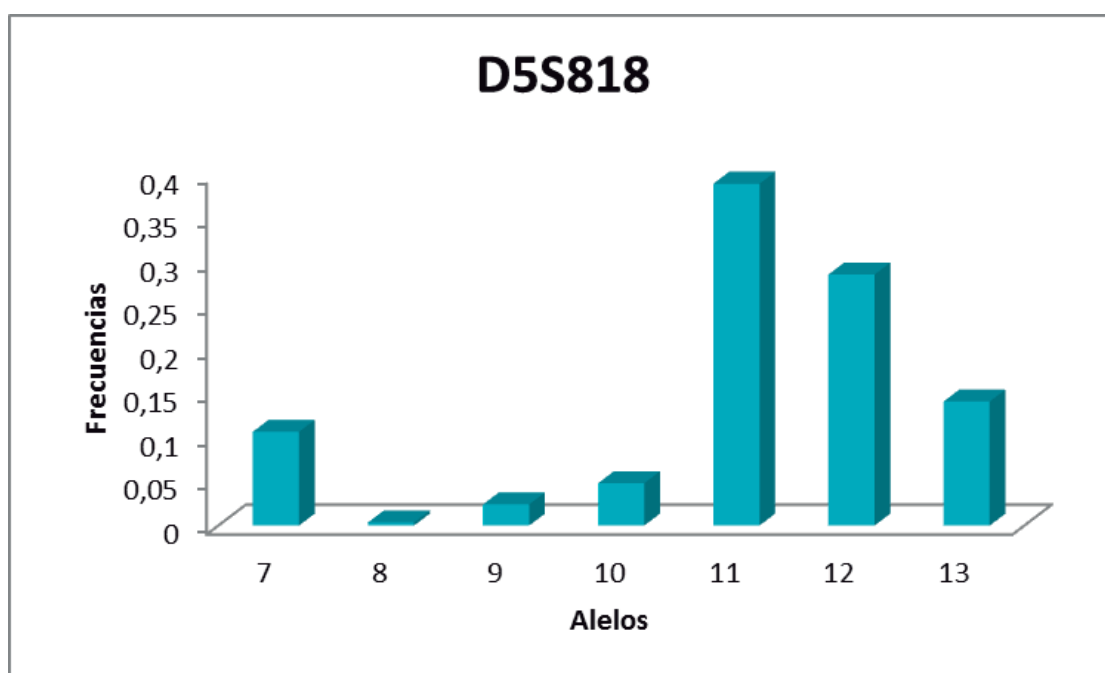


Figura 23 Distribución de las Frecuencias alélicas para el marcador D5S818

Para el marcador D5S818 se observó 7 alelos distintos. Las frecuencias oscilan entre 0.0035 (alelo 8) y 0,3897 (alelo 11).

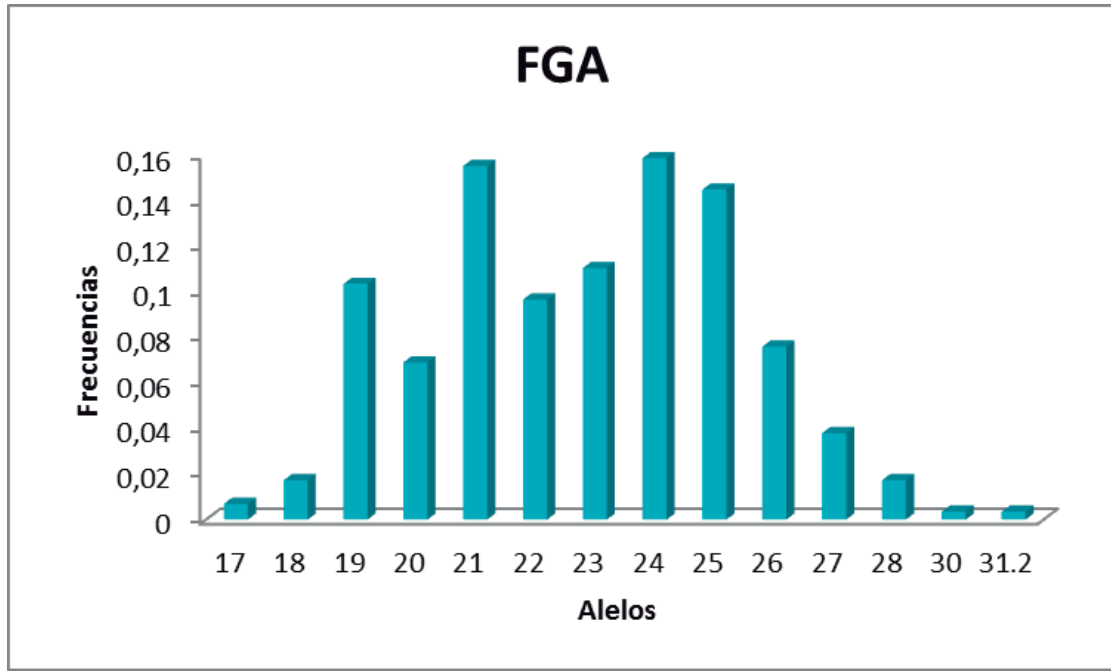


Figura 24 Distribución de las Frecuencias alélicas para el marcador FGA

Para el marcador FGA se observó 14 alelos distintos. Las frecuencias oscilan entre 0,0035 (alelos 30, 31.2) y 0,1586 (alelo 24).

El gráfico a continuación (Figura 25) presenta las frecuencias más altas observados en la población de Santa Cruz en cada uno de los STRs analizados.

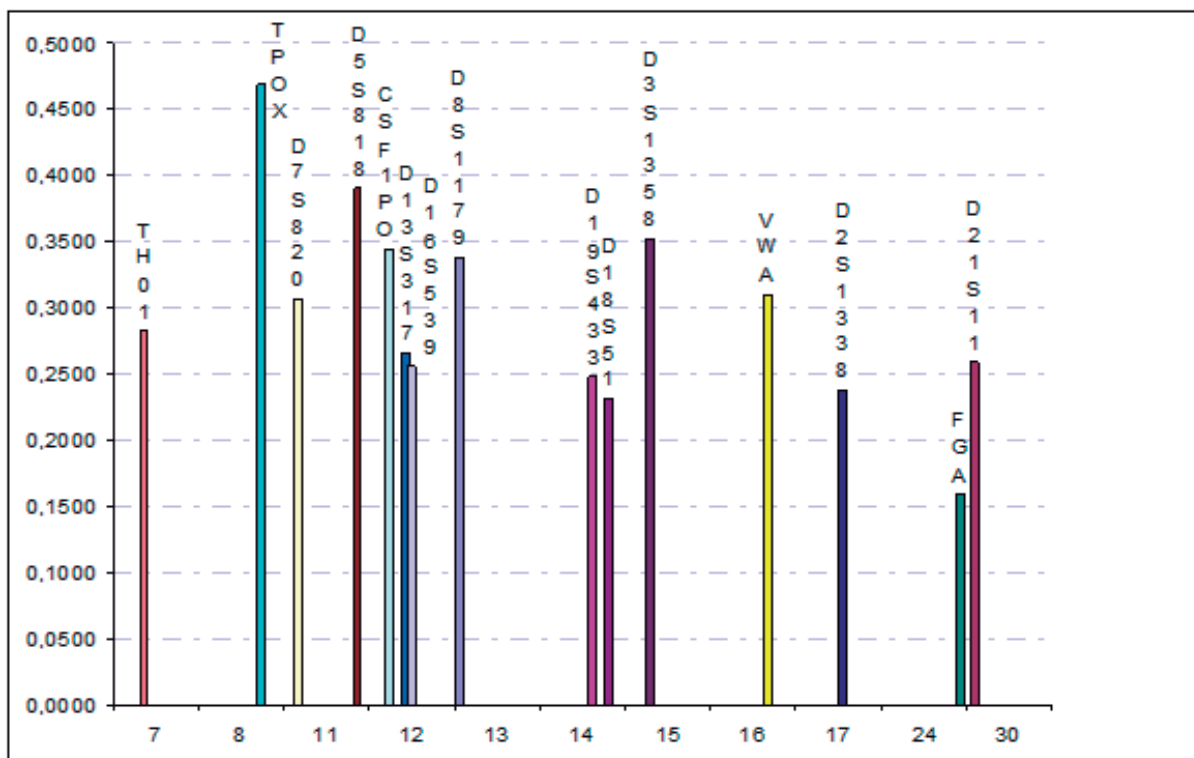
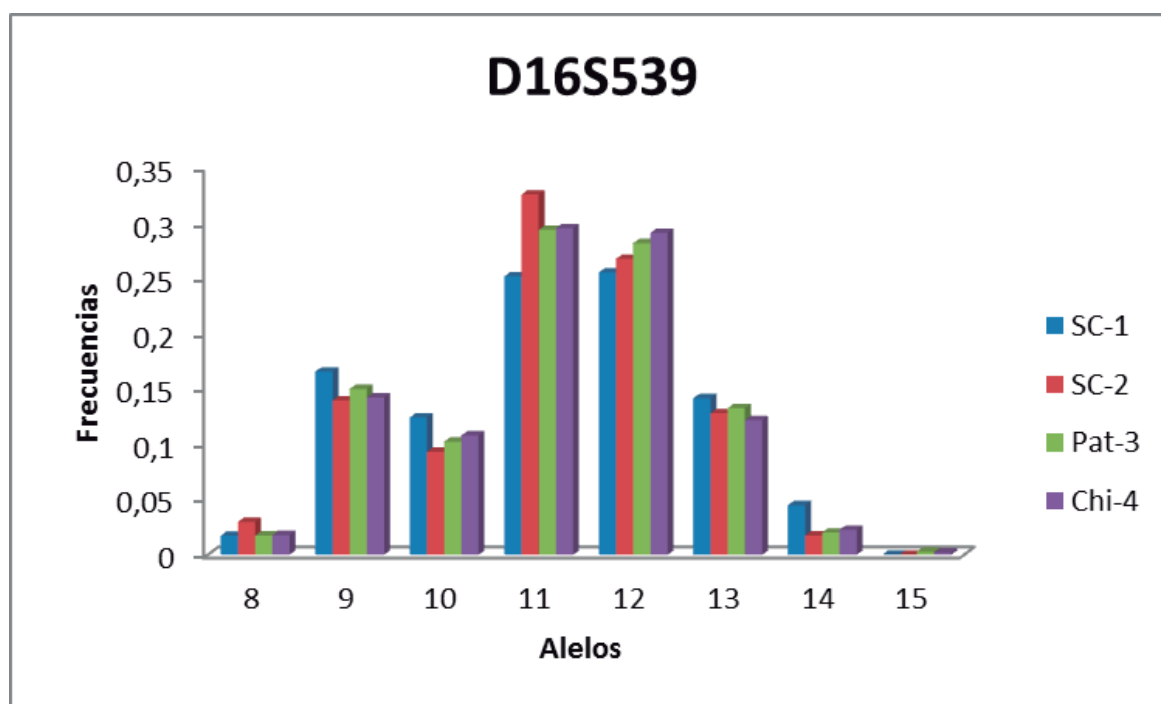


Figura 25 Distribución de frecuencias alélicas considerando únicamente los alelos más frecuentes para cada loci investigado.

La heterocigosidad observada ( $H_o$ ) promedio para la población total es de 0,7848, en la que todos los marcadores exhiben valores superiores al 71% excepto el sistema TPOX en el que la  $H_o$  fue de 62%, en concordancia con las características exhibidas por este marcador en otras poblaciones analizadas (Marino, M 2006 a, b, c). Los marcadores que exhibieron los porcentajes más altos de  $H_o$  fueron D18S51, D2S1338 y FGA.

