

**ESTIMACION DE LA VARIACION GENETICA DE LA POBLACION
SANTANEREANA MEDIANTE EL ESTUDIO DE MARCADORES
STR's**

SARA MARIA CASTILLO CASTRO

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE BIOLOGIA
BUCARAMANGA**

2004

**ESTIMACION DE LA VARIACION GENETICA DE LA POBLACION
SANTANEREANA MEDIANTE EL ESTUDIO DE MARCADORES
STR's**

SARA MARÍA CASTILLO CASTRO

**Trabajo de grado presentado como requisito para optar al título de
Bióloga**

DIRECTORA

NORMA CECILIA SERRANO DÍAZ

MD., M.Sc.

UNAB

CO-DIRECTORA

MARÍA CAROLINA PAÉZ LEAL

MD.

Departamento de investigaciones

UNAB

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE CIENCIAS

ESCUELA DE BIOLOGIA

BUCARAMANGA

2004

DEDICATORIA

A mis padres,

Henry y María Elena

Por su apoyo incondicional y motivación

A mi hermana,

Catalina

**Por brindarme su cariño y afecto y paciencia
en esos momentos difíciles**

AGRADECIMIENTOS

A la Dr. Norma Cecilia Serrano Díaz, por su apoyo, colaboración y respaldo científico durante el desarrollo de la presente investigación.

A la Dr. María Carolina Páez Leal, por su tiempo y dedicación, los cuales fueron indispensables para la culminación de este proyecto.

Al Dr. Luis Alfonso Díaz, por su colaboración en la parte estadística y por sus valiosos comentarios y correcciones en el escrito final.

Al Laboratorio Higuera Escalante FOS-CAL, Departamento de genética, por los datos suministrados para el desarrollo del presente trabajo y por permitirme trabajar en sus instalaciones durante la fase de almacenamiento de datos y análisis de los mismos.

Al laboratorio Yunis-Turbay, por su colaboración en el procesamiento de las muestras y sus posteriores resultados que fueron la base fundamental para el desarrollo del trabajo.

A Diego, por el apoyo incondicional en los diferentes momentos por los que pase durante la realización del proyecto.

A mis compañeros que siempre estuvieron dispuestos a colaborar y me ayudaron en esos pequeños detalles que hoy hacen parte de este trabajo.

A todas las personas que de una u otra forma me colaboraron durante el desarrollo de la presente investigación.

CONTENIDO

	Pág.
INTRODUCCIÓN	1
1. OBJETIVOS	4
1.1 OBJETIVO GENERAL	4
1.2 OBJETIVOS ESPECÍFICOS	4
2. MARCO TEÓRICO	5
2.1 GENETICA POBLACIONAL	5
2.1.1 Estructura poblacional	6
2.1.2 Estimadores genético-poblacionales	7
2.1.3 Equilibrio de Hardy-Weinberg	7
2.1.4 Factores que alteran las frecuencias en las poblaciones	8
2.1.5 Ligamiento	9
2.1.6 La heterocigocidad en las poblaciones	10
2.2 ASPECTOS BASICOS DE LOS STR's	11
2.2.1 Bases moleculares de los STR	15
2.2.2 STR tetraméricos autosómicos	15
2.2.3 Nomenclatura y clasificación de los alelos STR	16
2.2.4 Sistemas de análisis de los STR	17
2.2.5 Los STR y estudios poblacionales	18
2.3 EL MESTIZAJE	18
2.3.1 El proceso de mestizaje en Santander	19
2.3.2 Flujo Génico y proceso de mestizaje	20

2.4 ANÁLISIS DE AGRUPAMIENTO	21
3. MATERIALES Y MÉTODOS	23
3.1 TIPO DE ESTUDIO	23
3.2 ALMACENAMIENTO DE DATOS	23
3.3 DESCRIPCIÓN DE LA MUESTRA	23
3.4 VARIABLES UTILIZADAS	24
3.5 ASPECTOS ETICOS	25
3.6 MARCADORES	26
3.7 VARIABILIDAD GENETICA	26
3.7.1 Estimación de frecuencias y estimación de heterocigocidad observada y esperada	26
3.7.2 Pruebas para determinar equilibrio de Hardy-Weinbrg	27
3.7.3 Desequilibrio de ligamiento para loci localizados en el mismo cromosoma	28
3.8 COMPARACIÓN ENTRE POBLACIONES	28
3.8.1 Diferenciación entre poblaciones	28
3.8.2 Análisis de agrupamiento	30
4. RESULTADOS	31
4.1 ESTIMACIÓN DE FRECUENCIAS, EQUILIBRO DE HARDY-WEINBERG Y HETEROCIGOCIDAD OBSERVADA Y ESPERADA EN LA POBLACIÓN SANTANDEREANA	31
4.1.1 Locus TPOX	32
4.1.2 Locus vWA	34
4.1.3 Locus D18S51	37
4.1.4 Locus TH01	40
4.1.5 Locus D3S1358	42

4.1.6 Locus Penta D	45
4.1.7 Locus D7S820	49
4.1.8 Locus D13S317	51
4.1.9 Locus D5S818	54
4.1.10 Locus D16S539	57
4.1.11 Locus CSF1PO	60
4.1.12 Locus D21S11	62
4.1.13 Locus FGA	66
4.1.14 Locus D8S1179	69
4.1.15 Locus Penta E	72
4.2 DESEQUILIBRIO DE LIGAMENTO	79
4.3 COMPARACIÓN ENTRE POBLACIONES	79
4.3.1 Locus TPOX	81
4.3.2 Locus vWA	82
4.3.3 Locus D18S51	84
4.3.4 Locus TH01	84
4.3.5 Locus D3S1358	85
4.3.6 Locus Penta D	85
4.3.7 Locus D7S820	86
4.3.8 Locus D13S317	86
4.3.9 Locus D5S818	87
4.3.10 Locus D16S539	87
4.3.11 Locus CSF1PO	88
4.3.12 Locus D21S11	88
4.3.13 Locus FGA	89
4.3.14 Locus D8S1179	90
4.3.15 Locus Penta E	90
4.4 ANÁLISIS DE AGRUPAMIENTO	96
5. DISCUSIÓN DE RESULTADOS	100

5.1 FRECUENCIAS ALÉLICAS, GENOTÍPICAS, EQUILIBRIO DE HARDY- WEINBERG HETEROCIGOCIDAD	100
5.2 DESEQUILIBRIO DE LIGAMENTO	102
5.3 COMPARACIÓN ENTRE POBLACIONES	103
5.4 ANÁLISIS DE AGRUPAMIENTO	105
6. CONCLUSIONES	107
7. RECOMENDACIONES	109
BIBLIOGRAFÍA	111

LISTA DE FIGURAS

	Pág.
Figura 1. Distribución de las frecuencias alélicas en la población Santandereana para el locus TPOX.	32
Figura 2. Distribución de los genotipos observados y esperados en la población Santandereana para el locus TPOX.	34
Figura 3. Distribución de las frecuencias alélicas en la población Santandereana para el locus vWA.	35
Figura 4. Distribución de los genotipos observados y esperados en la población Santandereana para el locus vWA	36
Figura 5. Distribución de las frecuencias alélicas en la población Santandereana para el locus D18S51.	37
Figura 6. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D18S51	38
Figura 7. Distribución de las frecuencias alélicas en la población Santandereana para el locus TH01	41
Figura 8. Distribución de los genotipos observados y esperados en la población Santandereana para el locus TH01	42
Figura 9. Distribución de las frecuencias alélicas en la población Santandereana para el locus D3S1358	43
Figura 10. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D3S1358	45
Figura 11. Distribución de las frecuencias alélicas en la población Santandereana para el locus Penta D	46
Figura 12. Distribución de los genotipos observados y esperados en la población Santandereana para el locus Penta D	48

Figura 13. Distribución de las frecuencias alélicas en la población Santandereana para el locus D7S820	49
Figura 14. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D7S820	51
Figura 15. Distribución de las frecuencias alélicas en la población Santandereana para el locus D3S317	52
Figura 16. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D3S317	54
Figura 17. Distribución de las frecuencias alélicas en la población Santandereana para el locus D5S818	55
Figura 18. Distribución de las frecuencias alélicas en la población Santandereana para el locus D5S818	57
Figura 19. Distribución de las frecuencias alélicas en la población Santandereana para el locus D16S539	58
Figura 20. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D16S539	60
Figura 21. Distribución de las frecuencias alélicas en la población Santandereana para el locus CSF1PO	61
Figura 22. Distribución de los genotipos observados y esperados en la población Santandereana para el locus CSF1PO	62
Figura 23. Distribución de las frecuencias alélicas en la población Santandereana para el locus D21S11	64
Figura 24. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D21S11	66
Figura 25. Distribución de las frecuencias alélicas en la población Santandereana para el locus FGA	67
Figura 26. Distribución de los genotipos observados y esperados en la población Santandereana para el locus FGA	69
Figura 27. Distribución de las frecuencias alélicas en la población Santandereana para el locus DS1179	70

Figura 28. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D8S1179	72
Figura 29. Distribución de las frecuencias alélicas en la población Santandereana para el locus Penta E	73
Figura 30. Distribución de los genotipos observados y esperados en la población Santandereana para el locus Penta E	76
Figura 31. Dendograma basado en el método UPGMA con distancias Euclidianas para la población Santandereana y otras reportadas a nivel mundial	99

LISTA DE TABLAS

	Pág
Tabla 1. Caracterización de los 15 STR's estudiados en la población Santandereana	13
Tabla 2. Prueba para determinar el equilibrio de Hardy-Weinberg y Heterocigocidad observada y esperada en la población Santandereana	31
Tabla 3. Frecuencias alélicas encontradas en la población Santandereana para el locus TPOX	32
Tabla 4. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus TPOX	33
Tabla 5. Frecuencias alélicas encontradas en la población Santandereana para el locus vWA	34
Tabla 6. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus vWA	35
Tabla 7. Frecuencias alélicas encontradas en la población Santandereana para el locus D18S51	37
Tabla 8. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D18S51	39
Tabla 9. Frecuencias alélicas encontradas en la población Santandereana para el locus TH01	40
Tabla 10. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus TH01	31
Tabla 11. Frecuencias alélicas encontradas en la población Santandereana para el locus D3S1358	43
Tabla 12. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D3S1358	44

Tabla 13. Frecuencias alélicas encontradas en la población Santandereana para el locus Penta D	46
Tabla 14. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus Penta D	47
Tabla 15. Frecuencias alélicas encontradas en la población Santandereana para el locus D7S820	49
Tabla 16. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D7S820	50
Tabla 17. Frecuencias alélicas encontradas en la población Santandereana para el locus D13S317	52
Tabla 18. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D13S317	53
Tabla 19. Frecuencias alélicas encontradas en la población Santandereana para el locus D5S818	55
Tabla 20. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D5S818	56
Tabla 21. Frecuencias alélicas encontradas en la población Santandereana para el locus D16S539	58
Tabla 22. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D16S539	59
Tabla 23. Frecuencias alélicas encontradas en la población Santandereana para el locus CSF1PO	60
Tabla 24. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus CSF1PO	61
Tabla 25. Frecuencias alélicas encontradas en la población Santandereana para el locus D21S11	63
Tabla 26. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D21S11	64
Tabla 27. Frecuencias alélicas encontradas en la población Santandereana para el locus FGA	67

Tabla 28. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus FGA	68
Tabla 29. Frecuencias alélicas encontradas en la población Santandereana para el locus D8S1179	70
Tabla 30. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D8S1179	71
Tabla 31. Frecuencias alélicas encontradas en la población Santandereana para el locus Penta E	73
Tabla 32. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus Penta E	74
Tabla 33. Frecuencias alélicas para los 15 STR's estudiados en la población Santandereana	77
Tabla 34. Desequilibrio de ligamiento presente en la población Santandereana	80
Tabla 35. Frecuencias alélicas encontradas en la población Santandereana y en las 21 poblaciones comparadas, previamente reportadas en la literatura	80
Tabla 36. Resultados de la prueba de Chi-cuadrado Global para las 21 poblaciones comparadas teniendo como referencia la población Santandereana	83
Tabla 37. Heterocigocidad observada en las poblaciones comparadas	93
Tabla 38. Comparación de las heterocigocidades para todas las poblaciones basada en una prueba de Chi-cuadrado	95
Tabla 39. Matriz de distancias genéticas para la población Santandereana y otras poblaciones a nivel mundial	97

RESUMEN

1. **TITULO:** ESTIMACION DE LA VARIACION GENETICA DE LA POBLACION SANTANDEREANA MEDIANTE EL ESTUDIO DE MARCADORES STR's *

2. **AUTOR:** CASTILLO CASTRO SARA MARIA **

3. **PALABRAS CLAVES:** Estructura genética, Frecuencias alélicas, Población Santandereana, STR's, Mestizaje.

4. **CONTENIDO:**

La estructura genética de la población santandereana fue analizada usando 15 marcadores moleculares de tipo STR: FGA, TPOX, D8S1179, vWA, Penta E, D18S51, D21S11, THO1, D3S1358, Penta D, CSF1PO, D16S539, D7S820, D13S317, D5S818. Se reportan por primera vez frecuencias alélicas para esta población de los marcadores penta D y penta E. Las frecuencias alélicas fueron calculadas a partir de los genotipos usando el método de conteo directo. Posibles desviaciones del equilibrio de Hardy-Weinberg fueron determinadas por el método exacto; con el mismo programa se determinó el desequilibrio de ligamiento para los loci ubicados en el mismo cromosoma. Se realizó un análisis de agrupamiento para confrontar la población estudiada con trabajos previamente realizados en el ámbito mundial. Los resultados demuestran que todos los marcadores se encuentran en equilibrio de Hardy-Weinberg y presentan una alta heterocigocidad media observada debido a una gran diversidad alélica. Los sistemas ligados en el cromosoma 5 (CSF1PO y D5S818), se encuentran en equilibrio debido a que los alelos presentan frecuencias de recombinación alta y a la segregación independiente de los mismos. El análisis de agrupamiento muestra básicamente tres agrupaciones de las poblaciones de acuerdo a su localización geográfica. Las diferencias estadísticamente significativas entre Santander y las 21 poblaciones comparadas a nivel mundial reflejan una diferenciación genética que se ha podido formar en los diferentes continentes. La estructura genética de la población santandereana es producto de los diferentes procesos de mestizaje que han tomado lugar en los países de Sudamérica desde el año de 1542. El presente estudio proporciona una visión general de las frecuencias alélicas y genotípicas de la población santandereana, lo que a su vez contribuye al análisis de la estructura genética de dicha población.