En cuanto al análisis de equilibrio, 14 de los 15 marcadores estudiados en la población de Santa Cruz se encuentran en equilibrio de Hardy-Weinberg (EHW) ( $p > 0,05$ ), mientras que el sistema D16S539 se encontraría fuera del EHW ( $p < 0,05$ ). De acuerdo a Bonferroni (1936) trabajando con un nivel de significancia del 0.05, y considerando que el número total de sistemas es 15, existe un 54% de probabilidad de observar al menos un resultado significativo (aunque todos los test hubieren resultado no significativos). Ello permite, según el autor, realizar un ajuste al nivel de significancia. En este caso este valor es  $\alpha/n = 0.05/15 = 0.0033$ . Considerando este nivel de significancia el sistema mencionado también se encontraría dentro del equilibrio. Habiendo descartado posibles errores en la asignación alélica y la presencia de individuos relacionados en la base de datos, una distribución de frecuencias particular de este marcador podría atribuirse a un efecto local producto de las migraciones experimentadas en la región.

A continuación se presenta la distribución de frecuencias alélicas para este marcador en comparación con otros trabajos publicados (Borosky et al, 2014; Marino et al, 2006; Vergara et al 2012) *Figura 26*.



**Figura 26** Distribución comparativa de frecuencias para el marcador D16S539. Referencias: SC-1, población de Santa Cruz, presente trabajo  $N=145$ ; SC-2: población de Santa Cruz, Borosky et al, 2014,  $N=86$ ; Pat-3: población de Chubut y Río Negro, Marino et al, 2006,  $N=577$ ; Chi-4: población general de Chile, Vergara et al, 2012,  $N=732$

Es de destacar que en el trabajo de Vergara et al, realizado sobre un total de 732 individuos procedentes de distintas áreas geográficas de Chile, el sistema D16S539 también se encuentra fuera de equilibrio.

En el trabajo publicado por Borosky y colaboradores este sistema se encuentra en equilibrio, si bien el número de individuos es inferior y no se indica de que áreas geográficas proceden las muestras analizadas.

La ampliación de la presente base de datos permitirá corroborar el comportamiento de este marcador en la población de SC.

#### **IV-1.1. Análisis de parámetros estadísticos de interés forense (tabla 2):**

El marcador con el mayor índice de probabilidad de coincidencia (PC), es el TPOX (0,1704), seguido de CSF1PO (0,1284), mientras que el marcador con el índice más bajo es el FGA (0,0292) seguido del el D18S51 (0,0364) y del D21S11 (0,0455). La probabilidad de que dos individuos elegidos al azar tengan un genotipo idéntico (probabilidad de identidad o “match probability”) a lo largo de los 15 STRs analizados es  $1,35857E-17$  (equivalente a 1 en  $7,3607E+16$ )

Para el Poder de Discriminación (PD), el marcador que presenta el índice más alto es el FGA (0,9708), seguido del D18S51 (0,9636) y del D21S11 (0,9545), mientras que el marcador con el índice de PD más bajo es el TPOX (0,8296), seguido del CSF1PO (0,8716).

Por otra parte, en relación al contenido de información polimórfica (PIC) se observa que el marcador más informativo con el índice más alto es el FGA (0,8728), seguido del D18S51 (0,8533), mientras que el marcador que presenta el índice más bajo es el TPOX (0,6093), seguido de CSF1PO (0,6807). Valores de PIC inferiores a 0,5 resultan poco deseables en los test de parentesco.

En cuanto al Poder de Exclusión (PE), los marcadores más informativos son D2S1338 y D18S51 (0,7188 en ambos) seguido del FGA (0,7031), ya que son los más polimórficos permitiendo excluir a priori eficientemente en estudios de paternidad, mientras que el marcador menos informativo es el TPOX (0,3164), seguido del D3S1358 y D5S818 (0,4444 en ambos casos). Estos valores son concordantes con los IPT obtenidos. La probabilidad combinada de exclusión considerando los 15 STRs es superior al 99,999%.

#### **IV-2. Comparaciones interpoblacionales**

Se realizaron comparaciones interpoblacionales considerando las siguientes poblaciones:

- Población general de las provincias de Chubut y Río Negro, a fin de establecer comparaciones con otras provincias patagónicas (CHU, N=100; RN, N=100)

- Población general de Buenos Aires: esta población fue seleccionada por la historia demográfica que presenta el poblamiento de Santa Cruz (BA, N=100)

Estas bases de datos fueron previamente analizadas y publicadas por el Servicio de Huellas Digitales Genéticas (Marino et al., 2006 b,c; Sala y Corach, 2014).

Para estas comparaciones interpoblacionales se seleccionó un core de 13 STRs autosómicos, que son los comprendidos en el CODIS (Combined DNA IndexSystem), dado que han demostrado ser un conjunto muy informativo en el análisis de estructura de poblaciones (Budowle et al., 1999). Los STRs incluidos en éste core son: D3S1358, TH01, D21S11, D18S51, D5S818, D16S539, D7S820, D13S317, CSF1PO, VWA, D8S1179, TPOX y FGA.

En la Figura 27 se presenta en escala de tres colores diferentes el número promedio de diferencias entre pares de muestras en las poblaciones de Buenos Aires (BA); Chubut (CHU), Río Negro (RN) y Santa Cruz (SC). En este gráfico la densidad de color aumenta a medida que la diferencia es mayor. En la diagonal, en color naranja, se observa las diferencias dentro de cada población; sobre la diagonal, en verde, entre pares de poblaciones y debajo de la diagonal, en azul, el número neto de diferencias entre poblaciones (observación: considerar que si bien la intensidad de colores puede remitir a diferencias muy grandes, la escala es muy pequeña).

Como puede observarse las diferencias son máximas entre individuos dentro de cada población (línea en diagonal, representada en color naranja), mientras que las diferencias entre poblaciones (representado en verde) mayores son observadas entre las poblaciones de Buenos Aires respecto de Río Negro y Santa Cruz.

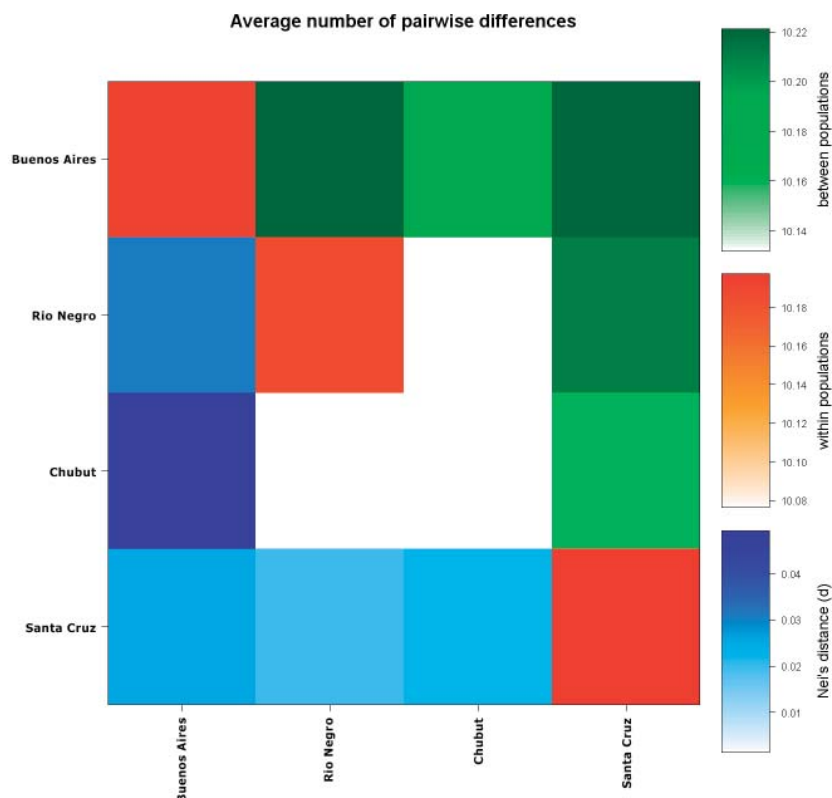


Figura 27 Número promedio de diferencias entre pares dentro y entre poblaciones.



Se determinaron las distancias genéticas existentes entre los grupos analizados. Las Tablas 3 y 4 muestran las matrices de distancia genéticas ( $F_{st}$ ) y de significancia, respectivamente.

**Tabla 3:** Distancias genéticas ( $F_{st}$ )

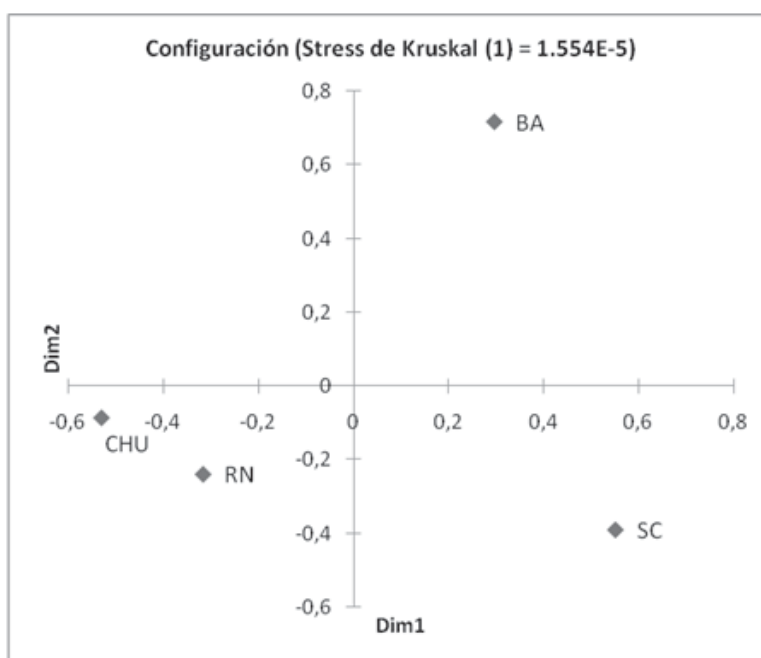
	BA	RN	Chu	SC
BA	0.00000			
RN	0.00307	0.00000		
Chu	0.00487	0.00013	0.00000	
SC	0.00254	0.00197	0.00218	0.00000