* Proyecto de grado

** Facultad de ciencias, Escuela de Biología

SUMMARY

1. **TITLE:** ESTIMATION OF THE GENETIC VARIATION OF THE SANTANDER POPULATION USING MARKERS STR's *

2. **AUTHOR:** CASTILLO CASTRO SARA MARIA **

3. **KEY WORDS:** Genetic structure, Allelic frequencies, Santander population, STR's, Admixture process.

4. **CONTENT:**

The genetic structure of the population of Santander was analyzed using 15 short tandem repeat: FGA, TPOX, D8S1179, vWA, Penta E, D18S51, D21S11, THO1, D3S1358, Penta D, CSF1PO, D16S539, D7S820, D13S317, D5S818. The allelic frequencies were calculated from the genotypes using the gene counting method. possible divergence from Hardy-Weinberg equilibrium were determined by the exact test method through the GenePop v.1.2 computer program. With the same program was determined the linkage disequilibrium for loci located on the same chromosome. Also was performed an agrupment analysis for made a comparison between the study population with previous studies to achieve a global level. The results show that all markers are in Hardy-Weinberg equilibrium, besides to present high observed heterozygosity caused by a big allele diversity. The system which are linkage in the chromosome five (CSF1PO and D5S818) are in Hardy-Weinberg equilibrium because of the continuous mix between populations, because the alleles present a high recombination frequencies or because the alleles segregation is a independent way, though there was situate in the same chromosome. The agrupment analysis show basically 3 groups which are agruped the population agreement with the geographic localization. The statistic difference between Santander and the other populations compared a global level show the genetic to isolate which have been formed through the different continents. The genetic structure for the Santander population is product for the different admixture process which have take a place in the different suramerican countries since 1542. This study give a general vision about the allelic and genotypic frequencies from the Santander population, this permit the construction of a genetic database that contribute to the analysis for the genetic structure in this population.

* Proyecto de grado

** Facultad de ciencias, Escuela de Biología

INTRODUCCIÓN

Las poblaciones humanas tienen una estructura dinámica, que pueden sufrir cambios drásticos en su volumen creciendo expansivamente, o pueden sufrir un declive y quedar pocos o ningún sobreviviente. Muchos procesos poblacionales, como la tasa de natalidad, las enfermedades, las migraciones y el clima influyen en la magnitud de una población. Cuando estos factores cambian, la estructura genética de una población puede verse comprometida.

La composición genética de las poblaciones naturales está también determinada por las fuerzas evolutivas que definen su variabilidad, tales como la mutación y la selección natural que actúan en contra de la dispersión, manteniéndola y regulándola a expensas de una mejor adaptabilidad (Klug & Cumming, 1999).

Otro tipo de factor que puede intervenir en la composición genética actual de una población es la mezcla racial, mejor conocida como mestizaje entre humanos, a consecuencia de las migraciones entre comunidades (Sandoval, 1993). El problema estriba en cómo medir la estructura genética de una población y cómo saber cuando se produce un cambio en la frecuencia de los alelos.

El carácter de nación mestiza, acentuadamente hispánica, como es el caso de Colombia, es el resultado de un rápido proceso de mestizaje iniciado con la conquista, y que se acentúa a fines de la época colonial, este proceso se debió no solo a la magnitud de la contribución demográfica española sino también a la debilidad de la población

indígena y a las buenas condiciones que para los españoles presentaba el territorio ocupado (Hernández, 2001).

Colombia presenta un estado intermedio de mestizaje muy característico, que es de gran interés en el estudio de las poblaciones a nivel genético, frente a países, como Argentina y Uruguay donde la influencia indígena fue poca, o frente a países donde el aporte indígena en la cultura y en la demografía fue bastante notorio, como fue el caso de México, Bolivia, Ecuador y Guatemala (Bernal, 2000).

La manera más directa de describir la estructura genética de una población es medir las frecuencias alélicas en un determinado número de locus marcadores que se encuentran distribuidos en dicha población (Yunis, *et al.* 2002).

Los polimorfismos de tipo STR (*Short Tandem Repeat*) son unidades de DNA repetidas en tándem formadas por 2 a 6 nucleótidos, que proporcionan una herramienta útil en el entendimiento de la diversidad genómica y las afinidades entre poblaciones humanas (Edwards, *et al.* 1992). Recientemente, estos marcadores están siendo utilizados para inferir la historia evolutiva humana (Bowcock, *et al.* 1994 & Jorde, *et al.* 1995).

El análisis de las frecuencias génicas de estos marcadores en una población, permite conocer la desviación de las frecuencias alélicas, reflejando así los cambios evolutivos y el grado de mestizaje que ha tenido la población en estudio (Puertas, 1992); además de permitir la estimación de nuevos datos para ayudar en el análisis de la estructura genética del grupo poblacional (Calo, *et al.* 2003), la comparación de las frecuencias con otras poblaciones, permite acercarnos al

entendimiento del proceso de mestizaje de nuestra región, que sumado a la evidencia antropológica aportará al conocimiento de las distancias relativas entre diferentes culturas y ayudará a determinar los orígenes ancestrales de nuestro grupo poblacional contemporáneo.

El objetivo del presente estudio es estimar los patrones de diversidad genómica de la población Santandereana por medio de 15 marcadores moleculares tipo SRT, y analizar las afinidades entre grupos poblacionales que han sufrido procesos de mestizaje diferentes, además de aumentar nuestro conocimiento acerca de las características genéticas que presenta nuestra población.

1. OBJETIVOS

1.1 OBJETIVO GENERAL

Caracterizar la composición genética de la población santandereana a partir del cálculo de las frecuencias alélicas de 15 marcadores del sistema STRs (*Short Tandem Repeat*) utilizados en las pruebas de paternidad.

1.2 OBJETIVOS ESPECÍFICOS

- ☆ Describir la composición genética de la población santandereana a estudio.
- ☆ Comparar las frecuencias alélicas y los genotipos de los marcadores estudiados en la población santandereana con los reportados en poblaciones nacionales que se asume presentan características de mestizaje similares.
- ☆ Cotejar las frecuencias alélicas y los genotipos, de los marcadores estudiados en la población santandereana con los reportados en poblaciones latinoamericanas y europeas asumiendo diferentes características de mestizaje.

2. MARCO TEÓRICO

2.1 GENÉTICA POBLACIONAL

Uno de los objetivos de la genética de poblaciones consiste en estudiar y describir los procesos microevolutivos que operan sobre la estructura génica de determinada población a lo largo de las generaciones. Es por esto, que el análisis de la afinidad genética entre poblaciones permite valorar el efecto de los procesos de microevolución. Para ello es necesario caracterizar genéticamente a los grupos que se comparan a partir de marcadores biológicos (Klug, *et al.* 1999).

Es conocido las limitaciones de los caracteres fenotípicos para establecer las distancias biológicas en términos genéticos, ya que son consecuencia de la interacción del medio ambiente con el genotipo. Todo lo contrario ocurre con los marcadores moleculares, como los sistemas de grupos sanguíneos, enzima o proteínas, en los que la variabilidad fenotípica suele estar controlada exclusivamente a nivel genético con modelos de herencia basados en uno o pocos loci. Debido a que es limitado lo que se puede aprender de un locus que es igual en cada individuo, este tipo de estudios poblacionales depende de los loci polimorficos (Sandoval, 1993).

Antes de la década de los 90, la diversidad genética usualmente se deducía por medio de polimorfismos clásicos como los mencionados anteriormente, pero, durante los últimos 20 años los nuevos avances tecnológicos, sobre todo la invención de la PCR (Reacción en Cadena de la Polimerasa), han permitido profundizar más aún en el

conocimiento de la diversidad genética humana, lo que ha hecho posible actualmente emplear enzimas de restricción para producir una clase útil de marcador al nivel del ADN, es decir Fragmentos de Restricción de Longitud Polimórfica (RFLP), ADN mitocondrial (ADN_{mit}), inserciones Alu, minisatélites, polimorfismo de nucleótido único (SNP) y microsatélite (STR). Los microsatélites probablemente son la más importante clase de marcador utilizado en la actualidad debido a sus altas tasas de polimorfismos, mutaciones rápidas y distribución aleatoria a través del genoma (Cooper, 1999).

2.1.1 Estructura poblacional

La estructura de la población se refiere a los resultados de varios factores evolutivos tales como tamaño poblacional, flujo poblacional y apareamiento no aleatorio (Fontdevilla, *et al.* 1999; Klug, *et al.* 1999; Hartl, 2000; Suzuki, *et al.* 1996). Los genetistas poblacionales investigan el patrón de variación encontrado entre los individuos al interior de un grupo, y cómo estos patrones varían geográficamente y evolucionan con el tiempo (Klug, *et al.* 1999).

Los individuos pueden llevar sólo dos alelos diferentes de un gen dado. Un grupo de individuos puede llevar un gran número de alelos diferentes, dando lugar a un reservorio de diversidad genética. La diversidad que tiene una población se puede medir con la Ley de Hardy-Weinberg (Klug, *et al.* 1999).

Como se puede concluir en esta primera parte, el dinamismo de las poblaciones le permite crecer y expandirse o disminuir y contraerse en función a fenómenos determinantes de densidad poblacional, como son las tasas de nacimientos o fallecimientos, los procesos migratorios y la

fusión con otras poblaciones. Esto tiene consecuencias importantes y, con el tiempo, puede dar lugar a cambios en la estructura genética de las poblaciones (Hartl, 2000).

2.1.2 Estimadores genético – poblacionales

La dinámica de la diversidad genética depende de muchos parámetros poblacionales que son difíciles de estimar. A la hora de decidir entre un gran número de alternativas, es importante contar con un buen juego de marcadores genéticos (Zhivotovsky, 2000).

Para poder identificar los mecanismos que permiten mantener la variabilidad genética, un requisito indispensable es contar con estimadores apropiados de su magnitud y una caracterización adecuada de los patrones que se pudieran encontrar. Una forma de estimar la variabilidad genética es utilizar secuencias de ADN que pueden ser obtenidas de la población. Son estos estimadores genético-poblacionales los que permiten saber qué tan distantes están dos poblaciones, si hay o no fragmentación en la población desde el punto de vista genético (estructura poblacional), o si la variabilidad encontrada puede ser o no explicada por selección natural (Nei *et al.*, 1993).

2.1.3 Equilibrio de Hardy-Weinberg

La Ley de Hardy-Weinberg es uno de los conceptos fundamentales en genética de poblaciones y tiene tres propiedades importantes: 1. las frecuencias alélicas predicen las frecuencias genotípicas; 2. en el equilibrio, las frecuencias alélicas y genotípicas no cambian de generación en generación, y 3. el equilibrio se alcanza con una sola

generación de apareamiento al azar (Klug, *et al.* 1999; Fontdevilla, *et al.* 1999)

La demostración de este equilibrio es sencilla e implica que las frecuencias génicas (la frecuencia de cada gen o alelo) permanezcan constantes de generación en generación, siempre que la población cumpla las siguientes condiciones ideales:

- Ser lo suficientemente grande como para que todos los cambios que se produzcan en ella sigan las leyes del azar; y además no debe existir inmigración ni emigración.
- Los organismos componentes de esa población han de ser diploides y de reproducción al azar (panmixia).
- En esta población no hay mutaciones ni selección natural, de modo que los individuos tienen las mismas probabilidades de reproducirse, independientemente de sus genotipos (Klug, *et al.* 1999).

2.1.4 Factores que alteran las frecuencias alélicas en las poblaciones

En la naturaleza las poblaciones naturales son dinámicas y cambian en tamaño y estructura como parte de sus ciclos biológicos. La Ley de Hardy-Weinberg establece una serie de condiciones iniciales que nos permiten estimar las frecuencias alélicas y genotípicas en una población, siempre y cuando se cumplan las siguientes características: Condiciones iniciales de apareamiento al azar, ausencia de selección y mutación, e igual viabilidad y fecundidad; sin embargo, es difícil

encontrar poblaciones naturales que cumplan estas condiciones, si algunas de estas no se cumple, las frecuencias alélicas o genotípicas cambiarán de una generación a la siguiente (Fontdevilla, *et al.* 1999).

La alteración genética de una población sólo puede darse por factores como mutaciones, selección, influencias casuales, convergencias o divergencias individuales; el cambio genético que surja significa la perturbación del equilibrio.

2.1.5 Ligamiento

Sabemos que la mayor parte de los cromosomas constan de un número muy grande de genes y tienen suficiente DNA como para codificar miles de esas unidades (Suzuki, *et al.* 1996). Se dice que los genes que forman parte del mismo cromosoma están ligados y demuestran ligamiento en cruces genéticos (Klug, *et al.* 1999).

Debido a que el cromosoma es la unidad de transmisión en la meiosis, y no el gen, los genes ligados no son libres para transmitirse independientemente. Por el contrario, los alelos de todos los loci de un cromosoma se transmitirán como una unidad en la formación de los gametos. Sin embargo, en muchos casos no ocurre así. En la primera profase meiótica, cuando los homólogos se aparean, puede tener lugar un intercambio recíproco de fragmentos de cromosomas. Este fenómeno, que se llama entrecruzamiento, da lugar a una mezcla o recombinación de los alelos entre homólogos (Hartl, 2000).