**Tabla 4:** Matriz de significancia (número de permutaciones 1000)

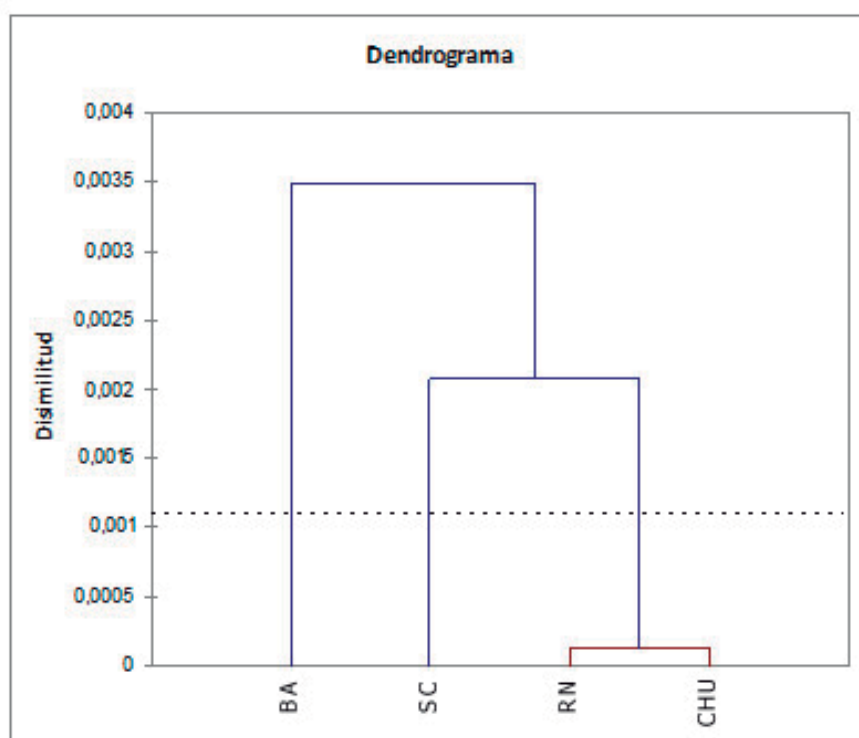
	BA	RN	Chu	SC
BA	-			
RN	0.00684+-0.0023*	-		
Chu	0.00000+-0.0000*	0.42480+-0.0171	-	
SC	0.00977+-0.0029*	0.01562+-0.0034*	0.00977+-0.0029*	-

\*Significativo ( $p < 0.05$ ). Número de permutaciones: 1000

Las distancias genéticas obtenidas fueron todas significativas, excepto la distancia entre Río Negro y Chubut, en concordancia con investigaciones previas (Sala y Corach, 2014). Como puede observarse en el gráfico de Multi Dimensional Scaling y Dendograma (Figuras 28 y 29), estas dos poblaciones constituyen un cluster, mientras que la población de Santa Cruz se encontraría en una posición relativamente equidistante entre éste y la población de Buenos Aires.



**Figura 28** Representación MDS (multi dimensional scaling) de las distancias genéticas ( $F_{st}$ , número de permutaciones 1000) entre las poblaciones de Buenos Aires (BA); Chubut (CHU); Río Negro (RN) y Santa Cruz (SC).



**Figura 29** Dendrograma de representación de las distancias genéticas entre las poblaciones de Buenos Aires (BA); Chubut (CHU); Río Negro (RN) y Santa Cruz (SC).

Finalmente, se realizó un Test de AMOVA (Analysis of Molecular Variation, Excoffier et al., 1992), en el que se consideraron los siguientes grupos: Grupo 1 (Buenos Aires), Grupo 2 (Río Negro +Chubut) y Grupo 3 (Santa Cruz). Los niveles jerárquicos de análisis fueron:

FST: variación dentro de las poblaciones respecto a la varianza total

FSC: variación entre poblaciones dentro de un grupo

FCT: varianza entre los grupos con respecto a la varianza total

La Tabla 5 presenta los resultados obtenidos.

**Tabla 5:** Resultados de AMOVA (número de permutaciones 1000)

Origen variación	Componente de variación	Porcentaje de variación	de	Indices de fijación	de	Significancia (p)
Entre grupos	0.01346 Va	0.26		FCT	0.00264	0.17595+- 0.00998
Entre poblaciones dentro de un grupo	0.00058 Vb	0.01		FSC	0.00011	0.41153+- 0.01445
Dentro de poblaciones	5.08273 Vc	99.72		FST	0.00276	0.00000+- 0.00000

Conclusión de AMOVA: el porcentaje de variación indica que la variación máxima (99,72%) puede ser atribuida a la variación entre los individuos dentro de una población (Fst). Respecto de los valores de FSC y FCT, variación entre grupos y entre poblaciones, ambos resultaron no significativos.

### IV-3. Análisis de estructura

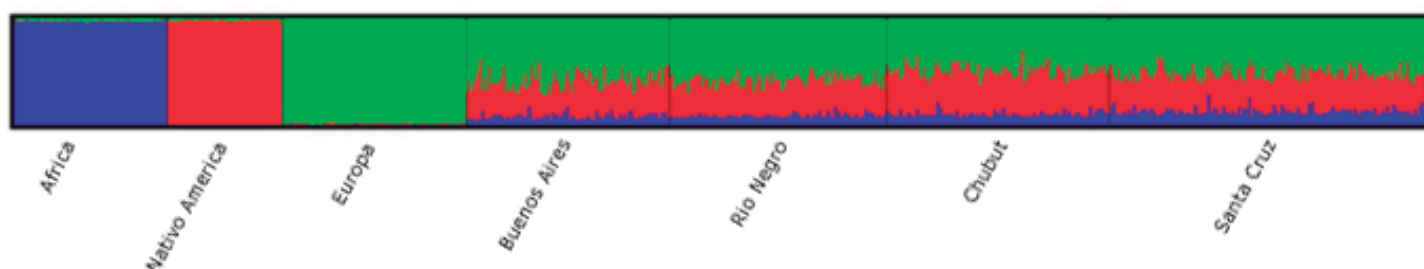
Para el análisis de estructura se empleó un panel de muestras del CEPH, que corresponden a individuos de tres grupos étnicos definidos: Africano, Europeo y Nativo americano. Además se incluyeron las muestras correspondientes a las poblaciones de Buenos Aires, Río Negro, Chubut y Santa Cruz.

Los resultados obtenidos se presentan en la Tabla 6 y la Figura 30.

Los resultados obtenidos indican que si bien las poblaciones tienen una composición genética similar, el componente africano, minoritario en todos los grupos, se presenta en mayor proporción en la población de Santa Cruz.

**Tabla 6** Composición Genética de las poblaciones

	Componente africano	Componente Nativo americano	Componente Europeo
Buenos Aires	0.084	0.338	<b>0.579</b>
Río Negro	0.091	0.339	<b>0.571</b>
Chubut	0.103	0.418	<b>0.479</b>
<b>Santa Cruz</b>	<b>0.130</b>	<b>0.363</b>	<b>0.507</b>



**Figura 30:** Bar Plot obtenido a partir del análisis de estructura.

Población 1: panel del CEPH, Africano; 2: CEPH, Nativo americano; 3: CEPH, Europeo; 4: Buenos Aires; 5: Río Negro; 6: Chubut y 7: Santa Cruz. Condiciones de análisis K=3, burning= 40000, MarkovChain Monte Carlo simulations= 60000.

#### **IV-4. Resultados de STRs de Cromosoma Y**

De las 145 muestras que constituyen la población en estudio, 44 corresponden a individuos varones. Sobre estas muestras se analizaron 23 STRs de cromosoma Y, cuyos haplotipos son presentados en la Tabla 7. Diez de las muestras fueron analizadas con el kit Y-Filer, y como consecuencia fueron tipificados con 17 Y-STRs.

Si bien los resultados constituyen un análisis preliminar ya que el número de individuos no es suficiente para conformar una base de datos de cromosoma Y (Evet and Gill, 1991; Carracedo et al., 2013), los resultados obtenidos indican la existencia de una gran diversidad, donde todos los haplotipos hallados son únicos. Por otro lado, cada uno de los 44 haplotipos, fueron ingresados en la base de datos mundial de YHRD encontrándose solo dos del total con una frecuencia de 1 en 25499 haplotipos. (YHRD release 49).

**Tabla 7:** Haplotipos de cromosoma Y en la Población de Santa Cruz

	D Y S 5 7 6	D Y S 3 8 9	D Y S 4 4 8	D Y S 3 8 9	D Y S 1 3 9	D Y S 4 8 1	D Y S 5 4 9	D Y S 5 3 3	D Y S 5 4 3	D Y S 5 3 8	D Y S 5 4 7	D Y S 5 3 5	D Y S 5 3 9	D Y S 4 3 9	D Y S 5 6 4	D Y S 3 9 3	D Y S 3 4 5	D Y S 3 8 5	D Y S 3 8 5	D Y S 4 5 6	G A T A H 4	F r e c u e n c i a	H a p l o t i p o		
HY1	19	13	20	30	13	10	22	12	12	10	15	20	22	24	12	11	12	13	15	16	17	17	12	0,02273	E1b1b
HY2	19	12	20	29	14	10	26	12	11	10	16	19	21	22	11	11	12	13	14	13	14	14	11	0,02273	I1
HY3	18	14	20	30	13	10	28	12	11	10	14	22	22	24	10	11	12	13	17	13	14	15	12	0,02273	E1b1b
HY4	17	13	18	30	13	10	22	12	13	11	15	17	23	25	11	13	9	13	17	11	11	16	11	0,02273	R1b
HY5	19	13	18	29	15	11	22	12	13	12	15	17	23	23	12	13	10	13	18	11	14	17	12	0,02273	R1b
HY6	19	13	20	31	13	10	26	12	10	10	14	19	22	24	12	11	12	12	18	17	18	17	12	0,02273	E1b1b
HY7	19	13	20	29	15	9	28	11	11	10	14	22	21	24	10	11	13	13	18	13	14	17	12	0,02273	E1b1b
HY8	16	13	19	29	14	10	23	12	12	12	16	17	23	24	12	13	11	13	17	12	14	17	12	0,02273	R1b
HY9	16	13	20	30	13	10	22	13	12	10	14	20	21	24	12	11	12	13	16	16	17	17	12	0,02273	E1b1b
HY10	18	14	20	32	15	10	27	14	12	10	14	19	23	23	11	12	12	16	15	16	16	16	11	0,02273	I2b
HY11	16	12	20	27	14	10	25	12	11	10	16	20	25	22	11	11	12	13	14	13	14	15	11	0,02273	I1
HY12	18	14	18	31	15	11	22	12	12	12	14	17	23	24	12	13	10	13	17	11	14	15	11	0,02273	R1b
HY13	18	13	19	29	14	11	23	13	13	13	15	17	24	24	12	13	10	13	18	11	14	16	12	0,02273	R1b
HY14	18	13	20	30	15	11	20	14	11	9	15	18	23	23	11	11	10	12	17	13	17	15	12	0,02273	J2a
HY15	19	14	20	31	13	10	25	12	13	11	14	17	22	23	13	16	10	14	16	15	19	15	12	0,02273	Q1a3a
HY16	20	13	20	30	14	10	24	13	11	10	14	18	21	23	12	11	9	12	17.2	13	15	15	11	0,02273	J1
HY17	19	13	20	30	13	10	22	12	12	10	14	19	21	24	13	11	11	13	15	16	19	17	11	0,02273	E1b1b
HY18	17	14	18	30	14	11	22	13	12	9	14	18	24	24	12	14	10	13	17	11	14	16	11	0,02273	R1b
HY19	18	13	19	29	15	10	22	12	11	12	15	17	23	24	11	14	11	13	19	11	15	15	12	0,02273	R1b
HY20	18	13	19	29	14	11	19	13	12	12	15	17	23	24	11	13	10	13	19	12	14	15	12	0,02273	R1b
HY21	19	14	21	31	13	10	24	13	12	11	14	17	22	25	13	14	10	13	20	14	15	15	13	0,02273	Q1a3a
HY22	15	14	21	31	15	11	27	12	11	11	14	19	21	21	11	11	13	13	16	16	17	15	12	0,02273	E1b1a
HY23	20	13	20	29	13	10	25	13	12	11	14	17	22	24	12	14	10	13	18	13	14	15	13	0,02273	Q1a3a
HY24	15	13	20	30	15	11	22	13	12	12	14	17	23	24	12	13	11	13	16	12	15	16	12	0,02273	R1b
HY25	19	13	20	30	13	9	27	13	11	13	14	16	23	24	13	14	10	13	18	14	16	15	12	0,02273	Q1a3a
HY26	18	13	19	28	14	11	21	11	13	12	15	18	23	24	13	13	9	13	17	11	15	15	12	0,02273	R1b
HY27	17	13	18	30	13	10	22	12	13	11	15	18	23	25	11	13	9	13	18	11	11	17	11	0,02273	R1b
HY28	19	12	20	29	14	11	22	12	12	12	14	17	24	23.3	12	13	9	13	16	11	14	15	12	0,02273	R1b
HY29	18	13	21	29	14	10	25	13	11	10	14	18	21	23	12	11	9	12	17.2	13	16	15	11	0,02273	J1
HY30	18	14	19	31	13	10	24	13	13	11	14	19	22	23	10	13	10	13	18	14	18	16	12	0,02273	Q1a3a
HY31	17	12	18	28	14	10	22	13	12	12	14	17	23	24	11	13	10	13	18	11	14	15	12	0,02273	R1b
HY32	16	12	20	28	13	10	23	12	12	11	14	18	22	21	12	14	11	14	19	14	19	16	11	0,02273	Q1a3a
HY33	17	14	20	31	13	10	25	12	12	11	14	17	22	23	13	16	11	14	16	15	19	15	12	0,02273	Q1a3a
HY34	16	13	20	30	14	10	22	13	11	9	16	16	23	23	11	11	10	12	17	13	18	15	11	0,02273	J2a
HY35	ND	12	19	28	14	11	ND	ND	ND	13	15	ND	24	23	12	13	ND	13	17	11	14	16	12	0,02273	R1b
HY36	ND	14	19	30	13	9	ND	ND	ND	10	14	ND	21	23	10	11	ND	13	17	13	14	16	13	0,02273	E1b1b
HY37	ND	14	18	30	14	10	ND	ND	ND	12	14	ND	23	23	13	13	ND	13	17	11	14	17	11	0,02273	R1b
HY38	ND	14	19	30	14	10	ND	ND	ND	9	16	ND	20	24	11	11	ND	13	16	13	16	16	11	0,02273	E1b1b
HY39	ND	14	20	30	13	9	ND	ND	ND	10	14	ND	22	24	10	11	ND	13	17	14	14	16	12	0,02273	E1b1b
HY40	ND	14	19	30	14	11	ND	ND	ND	12	15	ND	23	24	11	13	ND	13	17	12	15	15	12	0,02273	R1b
HY41	ND	13	19	30	14	11	ND	ND	ND	12	15	ND	23	24	13	13	ND	13	16	13	14	17	12	0,02273	R1b
HY42	ND	14	19	29	14	10	ND	ND	ND	12	15	ND	24	24	13	13	ND	13	20	11	13	15	12	0,02273	R1b
HY43	ND	13	20	29	14	11	ND	ND	ND	13	15	ND	23	24	12	13	ND	13	18	11	12	15	12	0,02273	R1b
HY44	ND	13	19	30	14	10	ND	ND	ND	12	15	ND	25	23	12	13	ND	13	16	11	14	16	12	0,02273	R1b