El grado de entrecruzamiento entre los loci de un cromosoma es proporcional a la distancia que los separa. Así, el porcentaje de

gametos recombinantes varia de dependiendo de que loci están siendo considerados (Fontdevilla, *et al.* 1999).

Los genes en distintos loci no recombinan libremente si están ligados en el mismo cromosoma y su independencia esta limitada por la tasa de recombinación, que depende de la distancia física citogenética y de la constitución cromosomica en la región que los contiene (Hartl, 2000).

Las poblaciones presentan muchas veces una distribución de genotipos multilocus en desequilibrio debido a causas diversas, tales como disminuciones en el tamaño poblacional, que genera desviaciones al azar del equilibrio, o el efecto de la selección sobre combinaciones génicas favorables. El desequilibrio se traduce en excesos y defectos de gametos respecto a lo que esperaríamos en el equilibrio (Fontdevilla, *et al.* 1999).

La necesidad de considerar las asociaciones entre genes de distintos loci para entender la dinámica de las poblaciones se aprecia inmediatamente en el estudio del desequilibrio gamético. El azar junto con la ausencia de recombinación dentro de la inversión son las causantes de la presencia de un elevado desequilibrio gamético (Futuyma, 1998).

2.1.6 La heterociogidad en las poblaciones

Una de las aplicaciones prácticas de la Ley de Hardy-Weinberg es la estima de la frecuencia de los heterocigotos en una población (Hartl, 2000).

La heterocigosis es una medida práctica de la diversidad génica y se define como la proporción de individuos heterocigóticos por locus o, lo que es lo mismo, de loci heterocigóticos por individuo (Fontdevilla, *et al.* 1999). Este parámetro es indicativo de la diversidad génica en la población de individuos diploides con apareamiento al azar, pero no lo es en poblaciones con otro tipo de apareamiento y, desde luego, no es aplicable en organismos haploides (Futuyma, 1998).

Este valor de heterocigocidad representa la probabilidad de que los dos alelos del mismo *locus* tomados al azar en la población sean distintos. La variabilidad genética de una población se mide como heterocigocidad media por *locus*. Esta es la mejor estima de la variabilidad de una población si el estudio se realiza sobre un número elevado de *loci* escogidos al azar, ya que entre ese número de *loci* existirán unos caracteres monomórficos y otros polimórficos con altas o bajas frecuencias de los alelos raros, al igual que en genoma (Martínez, 1999).

2.2 ASPECTOS BASICOS DE LOS STR's

El estudio de marcadores polimórficos del ADN se ha convertido en una herramienta imprescindible en el análisis genético de interés forense, así como en la investigación biológica de la paternidad (Martínez, 1999; Yunis, *et al.*2002).

La aparición de los STR's ocurrió a principios de la década de los 90's, gracias al desarrollo del proceso denominado *Reacción en Cadena de la Polimerasa* (Yunis, *et al.*2002).

La secuencia de DNA repetidas en *tándem* en el genoma humano son una fuente de marcadores útiles para estudios genéticos, diagnósticos

médicos (Hearne, *et al.* 1992), pruebas de paternidad (Gehrig, *et al.* 1999) e identificación personal en ciencias forenses (Busque, 1997 & Pinheiro, *et al.* 2000). Los STR's están formados por varias repeticiones de 2-7 pb, siendo precisamente, , la variación en el número de veces que se repite la unidad de secuencia la base de su polimorfismo genético (Martinez,1999). Estos microsatélites se encuentran dispersados ampliamente a través del genoma en una frecuencia estimada de 1 microsatélite por 6-10 kb (Oldroyd, *et al.* 1995). Por su pequeño tamaño pueden ser amplificados a partir de una mínima cantidad de material viejo o altamente degradado.

El descubrimiento de este tipo de marcadores moleculares, permitió superar la mayoría de las limitaciones que ofrecían otros tipos de marcadores utilizados anteriormente en pruebas de paternidad como los VNTR.

Entre las ventajas de utilizar marcadores de tipo STR para determinar la estructura genética de una población están: 1. Se puede obtener de pequeñas muestras de ADN, y aun en aquel que se encuentra degradado, 2. Tienen un alto grado de especificidad para cada género, y 3. Son de tamaño pequeño (Busque et al, 1996). Todas estas ventajas han hecho que la tipificación de STR's mediante PCR se haya convertido en uno de los métodos de elección en la identificación genética de vestigios biológicos, así como en las pruebas de paternidad. Martínez, *et al.* 2002).

Tabla 1. Caracterización de los 15 STR's estudiados en la población Santandereana. *Datos tomados de Short Tandem Repeat Internet Database

Marcador	Localización	No. De alelos	Referencia
TPOX	2p23-23pter	10	*
Vwa	12p12-pter	14	Calo, <i>et al.</i> 2003
D18S51	18q21.3	20	Calo, <i>et al.</i> 2003
TH01	11p15.5	13	*
D3S1358	3p	12	Calo, <i>et al.</i> 2003
Penta D	21q11-21	16	Yunis, <i>et al.</i> 2002
D7S820	7q	12	Calo, <i>et al.</i> 2003
D13S317	13q22-31	11	Calo, <i>et al.</i> 2003
D5S818	5q21-q31	11	Calo, <i>et al.</i> 2003
D16S539	16q22-24	9	Calo, <i>et al.</i> 2003
CSF1PO	5q33.3-34	11	*
D21S11	21	27	Calo, <i>et al.</i> 2003
FGA	4q28	32	*
D8S1179	8	12	Calo, <i>et al.</i> 2003
Penta E	15q	19	Yunis, <i>et al.</i> 2002

2.2.1 Bases moleculares de los STR's

Los STR son regiones de ADN repetitivo que se encuentran repartidas a lo largo del todo el genoma humano y están compuestos por una secuencia de 2-7 pb que se repiten en *tandem*. Un gran número de estas regiones presenta un alto grado de polimorfismo genético de longitud cuya base molecular es la variación en el número de unidades de repetición. Los STR polimorficos se encuentran tanto en regiones génicas como extragénicas en el genoma humano. Los STR localizados en regiones génicas se presentan tanto en intrones y en regiones flanqueantes como en regiones codificantes (Edwards, *et al.* 1992).

Debido a su abundancia en el genoma humano, a su naturaleza Polimórfica y a las posibilidad de ser estudiados fácilmente mediante técnicas de PCR, Los STR se han convertido en uno de los marcadores de elección en el mapeo del genoma y en los analisis de ligamiento genético, así como en el campo de la identificación genética humana (Martinez, 1999).

2.2.2 STR tetraméricos autosómicos

El analisis de los productos de amplificación de marcadores STR permite detectar, además del alelo real, otros productos de amplificación minoritarios denominados bandas *shutter* que se forman durante el proceso de amplificación como consecuencia de fenómeno de *slippage* de la ADN polimerasa (Hauge, *et al.* 1993). En el caso de STR diméricos este tipo de productos hace muy difícil la designación de

los alelos. Por el contrario, los STR tetraméricos no presentan este tipo resultados y en la mayoría de los casos el artefacto de amplificación se limita a la aparición de una única banda minoritaria que presenta una unidad de repetición menos que el fragmento mayoritario que representa el alelo real (Urquhart , *et al.* 1995).

Aunque la causa fundamental del polimorfismo de los STR es la variabilidad en el numero de repeticiones en *tandem*, en muchos casos a este tipo de variabilidad se suma una microvariación de las secuencias de las unidades de repetición, así como una variabilidad estructural de las unidades de repetición (Martínez, 1999).

2.2.3 Nomenclatura y clasificación de los Alelos STR

Tanto en los sistemas manuales como en los semiautomáticos la clasificación de los alelos se realiza mediante la utilización de Láderes alélicos, de tal forma que se compara en un mismo gel la movilidad electroforética de los productos de amplificación de la muestra en cuestión con la movilidad electroforética de los alelos del líder.

La ISFH (Comision de ADN de la International Society of Forensic Haemogenetics indica que la denominación de los alelos se debe realizar de acuerdo al número de unidades de repetición (UR). En caso de que un determinado alelos presente una UR incompleta será designado por el numero de UR completas y el número de pares de bases de la UR incompleta, separando estos dos números por un punto.

2.2.4 Sistemas de Análisis de los STR

El análisis de los *loci* STR en el campo de la identificación genética humana ha facilitado enormemente gracias a la posibilidad de amplificar de manera simultánea varios loci en una única multiplex-PCR. Entre las ventajas de amplificación múltiple se encuentran la posibilidad de obtener de varios *loci* lo que aumenta el poder de discriminación del sistema, disminuyendo así el tiempo de análisis, así como la cantidad de ADN y el tiempo (Martinez, 1999). Los productos de amplificación de una reacción Multiplex-PCR pueden ser separados de acuerdo a su tamaño mediante electroforesis en gel de poliacrilamida y posteriormente detectados mediante dos métodos: detección manual mediante tinción con nitrato de plata y detección fluorescente semiautomática (Martinez, 1999; Yunis, *et al.* 2002).

Más recientemente se han desarrollado formatos que permiten la amplificación y análisis de 15STR más el marcador de género (amelogenina) mediante métodos fluorescentes (Yunis, *et al.* 2002). La automatización del procedimiento de STR, con marcación mediante diferentes fluorocromos, y su observación con métodos fluorescentes, se fundamenta en que si varios *loci* presentan similar tamaño, la diferenciación de ellos se hace con la utilización de diferentes fluorocromos que harán a parecer las señales en canales diferentes, como picos o como bandas, con un color distinto en el análisis (Martinez, 1999; Yunis, *et al.* 2002).

2.2.5 Los STR y estudios poblacionales

Este tipo de marcadores genéticos deben estudiarse junto con las frecuencias alélicas y genotípicas de la población correspondiente para poder estimar las frecuencias de algún marco genético particular (Sandoval, 1993). Por este motivo, varios estudios genéticos que usan sistemas multialélicos se han hecho en poblaciones humanas de todo el mundo representando los grandes grupos a nivel geográfico. (Nei, *et al.* 1993; Deka, *et al.* 1995; Perez-Lezaun, *et al.* 1997; Zhivotovsky, *et al.* 2000). De igual forma, desde hace algunos años un gran número de poblaciones locales específicas se han investigado (Furedi, *et al.* 1995; Martín, *et al.* 1995; Rose, *et al.* 1996; Busque, *et al.* 1997; Yamamoto, *et al.* 1999; Gehrig, *et al.* 1999; Tourret, *et al.* 1999; Pinheiro, *et al.* 2000; Drobic, *et al.* 2000; Klitz, *et al.* 2000;). En los santanderes y sus alrededores los datos de frecuencias alélicas están disponibles para algunos marcadores de tipo STR (Lemus-López, 2000). Adicionalmente, existen trabajos previos que reportan frecuencias para distintas partes del país (Acosta, 2002) y para grupos poblacionales con diferentes tipos de mestizaje (Yunis *et al.*, 2000; Yunis *et al.*, 2000; Yunis *et al.*, 2001; Yunis *et al.*, 2002; Yunis *et al.*, 2002).

2.3 EL MESTIZAJE

Como mestizaje suele definirse al producto de la fusión de razas o grupos humanos diferenciados, cada uno de ellos con propiedades biológicas particulares que se reflejan en la diversidad racial. El híbrido resultante forma parte de una nueva estirpe con características propias derivadas de las étnias que entran en contacto y que las pueden segregar (Sandoval, 1993).

Existen varias metodologías para el análisis de la composición racial de una población. Krieger en 1965 desarrolló un modelo matemático basado en estimaciones de máxima verosimilitud, en un estudio sobre la composición étnica en Nordesteales en Brasil, con base en un modelo tri-étnico; posteriormente, Reed en 1969 empleó por primera vez un modelo bi-étnico para el estudio de los genes caucasoides en los Afro-Americanos. Este tipo de estudio posee una limitación ya que muy rara vez se conoce exactamente la composición genética de las poblaciones ancestrales (Reed, 1969).

En Latinoamérica existen pocos estudios genéticos sobre mestizaje; sin embargo, se pueden resaltar los trabajos realizados por Lisker (1986, 1988, 1990) para la población mexicana, Saldaña (1962) y Krieger (1965) para Brasil, y Sandoval (1993) para Colombia, en donde se reafirma la composición tri-étnica característica de Iberoamérica.

2.3.1 El proceso de mestizaje en Santander

Como se mencionó anteriormente, el mestizaje biológico se inicia en el momento de la conquista; la población que encontramos en distintas regiones engloba elementos indígena, europeo y mestizo de unos y otros. La proporción de cada uno de estos elementos raciales y el lapso de tiempo transcurrido determinaran en cada caso la forma de cómo se distribuye el total de la población (Comas, 1972).

La región andina, como lo indica su nombre, comprende la zona montañosa conformada por las tres cordilleras que cruzan el país. Se clasifican en esta región los departamentos de Boyacá, Cundinamarca, Tolima, Huila, Santanderes, Risaralda, Quindío, Caldas y parte de Antioquía, Cesar, Bolívar, Cauca, Nariño y Valle del Cauca. El grado a

nivel global de mestizaje para esta región fue reportada por Sandoval (1993) con base en 8 sistemas sanguíneos e indica un porcentaje de 63.8% de componente caucásico, 29.50% de indígena y 6.72% de negro siendo muy similares a los porcentajes propuestos para la zona de los Santanderes donde el aporte caucásico es de 70.50%, el indígena 25.7% y el negro 3.78%.

2.3.2 Flujo génico y proceso de mestizaje

Las relaciones evolutivas de las poblaciones humanas pueden ser reveladas solo cuando un gran número de loci son examinados (Nei & Roychoudhury, 1974, 1982).

El flujo génico originado por la mezcla es un factor determinante en la generación de la diversidad biológica de las poblaciones, de modo que para el conocimiento de la evolución de los grupos humanos, las migraciones y el mestizaje deben ser tenidos en cuenta (Hernández, 2001). A pesar de la abrumadora mayoría de pobladores colombianos que son producto del mestizaje, Colombia es un país esencialmente multiétnico, ya que cerca de 700.000 personas se clasifican a sí mismas como indígenas y algo más de tres millones pertenecen a los grupos negros, descendientes de personas de origen africano, llegadas a nuestro país desde muy temprano en la conquista. Esta diversidad étnica puede percibirse a todos los niveles, desde el más estrictamente biológico hasta el componente cultural (Bernal, 2000).