ND: no determinado

La asignación a haplogrupos fue realizada mediante el software Haplotype Predictor ([www.hprg.com](http://www.hprg.com)) y confirmada mediante la amplificación de SNPs característicos de haplogrupos (ver mat y metodos) (Figura 31).

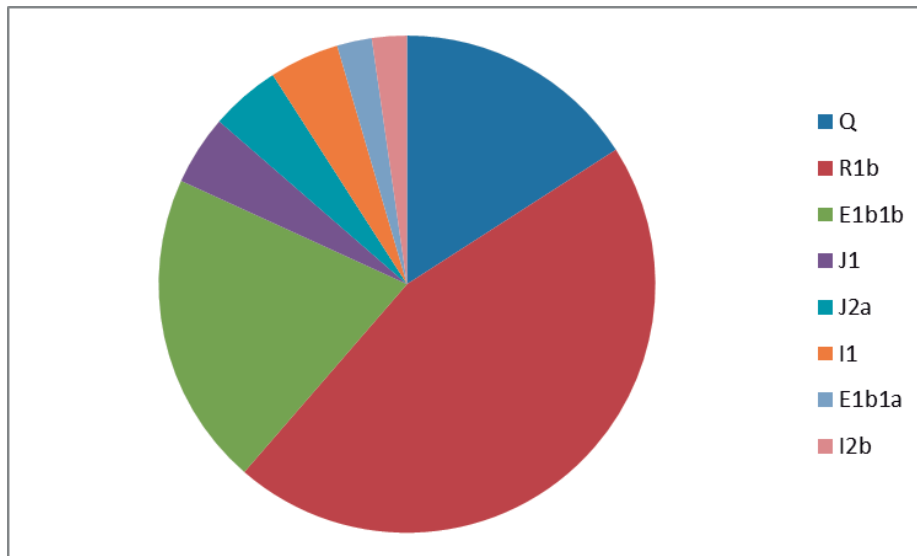


Figura 31 Frecuencia de Haplogrupos en población estudiada

Siete haplotipos (16%) corresponden al haplogrupo Nativo Americano Q1a3a, y los restantes son asignables a haplogrupos Europeos: R1b (45,45%), J1 (4,54%), J2a (4,54%), I1 (4,54 %) e I2b (2,27%); al haplogrupo E1b1b (20,45%)(asignable a poblaciones Europeas y Africanas –ver discusión-) y al haplogrupo africano E1b1a (2,27%)

La figura 32 muestra una representación de las distancias genéticas observadas entre las poblaciones que fueron comparadas. Las distancias genéticas de la población de Santa Cruz, para Y-STRs, resultaron significativas en relación a los grupos de comparación (BA, CHU y RN)( $p < 0.05$ ).

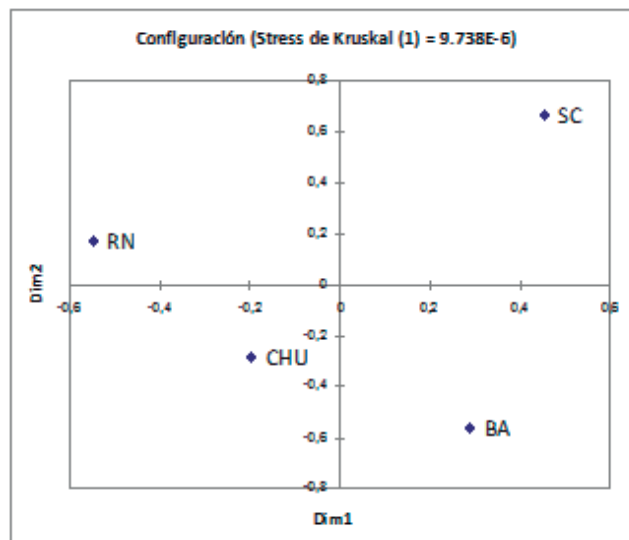
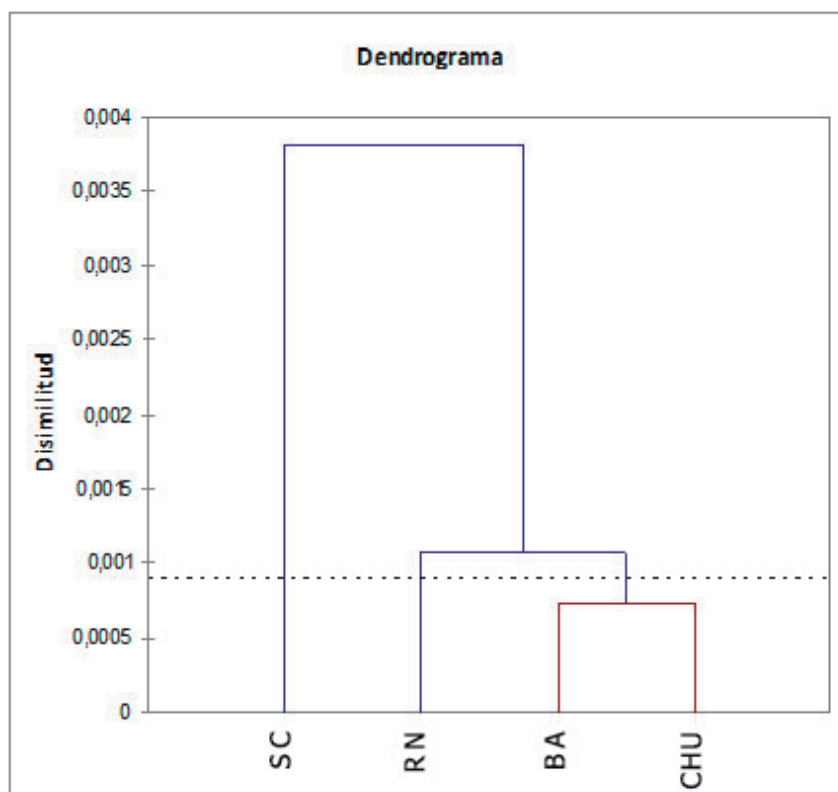


Figura 32 Representación MDS (multi dimensional scaling) de las distancias genéticas ( $F_{st}$ , Slatkin, número de permutaciones 1000). Los sistemas considerados fueron los correspondientes al haplotipo mínimo: DYS19, DYS389 I y II, DYS390, DYS391, DYS392, DYS393 y DYS385Aa/b.

El dendograma permite observar la posición relativa de la población en estudio respecto de los otros tres grupos (Figura 33).



**Figura 33** Dendrograma de representación de las distancias genéticas entre las poblaciones de Buenos Aires (BA); Chubut (CHU); Río Negro (RN) y Santa Cruz (SC), haplotipo mínimo de cromosoma Y.

# **V - DISCUSIÓN**



## V-DISCUSIÓN

El presente trabajo involucró el análisis de 145 muestras de la provincia de Santa Cruz a través de la tipificación de 15 STRs autosómicos. Además, las muestras de individuos varones fueron tipificadas con 23 Y-STRs y un conjunto de 5 SNPs. Si bien inicialmente los individuos muestreados fueron divididos en tres grupos de acuerdo a la procedencia geográfica y a las circunscripciones judiciales, un análisis preliminar permitió demostrar que las diferencias entre estos tres grupos fueron no significativas, lo que posibilitó trabajar posteriormente con una única base de datos unificada denominada Santa Cruz (SC).

Teniendo en cuenta la hipótesis planteada “si el uso de los 15 microsatélites autosómicos (STRs) propuestos para la investigación de la estructura genética de la población actual de la provincia, permiten correlacionar la información genética con la información histórica –demográfica disponible” se puede inferir que:

- Los sistemas analizados cumplen con las expectativas de Equilibrio de Hardy-Weinberg, si bien fue detectado un sistema (D16S539) que estaba fuera del equilibrio. No obstante, al aplicar la corrección de Bonferroni, este sistema también cumple con el EHW. Demostrar que la población en estudio se encuentra en equilibrio permite inferir que es una población cuyos individuos tienen un apareamiento aleatorio y que los eventos migratorios, que han marcado una dinámica poblacional en Santa Cruz, han tenido un efecto menor en la variación de las proporciones genotípicas de la misma, seguramente porque el movimiento demográfico se ha dado en ambos sentidos, tanto para inmigrar como para emigrar.
- La distribución de frecuencias alélicas exhibe una amplia variabilidad genética, con frecuencias relativamente bajas en todos los marcadores, excepto en el marcador TPOX, el que, al igual que en otros estudios poblacionales exhibe una variante alélica ampliamente representada en la población (alelo 8). En su conjunto, el set de marcadores analizados permite arribar a altos niveles de certeza tanto en estudios de identificación como en establecimiento de vínculos de parentesco. Este hecho se ve reflejado en los índices de Poder Discriminativo y de Exclusión y la probabilidad de coincidencia observados, los cuales muestran valores que están dentro del promedio de los estimados en otras poblaciones (Marino, et al, 2006 a, b, c).

Los marcadores STRs que más aportan para esta resolución, por su contenido de información polimórfica y sus valores de poder de discriminación y exclusión son el FGA, D18S51 y el D2S1338, datos coincidentes con la publicación de Borosky et al., 2014.

- Han sido detectadas numerosas microvariantes, pocos frecuentes en otras bases de datos o incluso no reportadas en STRbase (<http://www.cstl.nist.gov/strbase>) (microva-

riantes alélicas 7.2, 12.2, 16.2 y 17.2 en el marcador D19S433 y microvariante 12.1 en D7S820). El hallazgo de una microvariante en estudios de paternidad e identificación forense permite elevar los índices de verosimilitud, aumentando la certeza del estudio.

- Las distancias genéticas de la población de SC respecto de las poblaciones de Chubut, Río Negro y Buenos Aires, tomando como base los 13 STRs comprendidos en el CODIS, han sido significativas ( $p < 0.05$ ), colocando a la población de Santa Cruz en un cluster independiente de las restantes provincias patagónicas incluidas en el análisis.
- Mediante el uso de marcadores de herencia biparental, se puede evidenciar que la población de Santa Cruz, desde el punto de vista étnico está compuesta por un componente ancestral Europeo predominante (50,7%), luego le sigue el componente Nativo-americano (36,3%) y en menor medida el componente Africano (13%).

En ese sentido, la información histórica nos permite plantear una explicación respecto de la importante contribución genética europea y es que una de las corrientes migratorias más importante fue la que ingreso a través de Punta Arenas (Chile), lugar por donde ingresaron la mayor cantidad de contingentes inmigratorios europeos por contar con un importante centro portuario. Respecto a la menor proporción de Nativo-americano se puede correlacionar con la gran matanza de aborígenes durante las campañas militares llevadas a cabo durante la segunda mitad del siglo XIX (“La Conquista del Desierto”, 1878-1885). El porcentaje de componente Nativo-americano es similar a otras regiones estudiadas del país.

Respecto al menor aporte del componente africano, puede deberse a la abolición del tráfico de esclavos, o a la alta mortandad masculina en las guerras de 1810-1870. Este porcentaje es ligeramente mayor en comparación con las provincias de Chubut, Río Negro y Buenos Aires (Fejerman et al, 2005; Sala y Corach, 2014).

- Teniendo en cuenta la herencia uniparental (cromosoma Y) y a pesar de tratarse de un pequeño tamaño muestral ( $N=44$ ), se observa una importante diversidad de haplotipos, perteneciendo a los haplogrupos R1b (45%), linaje de mayor prevalencia en Europa, principalmente en el mundo occidental y región Mediterránea; J1 (5%), J2a (5%), I1(2%) e I2b (2%) linajes Europeos; haplogrupo E1b1b (20%), predominante en el noroeste, sur y este de África y en Oriente Medio, aunque también está presente en Europa en los Balcanes, Italia y España ; haplogrupo Nativo-americano Q1a3a (16%) exclusivo del continente americano; finalmente y en menor proporción se identificó el haplogrupo E1b1a (2%) que es el más frecuente en África-subsahariana.

Los registros demográficos indican que las inmigraciones europeas estuvieron compuestas por una proporción mayor de hombres respecto de mujeres, tanto en la época colonial como después de la independencia nacional en 1816 (Assadourian et al. 1986). De acuerdo a los censos nacionales de los años 1895 y 1914, la proporción de extranjeros era de 172 hombres por cada 100 mujeres. Este alto índice de mas-

culinidad de inmigrantes europeos se mantuvo constante hasta principios del siglo XX (Avena et al. 2006). Estos hechos podrían explicar la alta prevalencia de linajes de ascendencia europea en la región, que se correlaciona además con datos previamente publicados para el resto del país (Corach et al 2010).

- Las distancias genéticas entre Santa Cruz y los grupos de comparación (BA, CH y RN) resultaron significativas, al igual que las determinadas a partir de los marcadores autosómicos.
- El análisis de estructura basado en STRs autosómicos reveló que la población de Santa Cruz tiene como componentes principales el Europeo y el Nativo-americano. Sin embargo, se observó que tiene un componente africano ligeramente superior que los restantes grupos incluidos en el análisis (BA, RN y CHU). A este respecto, el análisis de haplogrupos de cromosoma Y reveló un 2% de Hg E1b1a. Este haplogrupo es el más frecuente en África-subsahariana, alcanzando hasta un 80% en el oeste africano (regiones geográficas de donde se traficaron esclavos hacia América durante los siglos XVI a XVIII). Por otro lado, si bien el haplogrupo E1b1b (20% de los haplotipos de SC) tiene una alta prevalencia en África, también se encuentra en algunas regiones de Europa mediterránea, por lo que no puede descartarse un origen europeo asociado a las migraciones provenientes del viejo continente.

Luego de este análisis, los 44 haplotipos encontrados, correspondería a un 82% de haplogrupos europeos, 16 % Nativo-Americano y 2% de Africano.

# **VI - CONCLUSIONES**

## VI-CONCLUSIONES

Desde el punto de vista de identificación individual, los STRs autosómicos son la herramienta básica en los laboratorios forenses por el elevado poder discriminativo, cualidad que ha quedado demostrado en un gran número de estudios poblacionales, incluido el presente trabajo.

### **Las conclusiones finales son:**

- El presente trabajo ha permitido generar una base de datos de frecuencias alélicas de la población general de Santa Cruz.
- El análisis realizado ha permitido demostrar que la población se encuentra en equilibrio de H-W para los marcadores investigados, cuya implicancia directa es la constancia en las frecuencias alélicas y genotípicas, permitiendo estimar frecuencias genotípicas esperadas.
- Los resultados han permitido establecer una correlación entre la información genética y la demográfica, y han permitido además inferir la composición étnica de la población en estudio.
- Si bien el análisis de haplotipos de cromosoma Y constituye un trabajo preliminar, ha permitido extraer conclusiones así como determinar a priori las prevalencia de haplogrupos característicos. Los linajes paternos europeos fueron los predominantes, seguidos por los nativo-americanos y en baja frecuencia, los africanos.
- Si bien existe información poblacional previa que ha sido publicada sobre esta provincia (Toscanini et al, 2007; Boroski et al, 2014), se destaca del presente trabajo que: el muestreo ha involucrado los distintos distritos poblacionales de la provincia y que ha permitido realizar una inferencia étnico-demográfica.
- La evaluación de los resultados obtenidos ha permitido demostrar que la presente base de datos puede ser incorporada al trabajo de rutina del laboratorio del LRIF a los efectos de estimar los Índices de verosimilitud correspondientes en estudios de filiación e identificación forense.

Perspectivas futuras:

### **A nivel de análisis poblacional:**

- Incorporación a la base de datos de cromosoma Y de nuevas muestras de individuos varones no relacionados con el objetivo de generar una base de datos de frecuencias ha-

plotípicas local para su empleo en genética forense y estudios genético-poblacionales.

- Se plantea además, como proyecto de continuación del presente trabajo, el análisis de linajes maternos mediante la secuenciación del D-Loop mitocondrial, con el objetivo de completar la caracterización molecular de marcadores de herencia uniparental.
- El análisis de la estructura genética de la población de Santa Cruz podrá ser ampliado mediante la caracterización de Marcadores Informativos de Ancestro (Indels y SNPs).
- Publicación de los resultados obtenidos en revistas con referato nacional e internacional reconocidas.

**Desde el punto de vista profesional y académico:**

El cumplimiento de estas perspectivas a futuro podría permitir defender un proyecto de investigación con el objeto de alcanzar un título de posgrado en el área.

## **Difusión de resultados:**

Parte de los resultados aquí presentados han sido enviados como resumen a la reunión de la ISFG (*26th Congress of the International Society for Forensic Genetics, 31 de agosto al 5 de septiembre, 2015, Cracovia, Polonia*)

### **ANALYSIS OF THE GENETIC STRUCTURE OF SANTA CRUZ PROVINCE AND IT COMPARISON WITH THE OTHER SOUTHERN PATAGONIAN PROVINCES OF ARGENTINA**

H. Cano<sup>1</sup>, S. Ginart<sup>2</sup>; M. Caputo<sup>2,3</sup>, D. Corach<sup>2,3</sup>, and A. Sala<sup>2,3</sup>

<sup>1</sup>Laboratorio Regional de Investigación Forense, Poder Judicial de Santa Cruz, Argentina.

<sup>2</sup>Servicio de Huellas Digitales Genéticas, School of Pharmacy and Biochemistry Universidad de Buenos Aires, Argentina

<sup>3</sup>Consejo Nacional de Investigaciones Científicas y Tecnológicas-CONICET, Argentina

Santa Cruz is the southern most continental Patagonian province (latitude ranging 46°00' S - 52°23' S and longitude 65°43' W - 73° 35' W). Its demographic history is complex and changing over time. The oldest settlers were the Aónikenk, arriving around 12000 years before present. The Mapuche penetration started from western Andes during the seventeenth and eighteenth centuries. Europeans arrived during nineteenth and early twentieth centuries. Since then, the region experienced a constant demographic fluctuation due to Argentinean and international migrations. Aiming to investigate the overall genetic features of forensic interest in the Province of Santa Cruz a total number of 645 individuals were selected. The sample set included: Santa Cruz (N=145), Chubut (N=102), Río Negro (N=99), Buenos Aires (N=93) and potentially parental population samples from CEPH (N=206). Analysis included: 15 autosomal STRs, 23 Y-STRs and a set of 6 Y-SNPs. Genetic distance was significant between Santa Cruz either to Chubut, Río Negro or Buenos Aires populations. Native American Y-specific haplogroup Q1a3a was present in around 15% of the Santa Cruz males and the haplogroup R1b was the most frequent European hg (45%). Structure analysis demonstrated that Santa Cruz population is composed by three ancestral contributors: European (50%), Native American (36%) and African (13%). Genetic characterization of this population allows increasing genetic knowledge of the country, in which previous studies have identified regional differences associated with the complex demographic process that took place in South America.

Estos resultados fueron presentados en forma de Poster en el 26th Congress of the International Society for Forensic Genetics, 31 de agosto al 5 de septiembre 2015 en Cracovia, Polonia.



Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: [www.elsevier.com/locate/FSIGSS](http://www.elsevier.com/locate/FSIGSS)



## Analysis of the genetic structure of Santa Cruz province and its comparison with the other Southern Patagonian provinces of Argentina

H. Cano<sup>a</sup>, S. Ginart<sup>b</sup>, M. Caputo<sup>b,c</sup>, D. Corach<sup>b,c</sup>, A. Sala<sup>b,c,\*</sup>

<sup>a</sup> Laboratorio Regional de Investigación Forense, Poder Judicial de Santa Cruz, Argentina

<sup>b</sup> Servicio de Huellas Digitales Genéticas, School of Pharmacy and Biochemistry, University of Buenos Aires, Argentina

<sup>c</sup> Consejo Nacional de Investigaciones Científicas y Técnicas-CONICET, Argentina

### ARTICLE INFO

#### Article history:

Received 1 September 2015

Accepted 14 September 2015

Available online xxx

#### Keywords:

Argentinean Patagonia

Bi and uniparentally genetic markers

### ABSTRACT

Santa Cruz is the southernmost Patagonian province of Argentina. Its demographic history is complex and changing over time. Aiming to investigate the overall genetic features of this province a set of markers routinely used in forensic casework were analyzed. The study involved 645 from Santa Cruz ( $N=145$ ), Chubut ( $N=102$ ), Rio Negro ( $N=99$ ), Buenos Aires ( $N=93$ ) and potentially parental population samples from CEPH panel ( $N=206$ ). We analyzed 15 autosomal STRs, 23 Y-STRs and 6 Y-SNPs. Autosomal STRs displayed statistically significant genetic distances between Santa Cruz and Chubut, Rio Negro and Buenos Aires populations. Native American Y-specific haplogroup Q1a3a was present in around 15% of the Santa Cruz males and the haplogroup R1b was the most frequent European hg (45%). Structure analysis demonstrated that three ancestral contributors that build to Santa Cruz gene pool were European (50%), Native American (36%) and African (13%). Genetic characterization of this population allowed to further increase genetic knowledge of the country, where previous studies have identified regional differences associated with the complex demographic process that took place in South America.

© 2015 Elsevier Ireland Ltd. All rights reserved.

### 1. Introduction

Santa Cruz province is the southernmost Patagonian province (latitude ranging 46°00'S–52°23'S and longitude 65°43'W–73°35'W). The origin of the current population of Santa Cruz (SC) is based on successive internal and external migration events, constituting an interesting model from molecular demography point of view. Being the second larger province of Argentina, after Buenos Aires, with 243,945 km<sup>2</sup>, is one of the less populated (after Tierra del Fuego, Antártida and South Atlantic Islands) with a population density of 1.1 hab/km<sup>2</sup>. The actual population size is 263,974 inhabitants, mostly concentrated in Rio Gallegos, capital of the province.

The populations of Santa Cruz had experienced fluctuations across a short period. The oldest settlers were the Aónikenk, arriving around 12,000 years before present. The Mapuche penetration started from western Andes during the seventeenth and eighteenth centuries. Europeans arrived during nineteenth and early twentieth centuries. In the recent past, the region

experienced a constant demographic fluctuation due to internal migrations (mostly from Buenos Aires province) and external migrations from neighboring countries (Chile and in a lesser degree Bolivia) accompanying economic fluctuations related to oil exploitation.

Aboriginal communities living in Santa Cruz, as well as in the rest of Argentinean Patagonia, come mainly from two ethnic groups: Mapuche and Tehuelche [1].

The overall objective of this study was to characterize the genetic composition of this population by means of Mendelian and uniparentally transmitted markers. In addition, a comparison with other Argentinean provinces, geographically or demographically related was carried out.

### 2. Materials and methods

#### 2.1. Sample donors

We analyzed 145 samples belonging to the most populated areas of Santa Cruz (SC) province. Population comparisons included Buenos Aires (BA,  $N=93$ ), Chubut (CHU,  $N=102$ ) and Rio Negro (RN,  $N=99$ ) (Fig. 1).

\* Corresponding author. Fax: +54 1149648281/8282.

E-mail address: [asala@ffyba.uba.ar](mailto:asala@ffyba.uba.ar) (A. Sala).





Fig. 1. Map of Argentina shown geographical localization of sampling places.

## 2.2. Genetic polymorphisms analyzed

15 autosomal STRs, included in Identifiler Plus kit (Applied Biosystems), 23 Y-STRs (Power Y-23, Promega) and six Y-SNPs, specific for Q1a3a, R1b, E1b, J, I2 and G2a Y-haplogroups (Hgs).

## 2.3. Statistical analysis

Population genetic parameters were determined by using the Software GenAlix v6.501. Hardy Weinberg Equilibrium and interpopulation analysis were determined with Arlequin v3.5.1.2. Analysis of the population composition was carried out with Structure v.2.3.4; parental populations (European, Native American and African) were selected from HGDP-CEPH panel ( $N = 206$ ). Haplogroup Predictor Software (<http://www.hprg.com>) and YHRD-global tool database ([www.yhrd.org](http://www.yhrd.org), release 49) were used for determine Y Chromosome Hgs, confirmed with specific Y-SNPs typing by Real Time PCR followed by HRM.

## 3. Results and discussion

Allele frequencies were determined by counting method. None except one autosomal STR (D16S539) departed from HWE however, after Bonferroni correction, all system met equilibrium. The combined match probability was  $1.358 \times 10^{-17}$  and the combined power of exclusion was higher than 99,999999%. FGA, D18S51 and D21S11 systems shown the highest PD and PE (ranging from 0.9708 to 0.9545; and 0.7188 to 7031, respectively) and the lowest values were observed in TPOX (PD = 0.8296 and PE = 0.3164), in accordance with other Patagonian provinces [2].

Males were represented in 44/145 samples. All the haplotypes were unique, and only 2 were found in the worldwide Y-haplotype database (YHRD, release 49). The haplogroup distribution was R1b (45%), E1b1b (20%), E1b1a (2%), J1 (5%), J2a (5%), I1 (2%) and I2b (2%). Native American Y specific hg Q1a3a was observed in 15% of the male population, higher than general Argentinean population (5%) [3].

Population comparison between Santa Cruz and Buenos Aires population (selected due its demographic relationship) and two Patagonian provinces, Chubut and Río Negro (geographically

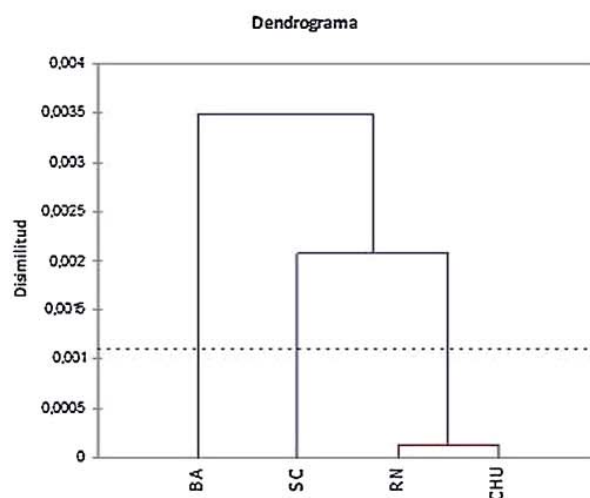


Fig. 2. Dendrogram representing the genetic distances between populations of Buenos Aires (BA); Chubut (CHU); Rio Negro (RN) and Santa Cruz (SC).

related to SC). Dendrogram from  $F_{st}$  genetic distances considering the 13 STRs Codis Core is shown in Fig. 2. All genetic distances, except RN-CHU, were significant ( $p < 0.05$ ). As can be observed, SC is located in between Buenos Aires and the cluster conformed by Chubut and Rio Negro.

Amova analysis considering three groups (BA, CHU+RN, and SC) shown no significant values for *Among groups* and *Among populations within groups* fixation indexes (FCT 0.00264,  $p = 0.176 +/ - 0.0098$ ; FSC 0.00011,  $p = 0.411 +/ - 0.0144$ , respectively).

A structure analysis was performed employing parental population's samples from HGDP-CEPH panel. Using biparentally transmitted markers, Santa Cruz population show a European ancestral component of 50.7%, Native-American component 6.3% and African component 13%, similar to those results obtained for populations included in the comparison (data not shown).

The results obtained could be correlated with the demographic history of Santa Cruz, reflecting in their genetic composition migratory movements of the recent past. This study let us to infer the ethnic composition of the studied population and generate a database for forensic purpose. Further analysis of mtDNA D-Loop sequence and Ancestral Informative Markers will complete the characterization of one of the most austral population of the planet.

## Conflict of interest

The authors report no declarations of interest.

## Acknowledgments

This work was support in part by grants 20020130100783BA-UBACyT to DC. DC, AS and MC are members of CONICET.

## References

- [1] A. Sala, D. Corach, Analysis of admixture and genetic structure of two Native American groups of Southern Argentinean Patagonia, *Mol. Biol. Rep.* 41 (2014) 1533–1543.
- [2] M. Marino, A. Sala, D. Corach, Genetic attributes of 15 autosomal STRs in the population of two patagonian provinces of Argentina, *Forensic Sci. Int.* 160 (2006) 84–88.
- [3] D. Corach, O. Lao, C. Bobillo, K. van Der Gaag, S. Zuniga, M. Vermeulen, K. van Duijn, M. Goedbloed, P.M. Vallone, W. Parson, P. de Knijff, M. Kayser, Inferring continental ancestry of Argentineans from autosomal, Y-chromosomal and mitochondrial DNA, *Ann. Human Genet.* 74 (2010) 65–76.