El territorio actual de Santander a la llegada de los españoles estaba poblado por varios grupos indígenas, sin embargo, con el tiempo se fue imponiendo el mestizaje en todo el territorio, formándose así las primeras poblaciones como pueblos de indios; ciudades como

Bucaramanga, Chalará, Oiba y Curití nacieron como pueblos de indios; otras poblaciones importantes como Girón, San Gil, Zapatoca y Socorro fueron pueblos de blancos fundados y habitados por mestizos y españoles (Rodríguez, 1968). El flujo génico desde regiones españolas hacia nuestro país, la mezcla y la posterior “homogenización” de la población colombiana tiene claramente una fuerte influencia sobre el patrón genético actual de esta región (Sandoval, 1993).

2.4 ANÁLISIS DE AGRUPAMIENTO

Los Análisis de agrupamiento cumplen un rol central en la biología: es el marco de referencia general para cualquier investigación. Los resultados de dichas investigaciones muestran las relaciones entre los grupos estudiados, las que mejoran las clasificaciones o asociaciones volviendo al punto de partida (Scrocchi, *et al.* 1992).

Cualquier grupo de objetos, sean animados o no, puede ser ordenado de acuerdo con las características que presenten. Los resultados diferirán de acuerdo con las pautas utilizadas para el ordenamiento. Como las relaciones que se pueden llegar a mostrar son distintas, cada tipo de diagrama representativo recibe una denominación particular, pero al tener una forma general semejante se denominan en conjunto dendogramas.

Dentro de los diferentes análisis de agrupamiento existe un línea de análisis llamada taxonomía numérica, n este caso lo que se busca es cuantificar los procedimientos utilizados para clasificar los organismos o en nuestro caso las poblaciones; para esto se hace necesario la utilización de computadoras y programas especiales ya que la

realización manual resulta ser demasiado complicada (Sneath & Sokal, 1973).

Entre las diferentes técnicas de agrupamiento de la de UPGMA (unweighted pair-group method using arithmetic averages) donde la agrupación se realiza utilizando valores de similitud entre cada par de componentes de los grupos que se unen dados en nuestro caso por una matriz de distancias genéticas producto de las frecuencias alélicas presentes en cada población.

3. MATERIALES Y METODOS

3.1 TIPO DE ESTUDIO: el estudio genético es clasificado como un estudio analítico, transversal, controlado, ya que los datos aquí evaluados pertenecen a un pequeño porcentaje de la población y es básicamente descriptivo basado en documentos o resultados de exámenes previamente practicados.

3.2 ALMACENAMIENTO DE DATOS

El estudio genético-poblacional se realizó a partir de los datos de 15 marcadores de tipo STRs utilizados en pruebas de paternidad proporcionados por el Departamento de Genética del Laboratorio Higuera Escalante de la FOS-CAL, colectados entre los meses de septiembre de 2001 a septiembre de 2003.

3.3 DESCRIPCION MUESTRA

La población de estudio esta conformada por 279 individuos no relacionados los cuales asistieron al laboratorio para practicarse pruebas de paternidad. El número de individuos estudios no representan un tamaño de muestra significativo para la población y por esto solo se hacen inferencias con respecto a la población Santandereana estudiada. El 39% de las personas que acudieron al laboratorio asistieron de forma privada mientras que el 61% fueron por citación del juzgado. Debido a los altos costos que conllevaría realizar un estudio poblacional con una muestra realmente representativa de determinada población se hace necesario implementar herramientas de

este tipo ya que pueden proporcionar una idea general de cómo se comporta la población a nivel genético.

3.4. VARIABLES

La base de datos se construyó en una hoja de cálculo de acuerdo con las siguientes variables disponibles en cada uno de los registros de las pruebas de paternidad.

Se tuvieron en cuenta dos tipos de variables: Variables sociodemográficas y variables Genéticas. Las primeras contemplan datos como raza: Registrada por la genetista a la hora de la toma de la muestra según los siguientes criterios:

A cada paciente se le clasifico por inspección de sus rasgos faciales. Color de la piel y tipo de cabello de acuerdo a si es de raza blanca, negra o indígena sea pura o con predominio marcado de alguna de ellas, de la siguiente manera:

<i>Características</i>	<i>Predominio de Raza</i>		
	<i>Indígena</i>	<i>Negra</i>	<i>Blanco</i>
Piel	Puede ser clara u oscura	oscura	Clara
Cabello	Liso (negro)	ensortijado(negro)	Liso(claro/ oscuro)
Pómulos	Sobresalientes	No sobres.	No sobres.
Labios	Delgados	Gruesos	Delgados
Nariz	Fina	ancha	Fina

En caso de poca claridad respecto a la raza de los pacientes, Los individuos fueron clasificados “Mestizos” con las diversas combinaciones posibles, teniendo en cuenta las características citadas en la tabla.

Adicionalmente se contemplaron variables como el parentesco verificado en la entrevista con los pacientes y comprobado con los apellidos.; Departamento del cual era natural el paciente, el cual hace referencia al lugar de nacimiento; Procedencia, para evitar un mayor sesgo en cuanto a la calidad de la población ya que se podría pensar que la posición social puede estar relacionada con la raza.

En cuanto a las variables genéticas contempladas se tuvo en cuenta el marcador a estudiar y el genotipo del individuo. Para el caso del marcador TPOX, el número de pacientes fue de 278 ya que el marcador no amplificó para dicho sistema en un individuo.

3.5. ASPECTOS ETICOS.

La Presente investigación se clasifica como un estudio de “Investigación de Riesgo Mínimo”, según el numeral b del Artículo 10, de la resolución 008430 de 1993 del Ministerio de Salud de la República de Colombia. Por lo tanto, el Comité de Ética de la Facultad de Medicina de la UNAB, considera que no hay necesidad de obtener un consentimiento Informado por escrito para la realización de la presente investigación. Esta decisión se tomó con base en el Parágrafo Primero del Artículo 16 de la Resolución anteriormente anotada.

3.6. MARCADORES

Los 15 marcadores de tipo STR utilizados fueron: Penta D, Penta E, D21S11, D3S1358, FGA, D8S1179, D18S51, CSF1PO, TPOX, TH01, vWA, D16S539, D7S820, D13S317, D5S818. Por primera vez se reportan frecuencias de los sistemas Penta D y penta E para la población santandererana.

3.7 VARIABILIDAD GENETICA

3.7.1 Estimación de las frecuencias y heterocigocidad esperada y observada

Las frecuencias genotípicas se calcularon por el método de conteo directo, es decir, se determinó cuántos individuos presentaban determinado genotipo y se dividió por el número total individuos. Esta estimación se realizó para cada marcador.

Las frecuencias alélicas fueron calculadas a partir de las frecuencias genotípicas de igual forma con ayuda del programa GenePop, versión 1.2 (Raymond and Rousset 1995). El algoritmo utilizado por el programa se basa en el desarrollo de la fórmula donde se conocen la frecuencias de los genotipos homocigotos y heterocigotos para cada loci.

$$\text{Frecuencia alélica} = AA + \frac{1}{2} Aa$$

Donde: AA = Frecuencia del genotipo homocigoto

Aa = Frecuencia del genotipo heterocigoto

La heterocigocidad observada y esperada para cada marcador se obtienen a partir de la sumatoria de las frecuencias de los genotipos heterocigotos.

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

Siendo p_i = Frec. Génica del alelo i en un locus de m alelos (Puertas 1992)

La heterocigocidad para la población fué calculada a partir de la frecuencia de los heterocigotos para cada locus y dividiendo por el número total de loci.

3.7.2 Pruebas para determinar equilibrio de Hardy-Weinberg.

Los cálculos para determinar equilibrio de Hardy-Weinberg fueron desarrollados realizados mediante el desarrollo del polinomio:

$$(p+q+r\dots+z)^2 = P^2+q^2+ r^2 +\dots+ 2pq + 2pr + \dots 2pz$$

partiendo del hecho de que:

p, q, r, \dots son las frecuencias alélicas para el locus i ésimo.

La prueba usual para validar los datos observados con HWE es la prueba de Chi-cuadrado. Esta prueba es estadísticamente simbolizada X^2 , y sobre la hipótesis de HWE, el X^2 tiene aproximadamente una distribución de Chi-cuadrado. El valor de X^2 es calculado así:

$$\chi^2 = \sum \frac{(obs - esp)^2}{esp}$$

Asociado a esto, el cálculo de grados de libertad es: número de clases de datos, menos uno, menos el número de parámetros estimados a partir de los datos (Klug, *et al.* 1999).

3.7.3 Desequilibrio de ligamiento para loci localizados en el mismo cromosoma

Asociaciones al azar de alelos a partir de loci genéticamente no relacionados pueden ofrecer evidencia para determinar la subheterogeneidad de la población, la cual será medida mediante un análisis de independencia de frecuencias genotípicas de pares de locus a partir de una tabla de contingencia (Busque et al, 1979).

El desequilibrio de ligamiento fué calculado mediante la creación de tablas de contingencia para todos los pares de loci en la población, posteriormente se realizó una prueba de probabilidad (Prueba exacta de Fisher) para cada tabla.

3. 8 COMPARACIÓN ENTRE POBLACIONES

3.8.1 Diferenciación entre poblaciones

La selección de trabajos a comparar se realizó mediante una búsqueda sistemática en una base de datos internacional

(www.ncbi.nlm.nih.gov) utilizando el término MeSh **Minisatellite repeats** y tomando artículos publicados en los últimos 5 años en donde se revisaran mínimo el 50% (8 sistemas) de los marcadores estudiados en la población santandereana analizada.

Los datos utilizados en la comparación fueron tomados de trabajos previos publicados en la literatura para diferentes poblaciones en el ámbito mundial (Yunis *et al*, 2000; Benítez-Paéz *et al*, 2003; Rey *et al*, 2003; Gómez *et al*, 2003; Luna-vázquez *et al*, 2003; Bydlowski *et al*, 2003; Andrade *et al*, 2003; Chiurrillo *et al*, 2003; Gehrig *et al*, 1999; Biondo *et al*, 2001; Anorin *et al*, 2001; Paredes *et al*, 2003); Yamamoto *et al*, 1999; Ahmed *et al*, 2001; Chantratita *et al*, 2001; Zhang *et al*, 2003; Seah *et al*, 2003 y Cho *et al*, 2003).

Se realizó una prueba de chi-cuadrado global con las frecuencias alélicas de cada marcador, para la diferenciación entre poblaciones (Raymond & Rousset, 1995). Esta prueba calcula un valor exacto de p para cada par de poblaciones teniendo como referencia la población santandereana. Estos cálculos fueron realizados con el programa Epiinfo v. 6.04 (Dean, et al. 1994), mediante el uso de tablas de contingencia.

Adicionalmente las poblaciones fueron comparadas por medio de las heterocigocidades observadas en cada una de las 21 poblaciones.

3.8.2 Análisis de agrupamiento

El uso de frecuencias alélicas en la construcción de dendogramas, dan la idea general de que las subpoblaciones se organizan de manera compacta representando una población.

Los cálculos de las distancias se realizaron a partir de una matriz de datos con las frecuencias alélicas de cada población. El análisis de agrupamiento se realizó mediante el algoritmo de agrupación jerárquica UPGMA y distancias Euclidianas. La construcción del respectivo dendograma para cada tipo de agrupación, fue realizada usando el programa Statistica 4.0

4. RESULTADOS

4.1 ESTIMACIÓN DE FRECUENCIAS, EQUILIBRIO DE HARDY-WEINBERG Y HETEROCIGOCIDAD OBSERVADA Y ESPERADA EN LA POBLACIÓN SANTANDEREANA.

La tabla 2 resume la probabilidad calculada para determinar el equilibrio o desequilibrio de Hardy-Weinberg de los 15 marcadores estudiados en la población santandereana. Todos los marcadores se encuentran en equilibrio de Hardy-Weinberg

Tabla 2. Prueba para determinar el equilibrio de Hardy-Weinberg y Heterocigocidad observada y esperada en la población Santandereana. En negrilla se resaltan los marcadores más polimorficos, los cuales a su vez son los que presentan una mayor heterocigocidad

Marcador	H Obs	He	H-W (p)	n
FGA	0,8996	0,8781	0,8196	279
TPOX	0,6511	0,6488	0,0701	278
D8S1179	0,7635	0,7886	0,7871	279
vWA	0,7526	0,7601	0,1366	279
Penta E	0,8887	0,9208	0,9931	279
D18S51	0,8678	0,883	0,8904	279
D21S11	0,8358	0,8382	0,4487	279
TH01	0,7458	0,744	0,8538	279
D3S1358	0,7599	0,864	0,1905	279
Penta D	0,839	0,8455	0,5243	279
CSF1PO	0,7069	0,7259	0,0777	279
D16S539	0,8319	0,7967	0,6647	279
D7S820	0,7744	0,793	0,9915	279
D13S317	0,8468	0,8184	0,6836	279
D5S818	0,7027	0,7311	0,0956	279

4.1.1 Locus TPOX

Se encontraron 6 alelos (Figura 1) para una población de 279 individuos, con frecuencias alélicas que varían entre 0.4838 para el alelo 8 y 0.0071 para el alelo 6 (Tabla 3).

Figura 1. Distribución de las frecuencias alélicas en la población Santandereana para el locus TPOX.

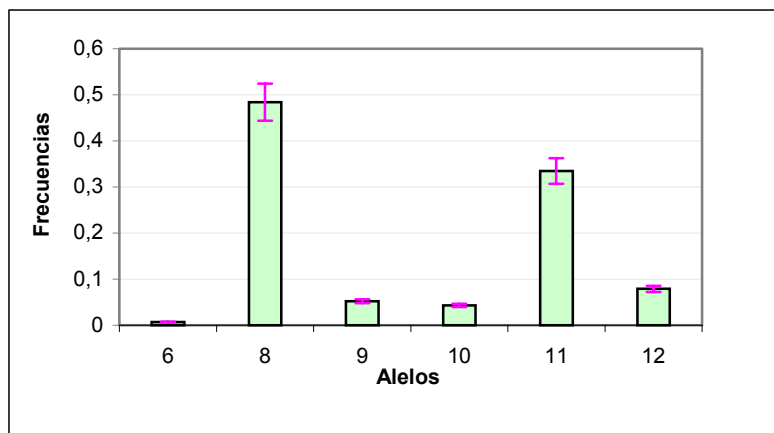


Tabla 3. Frecuencias alélicas encontradas en la población Santandereana para el locus TPOX.

Alelo	Ind.	F. alélica	EE
6	4	0,007194245	0,000598
8	269	0,48381295	0,0402157
9	29	0,052158273	0,0043355
10	24	0,043165468	0,003588
11	186	0,334532374	0,0278072
12	44	0,079136691	0,006578
6	556	1	

La población estudiada muestra 15 genotipos de 21 esperados, de los cuales 2 son homocigotos (13.33%) y 13 son heterocigotos (86.66%)

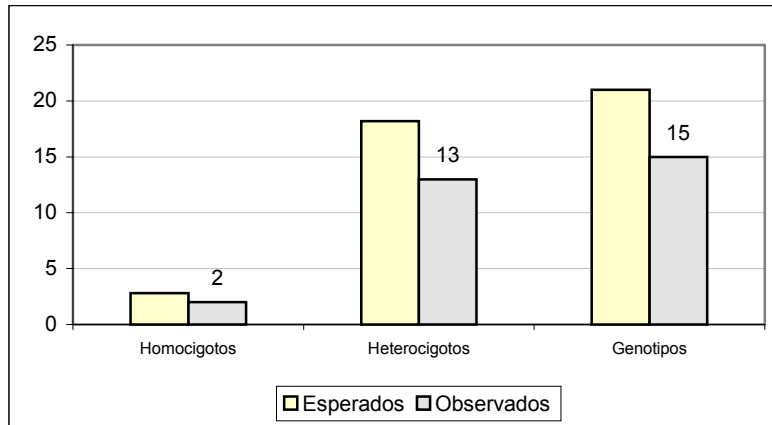
(Figura 2). El genotipo que predomina es el 11/8 presente en 91 individuos (32.73%) y los que se encuentran en menor proporción son el 9,3/7 y el 10/6 con un individuo cada uno (Tabla 4). La heterocigocidad observada presento un valor de 0.6511.

Tabla 4. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus TPOX.

Genotipos	Obs	Esp	Frecuencia
8/6	2	1.939	0,0071942
8/8	64	64.948	0,2302158
9,3/7	1	0.209	0,0035971
9,3 / 9	14	14.056	0,0503597
10 / 6	1	0.173	0,0035971
10/8	16	11.632	0,057554
10/9	2	1.254	0,0071942
11/ 8	91	90.151	0,3273381
11/ 9	7	9.719	0,0251799
11/10	3	8.043	0,0107914
11/11	33	31.000	0,118705
12/8	18	21.326	0,0647482
12/ 9	5	2.299	0,0179856
12/10	2	1.903	0,0071942
12/11	19	14.746	0,0683453
15	278		1

Los genotipos observados y esperados no presentan diferencia significativa ($X^2=1.71$, 9 g.l., $0.99 < p < 0.995$). El locus se encuentra en equilibrio de Hardy–Weinberg ($P=0.0701$).

Figura 2. Distribución de los genotipos observados y esperados en la población Santandereana para el locus TPOX.



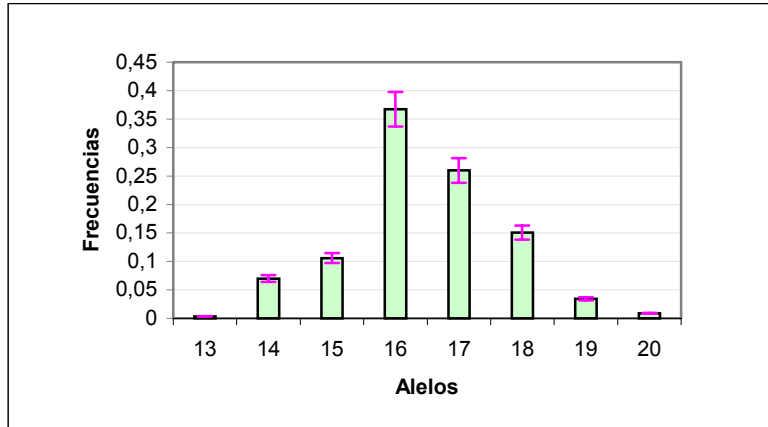
4.1.2 Locus vWA

Para este sistema se encontraron 8 alelos (Figura 3), siendo el más común el 16 con una frecuencia de 0.3673 y el que menos se detectó fue el 13 con una frecuencia de tan solo 0.0035 (Tabla 5).

Tabla 5. Frecuencias alélicas encontradas en la población Santandereana para el locus vWA.

Alelos	Ind.	F. alélica	EE
13	2	0,003584	0,000297396
14	39	0,069892	0,005799217
15	59	0,105735	0,008773175
16	205	0,367384	0,030483065
17	145	0,259857	0,021561192
18	84	0,150538	0,012490622
19	19	0,03405	0,00282526
20	5	0,008961	0,000743489
8	558	1	

Figura 3. Distribución de las frecuencias alélicas en la población Santandereana para el locus vWA.



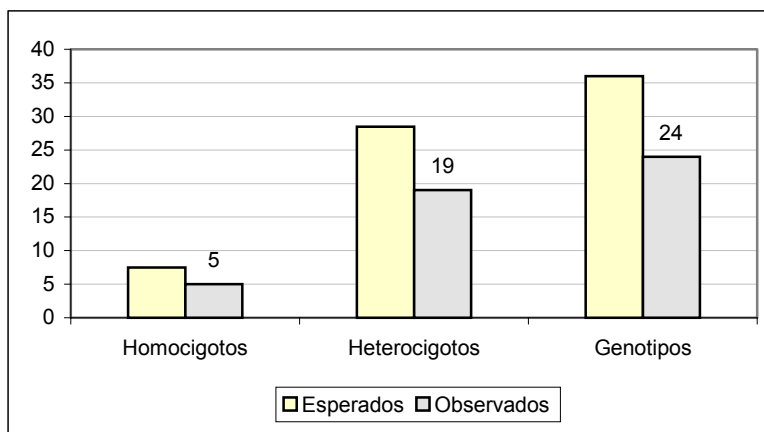
A diferencia del marcador anterior este locus presentó un mayor número de genotipos observados, encontrándose 24 de 36 esperados. El porcentaje de homocigotos y heterocigotos fué de 20.83% y 79.16% respectivamente (Figura 4). El genotipo 16/17 presente en 19.71% de la población es el de mayor frecuencia (Tabla 6). La Heterocigocidad observada para este marcador fue de 0.7526. Los genotipos observados y esperados no presentan diferencia significativa ($\chi^2=4, 16$ g.l., $p > 0.995$). No hubo desviaciones del equilibrio de Hardy-Weinberg ($P=0.1366$).

Tabla 6. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus vWA.

Genotipos	Obs	Esp	Frecuencia
14 /13	1	0.140	0,0035842
14 /14	1	1.330	0,0035842
15 /14	4	4.131	0,0143369
15 /15	4	3.072	0,0143369
16 /14	11	14.354	0,0394265
16 /15	20	21.715	0,0716846

16 /16	39	37.540	0,1397849
17 /14	13	10.153	0,046595
17 /15	20	15.359	0,0716846
17 /16	55	53.366	0,1971326
17 /17	18	18.743	0,0645161
18 /14	6	5.882	0,0215054
18 /15	7	8.898	0,0250896
18 /16	34	30.916	0,1218638
18 /17	15	21.867	0,0537634
18 /18	7	6.259	0,0250896
19 /13	1	0.068	0,0035842
19 /14	1	1.330	0,0035842
19 /16	7	6.993	0,0250896
19 /17	5	4.946	0,0179211
19 /18	5	2.865	0,0179211
20 /14	1	0.350	0,0035842
20 /17	1	1.302	0,0035842
20 /18	3	0.754	0,0107527
24	279		1

Figura 4. Distribución de los genotipos observados y esperados en la población Santandereana para el locus vWA.



4.1.3 Locus D18S51

Las frecuencias alélicas fueron desde 0.1666 para el alelo 17 hasta 0.0017 para el alelo 24 (Tabla 7), dentro de un grupo formado por 15 alelos (Figura 5).

Figura 5. Distribución de las frecuencias alélicas en la población Santandereana para el locus D18S51.

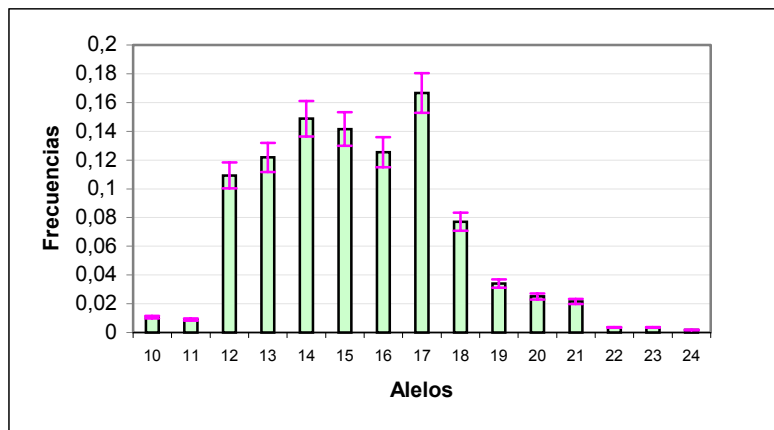


Tabla 7. Frecuencias alélicas encontradas en la población Santandereana para el locus D18S51.

Alelos	Ind.	F. alélica	EE
10	6	0,0107526	0,00089219
11	5	0,0089605	0,00074349
12	61	0,1093189	0,00907057
13	68	0,1218637	0,01011146
14	83	0,14874552	0,01234192
15	79	0,1415770	0,01174713
16	70	0,1254480	0,01040885
17	93	0,1666666	0,0138289
18	43	0,0770609	0,00639401
19	19	0,0340501	0,00282526
20	14	0,0250896	0,00208177
21	12	0,0215053	0,00178437

22	2	0,0035842	0,0002974
23	2	0,0035842	0,0002974
24	1	0,0017921	0,0001487
15	558	1	

En los 279 individuos que conforman la población de estudio se encuentran presentes un total de sesenta genotipos de 120 esperados, todos con frecuencias muy bajas para cada uno de ellos (Tabla 8), lo que se ve reflejado en un alto valor de heterocigocidad observada: 86.78. El número de genotipos homocigotos y heterocigotos observados es de 7 (11.66%) y 53 (88.33%) respectivamente (Figura 6). Los genotipos observados y esperados no presentan diferencia significativa ($X^2=30,264$ g.l., $0.97 < p < 0.99$). El marcador se encontró en equilibrio de Hardy –Weinberg ($P=0.8904$).

Figura 6. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D18S51.

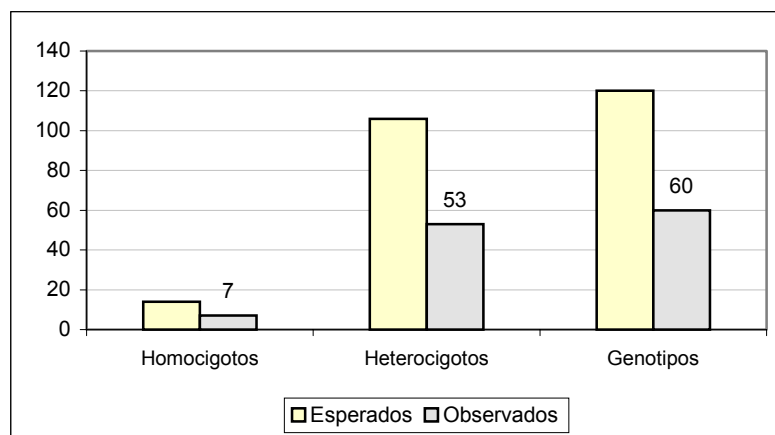


Tabla 8. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D18S51

Genotipos	Obs.	Esp.	Frecuencia
12 /11	1	0.548	0,00358423
12/12	6	3.285	0,02150538
13 /11	1	0.610	0,00358423
13/12	7	7.447	0,02508961
13 /13	7	4.090	0,02508961
14/10	2	0.894	0,00716846
14/11	1	0.745	0,00358423
14/12	8	9.090	0,02867384
14 /13	9	10.133	0,03225806
14 /14	7	6.110	0,02508961
15 /10	2	0.851	0,00716846
15 /12	4	8.652	0,01433692
15 /13	7	9.645	0,02508961
15 /14	12	11.772	0,04301075
15 /15	3	5.531	0,01075269
16/12	9	7.666	0,03225806
16 /13	5	8.546	0,01792115
16 /14	11	10.431	0,03942652
16 /15	11	9.928	0,03942652
16 /16	5	4.336	0,01792115
17 /10	2	1.002	0,00716846
17 /11	2	0.835	0,00716846
17 /12	12	10.185	0,04301075
17 /13	11	11.354	0,03942652
17 /14	15	13.858	0,05376344
17 /15	19	13.190	0,06810036
17 /16	10	11.688	0,03584229
17 /17	7	7.680	0,02508961
18 /12	3	4.709	0,01075269
18 /13	6	5.250	0,02150538
18 /14	5	6.408	0,01792115
18 /15	10	6.099	0,03584229
18 /16	5	5.404	0,01792115
18 /17	4	7.180	0,01433692
18 /18	2	1.621	0,00716846
19/12	2	2.081	0,00716846
19 /13	4	2.320	0,01433692

19 /14	2	2.831	0,00716846
19 /15	4	2.695	0,01433692
19 /16	3	2.388	0,01075269
19 /17	2	3.172	0,00716846
19 /18	1	1.467	0,00358423
20/12	2	1.533	0,00716846
20 /13	2	1.709	0,00716846
20 /14	2	2.086	0,00716846
20 /15	2	1.986	0,00716846
20 /16	2	1.759	0,00716846
20 /17	2	2.338	0,00716846
20 /18	1	1.081	0,00358423
21 /13	2	1.465	0,00716846
21 /14	2	1.788	0,00716846
21 /15	2	1.702	0,00716846
21 /16	3	1.508	0,01075269
21 /18	2	0.926	0,00716846
21 /20	1	0.302	0,00358423
22 /18	1	0.154	0,00358423
22 /19	1	0.068	0,00358423
23 /16	1	0.251	0,00358423
23 /18	1	0.154	0,00358423
24/12	1	0.110	0,00358423
60	279		1

4.1.4 Locus TH01

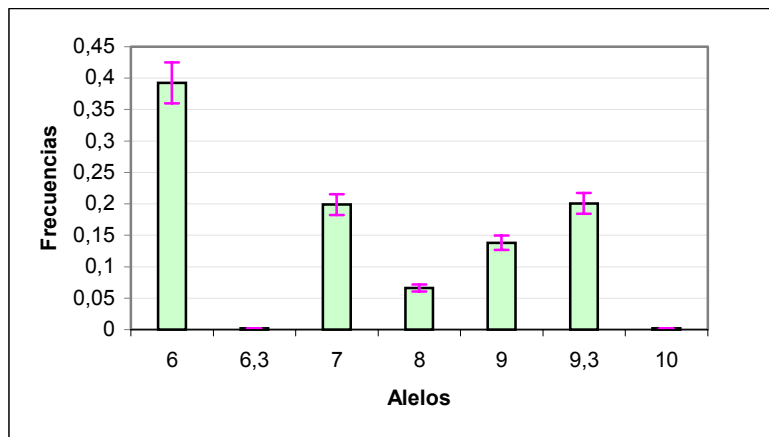
Siete alelos (Figura 7) conforman los 17 genotipos encontrados en la población de 28 esperados, siendo los alelos de mayor presencia el 6, el 9,3 y el 7 con frecuencias de 0.3924, 0.2007 y 0.1989 respectivamente (Tabla 9) y los de menor segregación el 6,3 y el 10 presentes con una frecuencia de 0.0017 cada uno.

Tabla 9. Frecuencias alélicas encontradas en la población Santandereana para el locus TH01

Alelo	Ind	F. alélica	EE
6	219	0,392473118	0,032564835
6,3	1	0,001792115	0,000148698
7	111	0,198924731	0,016505465

8	37	0,066308244	0,005501822
9	77	0,137992832	0,011449737
9,3	112	0,200716846	0,016654162
10	1	0,001792115	0,000148698
7	558	1	

Figura 7. Distribución de las frecuencias alélicas en la población Santandereana para el locus TH01.



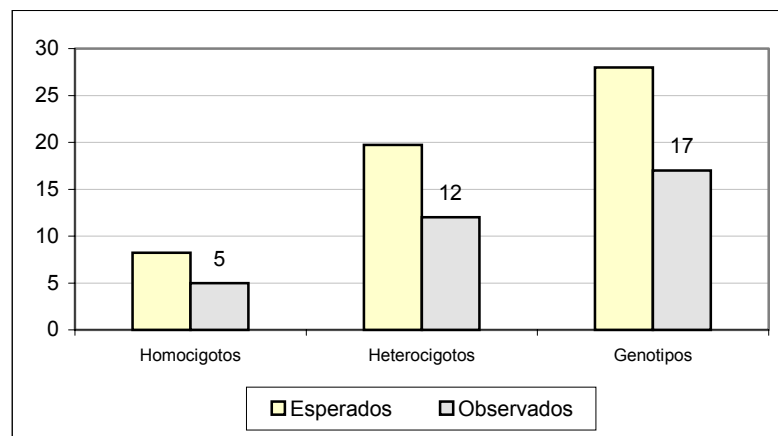
Se encontró un total de 17 genotipos de 28 esperados (Figura 8) con frecuencias que varían desde 0.1827 hasta 0.0035 siendo el genotipo más común el 6,3/6 presente en 18.27% de la población estudiada (Tabla 10). La heterocigocidad observada fue de 0.7458. No hay diferencia significativa entre el número de genotipos observados y esperados ($\chi^2=4.32$, 10 g.l., $0.90 < p < 0.95$). El marcador se encontró en equilibrio de Hardy–Weinberg ($P=0.8538$).

Tabla 10. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus TH01.

Genotipos	Obs	Esp	Frecuencias
6/6	39	42.856	0,139784946
6,3 / 6	1	0.393	0,003584229

6,3 / 6	51	43.643	0,182795699
7 / 7	10	10.961	0,035842294
8 / 6	16	14.548	0,05734767
8 / 7	5	7.373	0,017921147
8 / 8	1	1.196	0,003584229
9 / 6	25	30.275	0,089605735
9 / 7	15	15.345	0,053763441
9 / 8	6	5.115	0,021505376
9 / 9	9	5.253	0,032258065
9,3 / 6	47	44.036	0,168458781
9,3 / 7	20	22.320	0,071684588
9,3 / 8	8	7.440	0,028673835
9,3 / 9	13	15.483	0,046594982
9,3 / 9,3	12	11.160	0,043010753
10 / 6	1	0.393	0,003584229
17	279		1

Figura 8. Distribución de los genotipos observados y esperados en la población Santandereana para el locus TH01.



4.1.5 Locus D3S1358

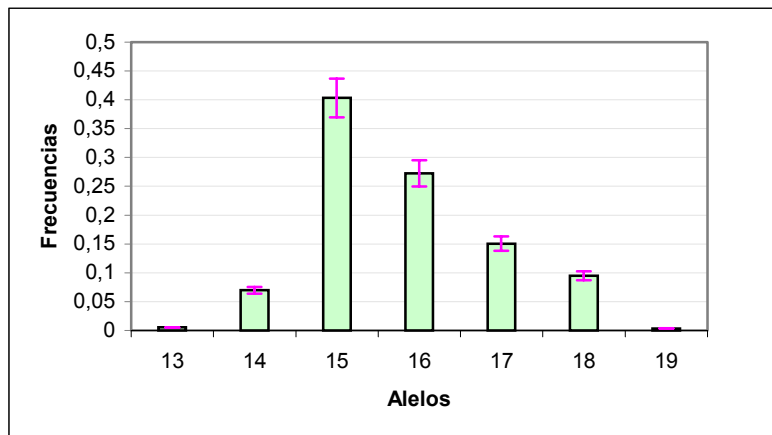
El número de alelos detectados para este marcador es de 7 (Figura 9). Los alelos 15 y 16 son los más comunes con frecuencias de 0.4032 y

0.2724 respectivamente, que equivalen a más de la mitad de la población estudiada (Tabla 11).

Tabla 11. Frecuencias alélicas encontradas en la población Santandereana para el locus D3S1358.

Alelo	Ind	F. alélica	EE
13	3	0,005376344	0,000446094
14	39	0,069892473	0,005799217
15	225	0,403225806	0,033457023
16	152	0,272401434	0,022602078
17	84	0,150537634	0,012490622
18	53	0,094982079	0,007880988
19	2	0,003584229	0,000297396
7	558	1	

Figura 9. Distribución de las frecuencias alélicas en la población Santandereana para el locus D3S1358.



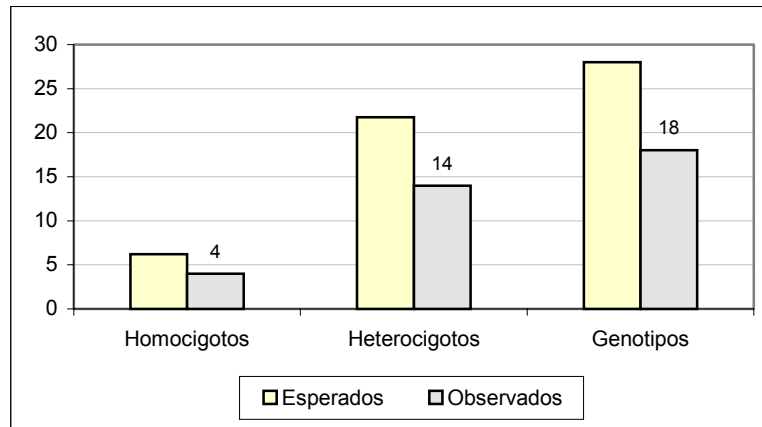
Se encontraron 18 genotipos observados de 28 esperados (Figura 10), con frecuencias relativamente constantes aunque sobresalió el genotipo 16/15 con una frecuencia de 0.2329 (Tabla12). Los genotipos homocigotos y heterocigotos presentan valores observados de 4 y 14

respectivamente (Figura 10) y la población tiene un valor de heterocigocidad observada de 0.7599.

Tabla 12. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D3S1358.

Genotipos	Obs	Esp	Frecuencia
15 /14	20	15.754	0,071684588
15 /15	41	45.242	0,146953405
16 /13	2	0.819	0,007168459
16 /14	15	10.643	0,053763441
16 /15	65	61.400	0,23297491
16 /16	19	20.603	0,068100358
17 /13	1	0.452	0,003584229
17 /14	3	5.882	0,010752688
17 /15	35	33.932	0,125448029
17 /16	21	22.923	0,075268817
17 /17	5	6.259	0,017921147
18 /14	1	3.711	0,003584229
18 /15	23	21.409	0,082437276
18 /16	11	14.463	0,039426523
18 /17	13	7.993	0,046594982
18 /18	2	2.474	0,007168459
19 /17	1	0.302	0,003584229
19 /18	1	0.190	0,003584229
18	279		1

Figura 10. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D3S1358.



No hay diferencia significativa entre el número de genotipos observados y esperados ($\chi^2=3.57, 11$ g.l., $0.975 < p < 0.99$). El equilibrio de Hardy-Weinberg se cumplió ($P=0.1905$).

4.1.6 Locus Penta D

De los 8 alelos encontrados (Figura 11) los que segregan con mayor frecuencia son el 10 y el 11 con frecuencias de 0.2186 y 0.1720 respectivamente, de manera contraria el menos segregado es el 5 y el 17 con una frecuencia de tan solo 0.0017 (Tabla 13).

Figura 11. Distribución de las frecuencias alélicas en la población Santandereana para el locus Penta D.

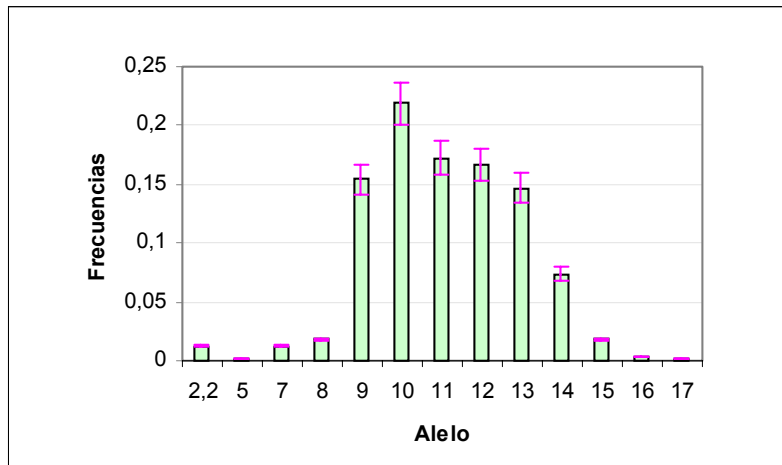


Tabla 13. Frecuencias alélicas encontradas en la población Santandereana para el locus Penta D.

Alelo	Ind	F. alélica	EE
2,2	7	0,012544803	0,001040885
5	1	0,001792115	0,000148698
7	7	0,012544803	0,001040885
8	10	0,017921147	0,001486979
9	86	0,154121864	0,012788018
10	122	0,218637993	0,018141141
11	96	0,172043011	0,014274996
12	93	0,166666667	0,013828903
13	82	0,146953405	0,012193226
14	41	0,073476703	0,006096613
15	10	0,017921147	0,001486979
16	2	0,003584229	0,000297396
17	1	0,001792115	0,000148698
14	558	1	

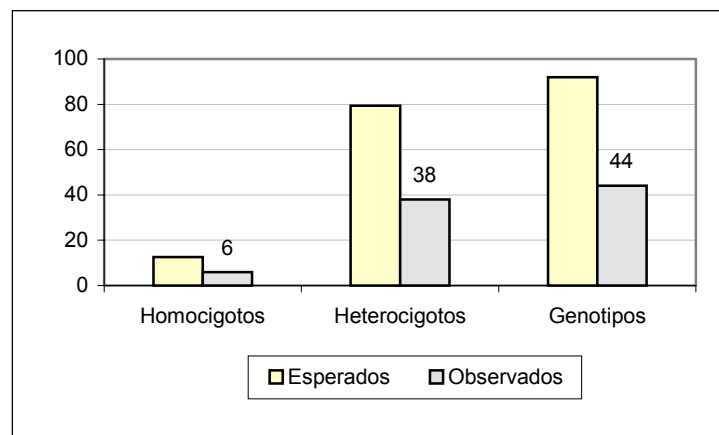
Este marcador presentó un total de 44 genotipos de 91 esperados. El heterocigoto 11/10 presentó una frecuencia de 0.0860 correspondiente a 24 individuos de la población (Tabla 14); Los genotipos homocigotos observados fueron 12 y los heterocigotos 32 (Figura 12). La heterocigocidad observada fue de 0.8468. No hay diferencia significativa entre el número de genotipos observados y esperados ($X^2=1.36$, 21 g.l., $p>0.995$). El equilibrio de Hardy-Weinberg se cumplió ($P=0.5243$).

Tabla 14. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus Penta D.

Genotipos	Obs	Esp	Frecuencia
9/ 5	1	1.081	0,003584229
9/ 7	1	1.081	0,003584229
9/ 8	3	1.544	0,010752688
9/ 9	6	6.562	0,021505376
10/ 2,2	1	0.219	0,003584229
10/ 7	1	1.533	0,003584229
10/ 8	1	2.190	0,003584229
10/ 9	19	18.837	0,068100358
10/10	15	13.251	0,053763441
11/ 7	1	1.206	0,003584229
11/ 8	2	1.724	0,007168459
11/ 9	13	14.822	0,046594982
11/10	24	21.027	0,086021505
11/11	9	8.187	0,032258065
12/ 5	3	1.169	0,010752688
12/ 8	3	1.670	0,010752688
12/ 9	15	14.359	0,053763441
12/10	23	20.370	0,082437276
12/11	19	16.029	0,068100358
12/12	4	7.680	0,014336918
13/ 5	2	1.031	0,007168459
13/ 7	1	1.031	0,003584229
13/ 9	12	12.661	0,043010753
13/10	17	17.961	0,0609319
13/11	11	14.133	0,039426523

13/12	13	13.691	0,046594982
13/13	9	5.962	0,032258065
14/ 5	1	0.515	0,003584229
14/ 7	2	0.515	0,007168459
14/ 8	1	0.736	0,003584229
14/ 9	8	6.330	0,028673835
14/10	6	8.980	0,021505376
14/11	4	7.066	0,014336918
14/12	6	6.846	0,021505376
14 /13	7	6.036	0,025089606
14/14	2	1.472	0,007168459
15/ 7	1	0.126	0,003584229
15 /9	1	1.544	0,003584229
15/11	4	1.724	0,014336918
15/12	2	1.670	0,007168459
15/14	2	0.736	0,007168459
16/ 9	1	0.309	0,003584229
16/12	1	0.334	0,003584229
17/13	1	0.147	0,003584229
44	279		1

Figura 12. Distribución de los genotipos observados y esperados en la población Santandereana para el locus Penta D.



4.1.7 Locus D7S820

Las frecuencias alélicas presentan variación en un rango de 0.0053 a 0.2706 para los alelos 14 y 10 respectivamente.(Tabla 15). El número de alelos encontrados fue de 8 (Figura 13).

Figura 13. Distribución de las frecuencias alélicas en la población Santandereana para el locus D7S820.

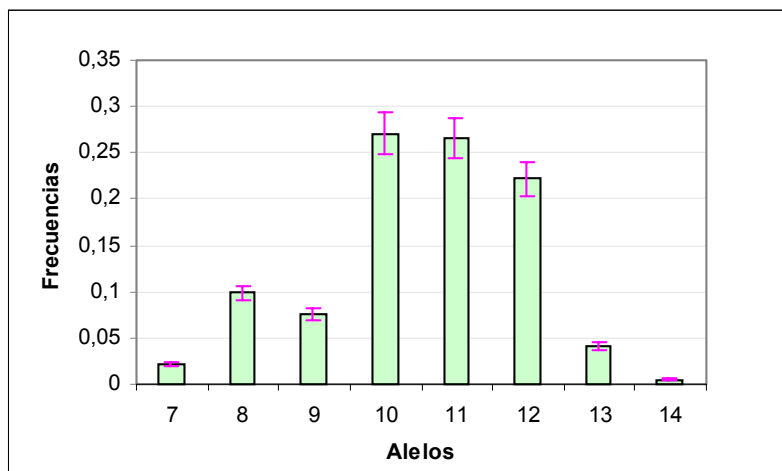


Tabla 15. Frecuencias alélicas encontradas en la población Santandereana para el locus D7S820.

Alelos	Ind	F. alélica	EE
7	12	0,021505	0,00178437
8	55	0,098566	0,00817838
9	42	0,075269	0,00624531
10	151	0,270609	0,02245338
11	148	0,265233	0,02200729
12	124	0,222222	0,01843854
13	23	0,041219	0,00342005
14	3	0,005376	0,00044609
9	558	1	

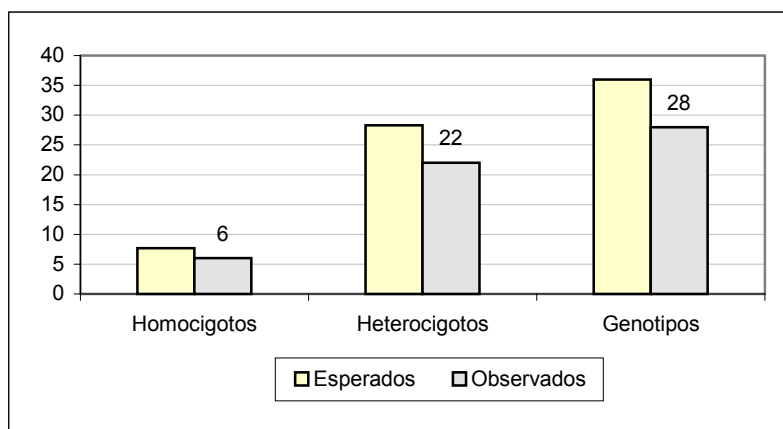
El número de genotipos presentes fue 28, de 36 esperados y el valor para la heterocigocidad observada fue de 0.7744. Las frecuencias genotípicas muestran una distribución equilibrada aunque al igual que con el marcador anterior el genotipo más común es el 11/10 con una frecuencia de 0.1362 (Tabla16). Los genotipos homocigotos y heterocigotos observados corresponden al 21.42% y 78.57% (Figura 14).

Tabla 16. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D7S820.

Genotipos	Obs	Esp	Frecuencias
8/ 7	2	1.185	0,00716846
8/ 8	3	2.666	0,01075269
9/ 8	4	4.147	0,01433692
9/ 9	3	1.546	0,01075269
10/ 7	3	3.253	0,01075269
10/ 8	14	14.910	0,05017921
10/ 9	11	11.386	0,03942652
10/10	21	20.332	0,07526882
11/ 7	4	3.189	0,01433692
11/ 8	17	14.614	0,0609319
11/ 9	10	11.160	0,03584229
11/10	38	40.122	0,13620072
11/11	20	19.530	0,07168459
12/ 7	3	2.671	0,01075269
12 / 8	9	12.244	0,03225806
12/ 9	10	9.350	0,03584229
12 /10	37	33.616	0,13261649
12/11	32	32.948	0,11469534
12 /12	14	13.691	0,05017921
13 /8	3	2.271	0,01075269
13/ 9	1	1.734	0,00358423
13/10	5	6.235	0,01792115
13/11	6	6.111	0,02150538
13/12	4	5.120	0,01433692
13/13	2	0.454	0,00716846
14/10	1	0.813	0,00358423
14/11	1	0.797	0,00358423

14/12	1	0.668	0,00358423
28	279		1

Figura 14. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D7S820.



No hay diferencia significativa entre el número de genotipos observados y esperados ($X^2=1.77$, 19 g.l., $p>0.995$). El equilibrio de Hardy-Weinberg se cumplió ($P=0.9915$).

4.1.8 Locus D13S317

Este locus presenta 8 variantes alélicas (Figura 15), con frecuencias en un rango de 0.0035 a 0.2831 (Tabla 17) siendo el alelo más común el 12 (28.31%) seguido del 11 (23.29%).

Se hallaron 29 genotipos de 36 esperados (Figura 16), el 20.68% son homocigotos y el 79.32% son heterocigotos (Tabla18). Aunque no hay grandes diferencias entre las frecuencias de los 29 genotipos

observados, el de mayor frecuencia fue 12/11 con una frecuencia de 0.1541. La heterocigidad observada fue de 0.839.

Tabla 17. Frecuencias alélicas encontradas en la población Santandereana para el locus D13S317.

Alelo	Ind	F. alélica	EE
8	59	0,105735	0,008773175
9	73	0,130824	0,010854945
10	41	0,073477	0,006096613
11	130	0,232975	0,019330724
12	158	0,283154	0,023494265
13	56	0,100358	0,008327081
14	39	0,069892	0,005799217
15	2	0,003584	0,000297396
8	558	1	

Figura 15. Distribución de las frecuencias alélicas en la población Santandereana para el locus D3S317.

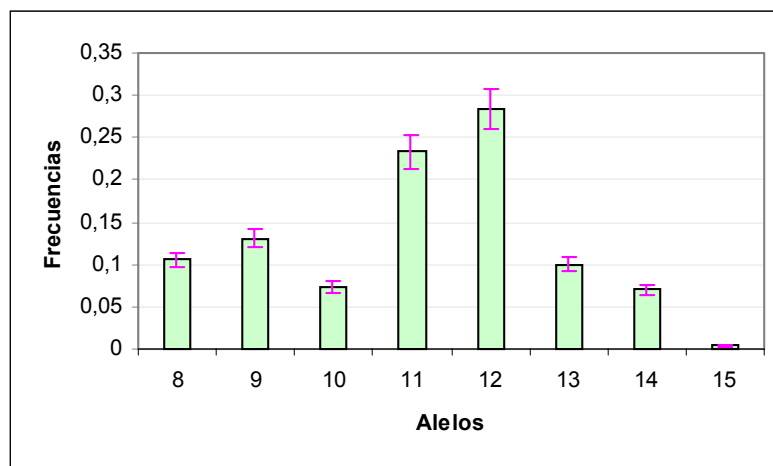
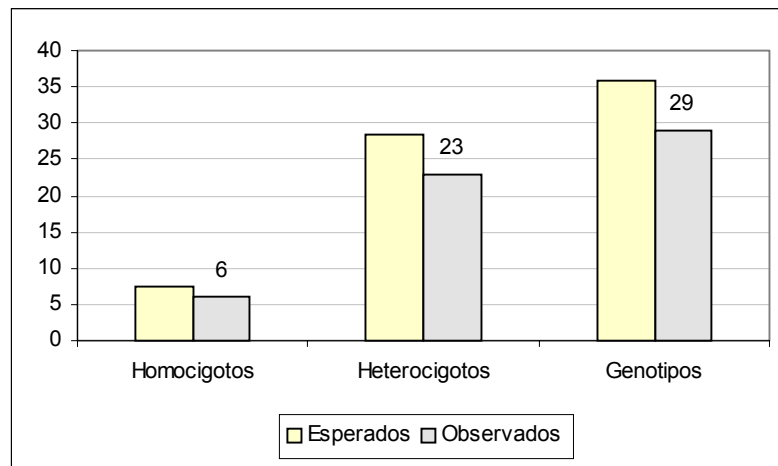


Tabla 18. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D13S317.

Genotipos	Obs	Esp	Frecuencias
8 /8	3	3.072	0,010752688
9/ 8	10	7.732	0,035842294
9 /9	2	4.718	0,007168459
10/8	6	4.343	0,021505376
10/ 9	8	5.373	0,028673835
10/10	2	1.472	0,007168459
11 /8	10	13.770	0,035842294
11/ 9	22	17.038	0,078853047
11/10	6	9.569	0,021505376
11/11	13	15.054	0,046594982
12/8	16	16.736	0,05734767
12/9	17	20.707	0,0609319
12/10	9	11.630	0,032258065
12/11	43	36.876	0,154121864
12/12	21	22.268	0,075268817
13 / 8	6	5.932	0,021505376
13/ 9	7	7.339	0,025089606
13/10	3	4.122	0,010752688
13/11	11	13.070	0,039426523
13/12	21	15.885	0,075268817
13/13	2	2.765	0,007168459
14/ 8	5	4.131	0,017921147
14/ 9	5	5.111	0,017921147
14/10	5	2.871	0,017921147
14 /11	12	9.102	0,043010753
14/12	9	11.063	0,032258065
14/13	3	3.921	0,010752688
15/12	1	0.567	0,003584229
15/13	1	0.201	0,003584229
29	279		1

Figura 16. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D3S317.



No hay diferencia significativa entre el número de genotipos observados y esperados ($\chi^2=25.04$, 30 g.l., $0.25 < p < 0.75$). El equilibrio de Hardy-Weinberg se cumplió ($P=0.6836$).

4.1.9 Locus D5S818

Se encontraron 8 alelos (Figura 17), con frecuencias alélicas que varían entre 0.3853 para el alelo 11 y 0.0017 para el alelo 8 (Tabla 19).

La población estudiada muestra 24 genotipos de 36 esperados, de los cuales 4 son homocigotos (16.66%) y 20 son heterocigotos (83.33%) (Figura 18). El genotipo que predomina es el heterocigoto 12/11 presente en 69 individuos (24.73%) (Tabla 20). La heterocigocidad observada presentó un valor de 0.7027.

Tabla 19. Frecuencias alélicas encontradas en la población Santandereana para el locus D5S818.

Alelo	Ind	F. alélica	EE
7	20	0,035842	0,002973958
8	1	0,001792	0,000148698
9	29	0,051971	0,004312238
10	35	0,062724	0,005204426
11	215	0,385305	0,031970044
12	177	0,317204	0,026319525
13	73	0,130824	0,010854945
14	8	0,014337	0,001189583
8	558	1	

Figura 17. Distribución de las frecuencias alélicas en la población Santandereana para el locus D5S818.

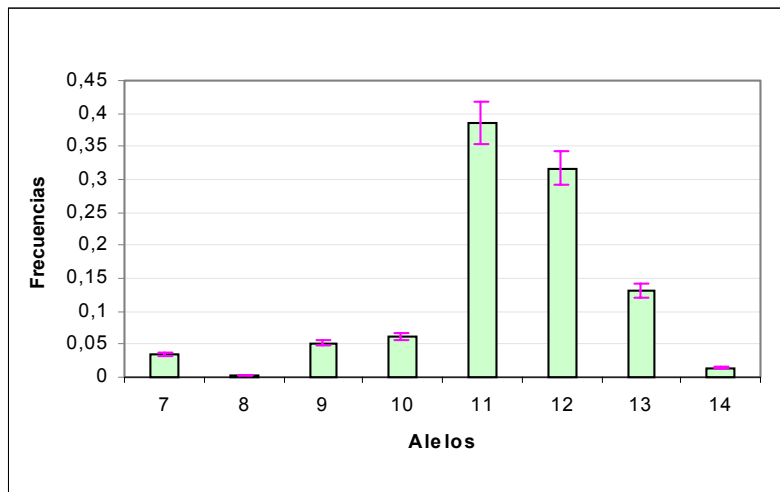
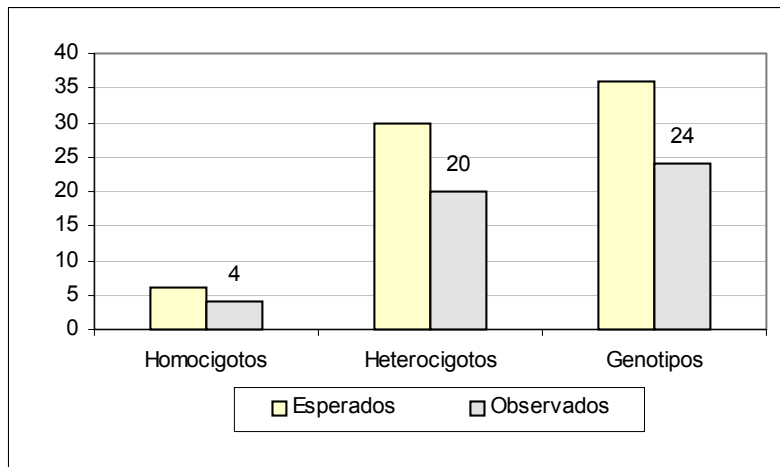


Tabla 20. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D5S818.

Genotipos	Obs	Esp	Frecuencias
9 , 7	2	1.041	,007168459
9 , 9	1	0.729	,003584229
10 , 7	1	1.257	,003584229
10 , 8	1	0.063	,003584229
11 , 7	4	7.720	,014336918
11 , 9	10	11.194	,035842294
11 ,10	13	13.510	,046594982
11 ,11	50	41.302	0,17921147
12 , 7	6	6.355	,021505376
12 , 9	10	9.215	,035842294
12 ,10	13	11.122	,046594982
12 ,11	68	68.321	,243727599
12 ,12	26	27.964	,093189964
13 , 7	7	2.621	,025089606
13 , 9	4	3.801	,014336918
13 ,10	6	4.587	,021505376
13 ,11	19	28.178	,068100358
13 ,12	24	23.197	,086021505
13 ,13	6	4.718	,021505376
14 , 9	1	0.417	,003584229
14 ,10	1	0.503	,003584229
14 ,11	1	3.088	,003584229
14 ,12	4	2.542	,014336918
14 ,13	1	1.048	,003584229
24	279		1

Los genotipos observados y esperados no presentan diferencia significativa ($X^2=4$, 16 g.l., $p > 0.995$). El locus se encuentra en equilibrio de Hardy–Weinberg ($P=0.0956$).

Figura 18. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D5S818.



4.1.10 Locus D16S539

El número de alelos detectados para este marcador es de 8 (Figura 19). Los alelos 11 y 12 son los más comunes presentando frecuencias de 0.2867 y 0.2491 respectivamente (Tabla 21), el alelo 15 fue el que menos se segregó con una frecuencia de 0.0017.

Se encontraron 25 genotipos observados de 36 esperados, con frecuencias relativamente homogéneas aunque sobresalió el genotipo 12/11 con una frecuencia de 0.1648 (tabla 22). Los genotipos homocigotos y heterocigotos presentan valores observados de 4 (16%) y 21(84%) respectivamente (figura 20). La población tiene un valor de heterocigocidad observada de 0.8319. No hay diferencia significativa entre el número de genotipos observados y esperados ($X^2=3.36$, 17 g.l., $p>0.995$). El locus se encuentra en equilibrio de Hardy–Weinberg ($P=0.6647$).

Tabla 21. Frecuencias alélicas encontradas en la población Santandereana para el locus D16S539.

Alelo	Ind	F. alélica	EE
8	8	0,01434	0,001189583
9	82	0,14695	0,012193226
10	84	0,15054	0,012490622
11	160	0,28674	0,023791661
12	139	0,2491	0,020669005
13	71	0,12724	0,010557549
14	13	0,0233	0,001933072
15	1	0,00179	0,000148698
8	558	1	

Figura 19. Distribución de las frecuencias alélicas en la población Santandereana para el locus D16S539.

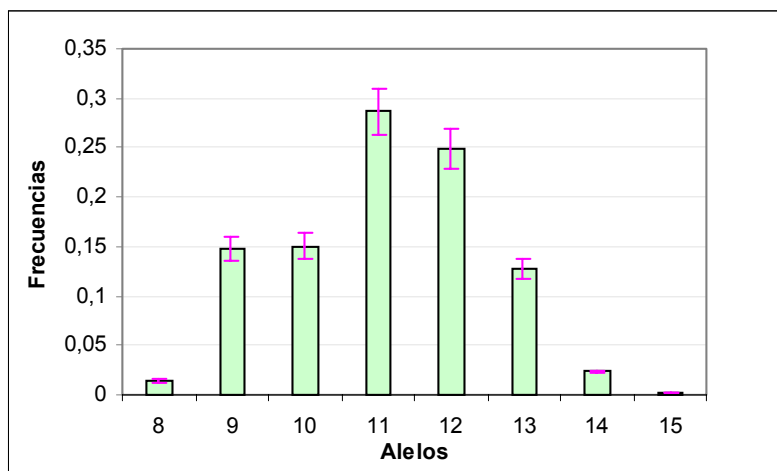
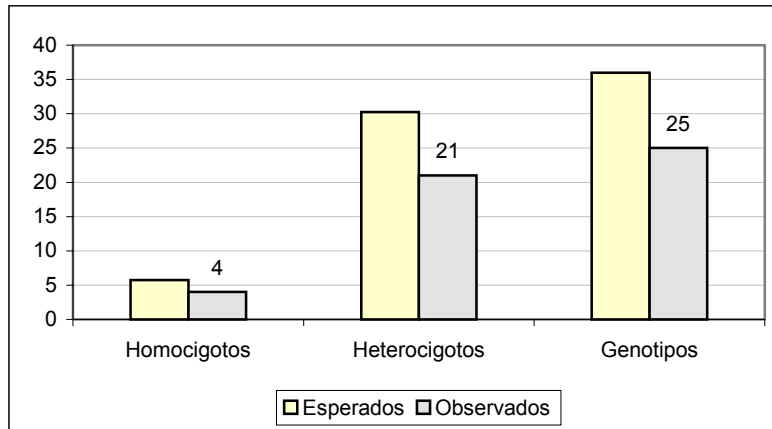


Tabla 22. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D16S539.

Genotipos	Obs	Esp	Frecuencia
9/9	7	5.962	0,025089606
10/ 8	1	1.206	0,003584229
10/ 9	14	12.366	0,050179211
10/10	8	6.259	0,028673835
11/8	5	2.298	0,017921147
11/9	21	23.555	0,075268817
11/10	21	24.129	0,075268817
11/11	21	22.837	0,075268817
12/ 8	1	1.996	0,003584229
12/ 9	18	20.463	0,064516129
12/10	23	20.962	0,082437276
12/11	46	39.928	0,164874552
12/12	14	17.219	0,050179211
13/ 9	11	10.452	0,039426523
13/10	8	10.707	0,028673835
13/11	20	20.395	0,071684588
13/12	22	17.718	0,078853047
13/13	4	4.461	0,014336918
14/ 8	1	0.187	0,003584229
14/ 9	4	1.914	0,014336918
14/10	1	1.961	0,003584229
14/11	4	3.734	0,014336918
14/12	1	3.244	0,003584229
14/13	2	1.657	0,007168459
15/11	1	0.287	0,003584229
25	279		1

Figura 20. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D16S539.



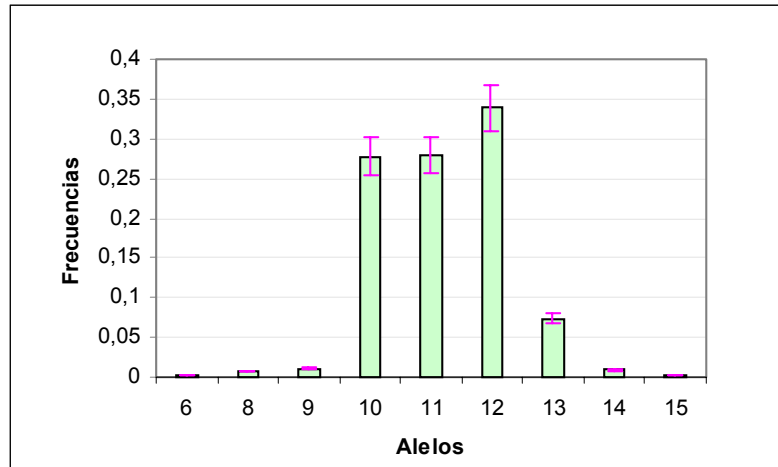
4.1.11 Locus CSF1PO

En un grupo de 9 alelos, los más frecuentes para este locus son el 12 (33.87%) y el 11 (27.95%), constituyendo más de la mitad de la población en estudio (Figura 21), los alelos 6, 8, 9 14 y 15 están muy poco segregados dentro de la población (tabla23).

Tabla 23. Frecuencias alélicas encontradas en la población Santandereana para el locus CSF1PO.

Alelo	Ind	F. alélica	EE
6	1	0,00179	0,000148698
8	4	0,00717	0,000594792
9	6	0,01075	0,000892187
10	155	0,27778	0,023048171
11	156	0,27957	0,023196869
12	189	0,33871	0,028103899
13	41	0,07348	0,006096613
14	5	0,00896	0,000743489
15	1	0,00179	0,000148698
9	558	1	

Figura 21. Distribución de las frecuencias alélicas en la población Santandereana para el locus CSF1PO.



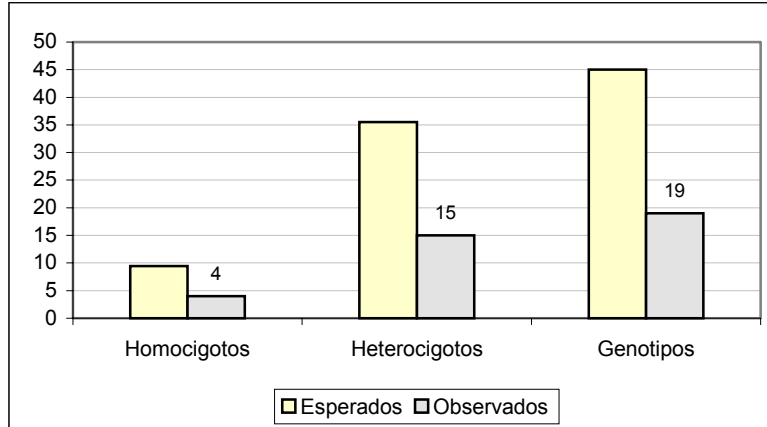
Se encontraron 19 genotipos de los 45 que deberían haberse formado en los individuos de esta población (Figura 22). Las frecuencias genotípicas varían desde 0.0035 a 0.2043 (tabla 24). Este marcador en la población estudiada presenta un valor de heterocigocidad observada de 0.7069. No hay diferencia significativa entre el número de genotipos observados y esperados ($X^2=15.02$, 10 g.l., $0.10 < p < 0.25$). El locus se encuentra en equilibrio de Hardy–Weinberg ($P=0.077$).

Tabla 24. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus CSF1PO.

Genotipos	Obs	Esp	Frecuencia
10/6	1	0.278	0,003584229
10 /9	2	1.670	0,007168459
10/10	20	21.427	0,071684588
11/8	2	1.120	0,007168459
11/ 9	3	1.680	0,010752688
11/10	44	43.411	0,157706093

11 /11	24	21.706	0,086021505
12 / 8	2	1.357	0,007168459
12 /10	57	52.594	0,204301075
12 /11	41	52.934	0,146953405
12 /12	37	31.896	0,132616487
13 /10	8	11.409	0,028673835
13 /11	17	11.483	0,0609319
13 /12	14	13.912	0,050179211
13 /13	1	1.472	0,003584229
14 /9	1	0.054	0,003584229
14/10	3	1.391	0,010752688
14/12	1	1.697	0,003584229
15/11	1	0.280	0,003584229
19	279		1

Figura 22. Distribución de los genotipos observados y esperados en la población Santandereana para el locus CSF1PO.



4.1.12 Locus D21S11

Este es uno de los marcadores más polimórficos, para este sistema se encontraron 15 alelos (Figura 23), siendo los más comunes el 30 con una frecuencia de 0.2831 y el 29 con 0.2204 (Tabla 25).

Tabla 25. Frecuencias alélicas encontradas en la población Santandereana para el locus D21S11.

Alelo	Ind	F. alélica	EE
27	15	0,0269	0,002230468
28	64	0,1147	0,009516664
29	123	0,2204	0,018289839
29,2	1	0,0018	0,000148698
30	158	0,2832	0,023494265
30,2	16	0,0287	0,002379166
31	29	0,052	0,004312238
31,2	51	0,0914	0,007583592
32	11	0,0197	0,001635677
32,2	63	0,1129	0,009367966
33	3	0,0054	0,000446094
33,2	19	0,0341	0,00282526
34,2	3	0,0054	0,000446094
35	1	0,0018	0,000148698
37	1	0,0018	0,000148698
15	558	1	

Este locus presentó un número de genotipos observados de 52 de 105 esperados. El porcentaje de homocigotos y heterocigotos fué de 13.46% y 86.53% respectivamente (Figura 24). El genotipo 30/29 presente en 41 individuos es el de mayor frecuencia (Tabla 26). La