## LISTA DE ABREVIATURAS Y SIGLAS

**ADN** Ácido desoxirribonucleico

**AMOVA** Análisis de Varianza Molecular

**BA** Buenos Aires

**C1E** Primera Circunscripción Este

**C1O** Primera Circunscripción Oeste

**C2** Segunda Circunscripción

**CEPH** Centro de Estudios de Polimorfismo Humano

**CH** Chubut

**CODIS** Combined DNA Index System

**CORE** Núcleo

**ECPI** Encuesta Complementaria de Pueblos Indígenas

**EHW** Equilibrio de Hardy Weinberg

**EMPOP** Base de Datos de ADN mit

**ENFSI** European Network Forensic Science Institutes

**GHEP-ISFG**, Grupo Español y Portugués de la Sociedad Internacional de Genética Forense

**G-W** Garza-Williamson

**He** Heterocigosidad esperada

**Ho** Heterocigosidad observada

**INDEC** Instituto Nacional de Estadística y Censos

**InDel** Polimorfismo de Inserción-Delección

**IPT** Índice de Paternidad Típico

**Locus** Localización en el genoma

**LRIF** Laboratorio Regional de Investigación Forense

**MLP** Sondas Multi Locus

**mt DNA ADN** mitocondrial

**NCBI** Centro Nacional de Información Biotecnológica

**NO** Noroeste

**NRY** Región no recombinante del Cromosoma Y

**PC** Probabilidad de Coincidencia

**PCR** Reacción en Cadena de la Polimerasa

**PD** Poder de discriminación

**PE** Poder de exclusión

**PIC** Contenido de Información Polimórfica

**RN** Río Negro

**SC** Santa Cruz

**SLP** Sonda de LocusSimple

**SNP** Polimorfismo de Nucleótido Simple

**STRs** Repeticiones cortas en Tandem

**YHRD** Base de Datos de Referencia de cromosoma Y

**YCC** Consorcio del Cromosoma Y

**YCF** Yacimientos Carboníferos Fiscales

**YPF** Yacimientos Petrolíferos Fiscales

# **VII - BIBLIOGRAFIA**

- Acreche, Noemí et al. Diversidad biológica humana en la provincia de Salta. Cuad. Fac. Humanid. Cienc. Soc., Univ. Nac. Jujuy [online]. 2004, n.22 [citado 2015-04-07], pp. 171-194.
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999). "Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA". *Nat Genet* 23 (2): 147. doi:10.1038/13779.
- Avena S., Goycochea A., Rey J., Agosti J., Carnese F.R., Análisis de la Participación del Componente Indígena en una Muestra Hospitalaria de la Ciudad de Buenos Aires, *Revista Argentina de Antropología, Biol. 2* (1):211-226 (1999).
- Avena S., Goycochea A., Dugoujon J., Slepoy M., Slepoy A. y Carnese F.R., Análisis Antropogenético de los Aportes Indígena y Africano en Muestras Hospitalarias de la Ciudad de Buenos Aires, *Revista Argentina de Antropología, Biol. 3* (1):79-99 (2001).
- Avena S, Goycochea A, Rey J., Dugoujon J., Dejean C., Carnese F. Mezcla Genica en una muestra poblacional de la ciudad de Buenos Aires. *MEDICINA (Buenos Aires)* 2006; 66: 113-118.
- Avena SA, et al. FR.2009. Mezcla génica y linajes uniparentales en Comodoro Rivadavia (provincia de Chubut, Argentina). *RevArgAntropBiol* 11:25-41.
- Avena SA et al FR. 2010. Mezcla génica y linajes uniparentales en Esquel (prov. de Chubut). Su comparación con otras muestras poblacionales argentinas. *J Basic ApplGenet* 21:1-14.
- Avena Sergio, Via Mark, Ziv Elad, Pérez-Stable Eliseo J, Gignoux Christopher, Dejean Cristina, Huntsman Scott, Torres-Mejía Gabriela, Dutil Julie, Matta Jaime L, Beckman Kenneth, González Burchard Eesteban, Parolin Maria L, Goicochea Alicia, Acreche Noemi, Boquet Mariel, Rios Part Maria del Carmen, Fernández Vanesa, Rey Jorge, Stern Mariana C, Carnese Francisco R, Fejerman Laura. (2012). Heterogeneity in genetic admixture across different regions of Argentina and implications for medical research. *PLoS ONE* 7(4): e34695.
- Assadourian C. Beato G., Chiaramonte JC (1986) *Argentina: De la Conquista a la Independencia*. Buenos Aires Paidós.
- Bailliet, Graciela et al. Antecedentes y nuevos aportes en el estudio del Cromosoma Y en poblaciones humanas sudamericanas. *BAG, J. basic appl. genet.* [online]. 2011, vol.22, n.1, pp. 0-0. ISSN 1852-6233.
- Ballantyne KN, Keerl V, Wollstein A, Choi Y, Zuniga SB, Ralf A, Vermeulen M, de Knijff P, Kayser M: A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *ForensicSciIntGenet* 2012, 6:208–218.
- Barbería, Elsa. *Los dueños de la tierra en la Patagonia Austral. (1880 - 1920) Río Gallegos*, Ed. de la UFPA. 1995.
- Bonatto SL, Salzano FM (1997a) Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. *Am J Hum Genet* 61:1413– 1423.
- Bonferroni C E *Teoria Statica delle classi e calcolo delle probabilità*. Pubblicazioni del R. Istituto Superiore di Scienze Economiche e Commerciali di Firenze 1936 8:3-62.
- Borosky A, Toscanini U, Gómez A, Parolín M L, Basso N., Vullo C Forensic population data for 20 STR loci in Argentina *FSI: Genetics*, Volume 13, 2014, Pages e20-e21.
- Brinkmann, B. et al. (1996b). Population genetic comparisons among eight populations using allele frequency and sequence data from three microsatellite loci. *European Journal of Human*

Genetics, 4, 175-182.

- Brinkmann, B. et al. (1998). Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *American Journal of Human Genetics*, 62: 1408-1415.
- Budowle B, Moretti TR, Niezgodá SJ, Brown BL: CODIS and PCR-based short tandem repeat loci: law enforcement tools, *Proceedings of the Second European Symposium on Human Identification*. Madison, WI: Promega Corporation; 1998:73-88.
- Budowle, B et al (1999). Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahiamians, Jamaicans and Trinidadians. *Journal of forensic Sciences*, 44 (6): 1227-86.
- Budowle, B. (2004). SNP typing strategies. *Forensic Sciences International*, 146S: S139-S142.
- Butler, J. M. *Forensic DNA Typing: Biology and Technology Behind STR Markers* (Academic Press, 2001).
- Butler J. M. (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. *Journal of Forensic Science* 51, 253-265.
- Cabrero J. y Juan Pedro M. Camacho *Fundamentos de Genética de Poblaciones-cap. 6 2002 Departamento de Genética, Facultad de Ciencias. Universidad de Granada.*
- Carnese F.R., Caratini A. and Goycochea A., *Demography and Blood Genetics of Argentinean mapuche Indians*, *Int. Jour. Of Anthropol.* Vol11(2-4):33-42 (1996).
- Carracedo A, Butler J.M, Gusmao L., Linacre A., Parson W. (2013) *New guidelines for the publication of genetic population data. Forensic Science International: Genetics*, Volume 7, 217-220.
- Casamiquela R: *Rectificaciones y ratificaciones hacia una interpretación definitiva del panorama etnológico de la Patagonia y área septentrional adyacente*. Bahía Blanca: Instituto de Humanidades: Universidad Nacional del Sur Press; 1965.
- Cavalli-Sforza y Bodmer, 1981 *Genetica de las Poblaciones Humanas*.
- Chakraborty, J.T. et al. (1994). Variation in Short Tandem Repeat sequences-a survey of twelve microsatellite loci for use as forensic identification markers. *International Journal of Legal Medicine*, 107: 13-20.
- Coble MD, Loreille OM, Wadhams MJ, Edson SM, Maynard K, Meyer CE, Niederstätter H, Berger C, Berger B, Falsetti AB, Gill P, Parson W, Finelli LN: *Mystery solved: the identification of the two missing Romanov children using DNA analysis. PLoS One* 2009, 4:e4838.
- Corach, D. et al. (2001). Routine Y-STR typing in forensic casework. *Forensic Science International*, 118(2): 131-135.
- Corach, D.; Marino, M. and Sala, A (2006). Relevant genetic contribution of amerindian to the extant population of Argentina. *International Congress Series*, 1288, 397-399.
- Corach, D. et al. (2010). Inferring Continental Ancestry of Argentineans from Autosomal, Y-Chromosomal and Mitochondrial DNA. *Annals of Human Genetics*, 74, 65-76.
- Da Costa Francez PA, Rodrigues EM, de Velasco AM, dos Santos SE. *Insertion-deletion polymorphisms: utilization on forensic analysis. Int J Legal Med.* 2012, 126(4):491-496.
- Edwards A, Civitello A, Hammond HA, Caskey CTDNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet.*1991 , 49:746-756.

- Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*. 1992, 12,:241-253.
- ENFSI survey on DNA Databases in Europe. December 2011, [www.enfsi.eu].
- Escalada F: El complejo tehuelche. Estudios de etnografía patagónica. Buenos Aires: Instituto Superior de Estudios Patagónicos Press; 1949.
- Evet, I.W., Gill P. (1991). A discusión of the robustness of methods for assessing the evidential value of DNA single locus profiles in crime investigations. *Electrophoresis* 12, 226-230.
- Excoffier, L. and H.E. L. Lischer (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564-567.
- Excoffier, L; Smouse, Pe; Quattro, Jm (Jun 1992). "Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data". *Genetics* 131 (2): 479-91.
- Fejerman Laura, Francisco R. Carnese, Alicia S. Goicoechea, Sergio A. Avena, Cristina B. Dejean, and Ryk H. Ward. (2005) African Ancestry of the Population of Buenos Aires. *AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY* 128:164-170 (2005).
- Gill, P., Jeffreys, A. J. & Werrett, D. J. Forensic application of DNA 'fingerprints'. *Nature* 318, 577-579 (1985).
- Gill, P. et al. Identification of the remains of the Romanov family by DNA analysis. *Nature Genet.* 6, 130-135 (1994).
- Gill, P. et al. (1995). Automated short tandem repeat (STR) analysis in forensic casework - a strategy for the future. *Electrophoresis*, 16:1543-1552.
- Gill P, Fereday, N Morling, PM Schneider (2006) The evolution of DNA databases recommendations for new European STR loci. *For SciInt*, 242-244.
- Goycoechea A., Soria M., Haedo A., Crognioier E. y Carnese F., Distancias Genéticas en Poblaciones Aborígenes de la Argentina, *Revista Argentina de Antropología, Biol.* 1(1):153-166 (1996).
- Hammer, M. F. et al. (2005). Population structure of Y chromosome SNP haplogroups in the United States and forensic implications for constructing Y chromosome STR databases. *Forensic Sciences International*, 164, 45-55.
- INDEC : <http://www.indec.mecon.ar>
- International Society of Genetic Genealogy (2014). Y-DNA Haplogroup Tree 2015 version 10.20, Fecha: [ : 28 February 2015 ], <http://www.isogg.org/tree/>. Date of access 02/02/2015.
- Jeffreys, A. J., Wilson, V. & Thein, S. L. Individual-specific 'fingerprints' of human DNA. *Nature* 316, 76-79 (1985).
- Jobling, M. A. et al. (2001). Y-chromosomal SNP haplotype diversity in forensic analysis. *Forensic Science International*, 118(2): 158-162.
- Jobling MA, Hurles ME, Tyler-Smith C: Chapter 15: Identity and identification. In *Human Evolutionary Genetics*. Abingdon: Garland Science; 2003:474-497.
- Karafet, T. M. et al. (1999). Ancestral asian source(s) of new world Y-chromosome founder haplotypes. *American Journal of Human Genetics*, 64: 817-831.
- Kayser, M. et al. (1997). Evaluation of Y-chromosomal STRs: a multicenter study. *International Journal of Legal Medicine*, 110(3), 125-133.

- Kimpton, C. P. et al. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods Appl.* 3, 13–22 (1993).
- Landsteiner, K. Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe. *Zentralbl. Bakteriologie*. 27, 357–362 (1900).
- Lao O, van Duijn K, Kersbergen P, de Knijff P, Kayser M. Proportioning whole-genome single-nucleotide-polymorphism diversity for the identification of geographic population structure and genetic ancestry. *Am J Hum Genet.* 2006 (4):680-690.
- Lao O, Vallone PM, Coble MD, Diegoli TM, van Oven M, van der Gaag KJ, Pijpe J, de Knijff P, Kayser M. Evaluating self-declared ancestry of U.S. Americans with autosomal, Y-chromosomal and mitochondrial DNA. *Hum Mutat.* 2010 Dec;31(12):E1875-1893.
- Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, Myers RM. Worldwide human relationships inferred from genome-wide patterns of variation. *Science.* 2008 Feb 22;319(5866):1100-4. doi: 10.1126/science.1153717.
- Martin PD, Schmitter H, Schneider PM: A brief history of the formation of DNA databases in forensic science within Europe. *Forensic Sci Int* 2001, 119:225–231.
- Manfred Kayser, Michael Krawczak, Laurent Excoffier, Patrick Dieltjes, Daniel Corach, Vincente Pascali, Christian Gehrigh, Luigi F. Bernini, Jørgen Jespersen, Egbert Bakker, Lutz Roewer, Peter de Knijff. An Extensive Analysis of Y-Chromosomal Microsatellite Haplotypes in Globally Dispersed Human Populations *Am J Hum Genet.* 2001; 68(4): 990–1018.
- Marino, M., Sala, A. y Corach, D. (2006a). Genetic analysis of the populations from Northern and Mesopotamian provinces of Argentina by means of 15 autosomal STRs. *Forensic Science International*, 160(2-3), 224-230.
- Marino, M., Sala, A. y Corach, D. (2006b). Genetic attributes of 15 autosomal STRs in the population of two patagonian provinces of Argentina. *Forensic Science International*, . 2006 Jun 27;160(1):84-8. PMID: 16024203 [PubMed - indexed for MEDLINE]. ISSN:0379-0738.
- Marino, M., Sala, A. y Corach, D. (2006c). Population genetic analysis of 15 autosomal STRs loci in the central region of Argentina. *Forensic Science International*, 10, 161(1), 72-77.
- Marino, M. et al (2007). Genetic attributes of the YHRD minimal haplotype in 10 provinces of Argentina. *Forensic Science International: Genetics* 1, 129-133.
- Marino, M. et al (2008). Inferring genetic sub-structure in the population of Argentina using fifteen microsatellite loci. *Forensic Science International: Genetics Supplement Series* 1, 350-352.
- Marino M, Sala A, Corach D. Genetic attributes of 15 autosomal STRs in the population of two patagonian provinces of Argentina *Forensic Sci Int.*
- Martinez Sarazola, Nuestros paisanos los indios, Emecé, 2005.
- Mills, R. E., Luttig, C. T., Larkins, C. E., Beauchamp, A., Tsui, C., Pittard, W. S., Devine, S. E., An initial map of insertion and deletion (INDEL) variation in the human genome. *Genome Res.* 2006, 16, 1182–1190.
- Mullis, K. and Faloona, F. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* 155: 335-350.
- Nagy M, Otremba P, Krüger C, Bergner-Greiner S, Anders P, Henske B, Prinz M, Roewer L. Optimi-



- zation and validation of a fully automated silica-coated magnetic beads purification technology in forensics. *Forensic Sci Int.* 2005,152(1):13-22.
- Nakamura Y, Leppert M, Fujimoto E, White R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* 235, 1616-1622.
- Nassir R, Kosoy R, Tian C, White PA, Butler LM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, Seldin MF. An ancestry informative marker set for determining continental origin: validation and extension using human genome diversity panels. *BMC Genet.* 2009 Jul 24;10:39. doi: 10.1186/1471-2156-10-39.
- National Research Council Report: DNA Technology in Forensic Science (NRC I), 1992.
- Palatnik Marcos (1968). Grupos sanguíneos en Ranqueles de Argentina. *Sangre* 13: 31-60. Barcelona
- Parson W, Arne Dür. EMPOP: A forensic mtDNA database. *Forensic Science International: Genetics* Volume 1, Issue 2, 2007, 88–92.
- Peakall R, Smouse PE. GenAEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes.* 2006;6:288–295.
- Peakall R1, Smouse PE. GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. . *Bioinformatics.* 2012 Oct 1;28(19):2537-9.
- Pierini M. Milagros - Título de la ponencia: “Las Migraciones en Santa Cruz: su análisis desde la perspectiva de los Derechos Humanos. Marzo 2010.
- Pigafetta, Antonio. Primer viaje alrededor del mundo relato escrito por el caballero Antonio Pigafetta, editor Manuel Walls y Merino, Imprenta de Fortanet, Madrid, 1899.
- Pizzolitto G. Distribución de la población y migraciones internas en Argentina: sus determinantes individuales y regionales. UNLP 2006.
- Polymeropoulos, M. H. et al. (1992). Tetranucleotide repeat polymorphism at the human myelin basic protein gene (MBP). *Human Molecular Genetics*, 1: 658.
- Pritchard Jonathan K., Matthew Stephens and Peter Donnelly (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155: 945-959.
- Quinque D, Kittler R, Kayser M, Stoneking M, Nasidze I: Evaluation of saliva as a source of human DNA for population and association studies. *AnalBiochem* 2006, 15:272-277.
- Quiroga Micheo E, Vilaseca Alicia, Bonder M, Quiroga Vergara E (1988). Frecuencia de los grupos sanguíneos y análisis de la progresiva disminución del factor Rh negativo en la población argentina. *Medicina (Buenos Aires)* 48:355-360.
- Roewer, L. et al. (1996). Analysis of molecular variances (AMOVA) of Y-chromosome specific microsatellites in two closely related human populations. *Human Molecular Genetics*, 5 (7), 1029-1033.
- Roewer, L. et al. (2000). A new method for the evaluation of matches in nonrecombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Sciences International*, 114: 31–43.
- Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, de Knijff P, Jobling MA, Tyler-Smith C, Krawczak M: Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Hum Genet* 2005, 116:279–291.

- Sala, A., Penacino, G. y Corach, D. (1998). Comparison of allele frequencies of eight STR loci from Argentinian Amerindian and European populations. *Human Biology*, 70(5), 937-947.
- Sala, A., Penacino, G y Corach, D. (1999). Reference Database of Hypervariable Genetic Markers of Argentina: Applications for Molecular Anthropology and Forensic Casework. *Electrophoresis*, 208, 1733-1739.
- Sala A, C. F. Argüelles, M.E. Marino, C. Bobillo, A. Fenocchio and D. Corach. Genetic analysis of six communities of Mbyá-Guaraní inhabiting North Eastern Argentina by means of nuclear and mitochondrial polymorphic markers. *Human Biology*, 2010, 82: 433-456.
- Sala A and Corach D. Analysis of admixture and genetic structure of two Native American groups of Southern Argentinean Patagonia. *MolBiol Rep* (2014) 41:1533-1543.
- Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, Pereira R, Gusmão L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, Santos SE. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat.* 2010 Feb;31(2):184-90. doi: 10.1002/humu.21159.
- Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: A software for population genetics data analysis, version 2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.
- STRBase: <http://www.cstl.nist.gov/strbase>.
- Structure 2.3.4 software. The basic algorithm was described by Pritchard, Stephens & Donnelly (2000). Extensions to the method were published by Falush, Stephens and Pritchard (2003), and (2007) and Hubisz, Falush, Stephens and Pritchard (2009).
- Tarragó, M.N. Los pueblos originarios y la conquista. Editorial Sudamericana. Buenos Aires. 2000. ISBN 950-07-1792-1.
- Tautz, D. (1993). Notes on the definition and nomenclature of tandemly repeated DNA The Y Chromosome Consortium. A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups *Genome Res.* 2002; 12: 339-348.
- The Evaluation of Forensic DNA Evidence. National Research Council. NATIONAL ACADEMY PRESS Washington, D.C. 1996.
- The Y Chromosome Consortium. A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups *Genome Res.* 2002; 12: 339-348.
- Torróni, A. et al. (1992). Native american mitochondrial DNA analysis indicates that the amerind and the nadene populations were founded by two independent migrations. *Genetics*, 130, 153-162. sequences. In: *DNA fingerprinting: state of the science*. Ed. by SDJ Pena, R.
- Toscanini Ulises, Leonor Gusmao, Gabriela Berardi, Antonio Amorim, Angel Carracedo, Antonio Salas, Eduardo Raimondi Testing for genetic structure in different urban Argentinian populations. *Forensic Science International* 165 (2007) 35-40.
- Toscanini U., et al., Y chromosome microsatellite genetic variation in two Native American populations from Argentina: Population stratification and mutation data, *Forensic Sci. Int. Gene.* (2008), doi:10.1016/j.fsigen.2008.03.001.
- Toscanini U, L. Gusmao, G. Berardi, A. Gomez, R. Pereira, E. Raimondi, Ancestry proportions in urban populations of Argentina, *Forensic Sci. Int. Genet.* 3 (2011) e387-e388.

- Urquhart, A. et al. (1994). Variation in short tandem repeat sequences. *International Journal of Legal Medicine*; 107: 13-20.
- Vali, U., Brandstrom, M., Johansson, M., Ellegren, H., Insertion-deletion polymorphisms (indels) as genetic markers in natural populations *Biomed.Chromatogr.Genet.*2008, 9, 8.
- van Oven M, Kayser M (2009). "Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation". *Human Mutation* 30 (2): E386–394.
- Venter, J. C. et al. (2001). The sequence of the Human Genome. *Science*, 291: 1304-1351.
- Vigilant, L. et al. (1991). African populations and the evolution of human mitochondrial DNA. *Science*, 253 (5027): 1503-1507.
- Weber, J. L. y May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* 44: 388-396.
- Weber, J. L., David, D., Heil, J., Fan, Y., Zhao, C., Marth, G., Human diallelic insertion/deletion polymorphisms. *Am. J. Hum. Genet.* 2002, 71, 854–862.
- Willuweit S., Roewer L. (2015), 'The new Y Chromosome Haplotype Reference Database.', *Forensic SciInt Genet* 15, 43-8.
- Yang, N., Li, H., Criswell, L. A., Gregersen, P. K., Alarcon-Riquelme, M. E., Kittles, R., Shigeta, R. et al., Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Hum. Genet.* 2005, 118, 382–392.
- Zuccarelli G. et al. / *Forensic Science International: Genetics* 5 (2011) 105–108.

## Resumen:

La población actual de la provincia de Santa Cruz se ha conformado sobre la base de sucesivos eventos de migración, tanto interna como externa, constituyendo un modelo de análisis de gran interés desde el punto de vista genético-poblacional.

El objetivo general del presente trabajo fue caracterizar la composición genética de esta población, mediante el análisis de polimorfismos de herencia mendeliana y uniparental (cromosoma Y).

Se analizaron 145 muestras de individuos no relacionados genéticamente, procedentes de tres Circunscripciones Judiciales. El análisis poblacional incluyó 15 STRs autosómicos, 23 Y-STRs y 5 SNPs de cromosoma Y.

Se determinó un conjunto de parámetros genético-poblacionales y forenses (frecuencias alélicas y genotípicas, índices de diversidad, equilibrio de Hardy-Weinberg, Poder Discriminativo y de Exclusión, Índice de Paternidad Típico, Probabilidad de Coincidencia), estableciendo además comparaciones con otras poblaciones (Chubut; Río Negro y Buenos Aires).

Los resultados permitieron determinar la relación existente entre SC y otras poblaciones en cuanto a su estructura genética. Por otro lado pudo inferirse la estructura étnica de Santa Cruz en referencia al componente Europeo, Nativo Americano y Africano.

El análisis de haplogrupos de cromosoma Y permitió determinar un fuerte componente Europeo en los linajes obtenidos.

Estos resultados son consistentes con la historia migratoria en SC y Argentina, en los que se destaca el impacto de la contribución genética masculina europea a la población actual.

Los resultados obtenidos tienen importancia, por un lado, desde el punto de la investigación básica por su aporte a la genética de poblaciones, y por otro, en el área de la Genética Forense, aportando a la caracterización de poblaciones y generación de banco de datos de referencia.

## **Abstract:**

*The current population of the Santa Cruz province (SC) has been formed on the basis of successive migration events, both internal and external, constituting a model of great interest from the population-genetic point of view.*

*The overall objective of this study was to characterize the genetic composition of this population by analyzing polymorphisms with Mendelian and uniparental inheritance (Y chromosome).*

*145 samples from genetically unrelated individuals from three Judicial Districts were analyzed. The population's analysis included 15 autosomal STRs, 23 Y-STRs and five Y-SNPs.*

*A set of population-genetics and forensic parameters (allele and genotype frequencies, diversity indexes, Hardy-Weinberg equilibrium, Discriminative Power and Power of Exclusion, Typical Paternity Index, Matching Probability) as well as comparisons with other populations (Chubut, Rio Negro and Buenos Aires) was determined.*

*The analysis of the genetic structure allowed establishing relationships with other populations. Moreover, the ethnic composition in relation to European, Native-American and African component was inferred.*

*The analysis of Y chromosome haplogroups allowed determining a strong European component.*

*These results are consistent with the migration history in SC and Argentina, where the impact of European male genetic contribution to the current population stands out.*

*The results are important, first, from the point of basic research for his contribution to population genetics, and secondly, in the area of Forensic Genetics, contributing to the characterization of populations and reference database generation.*

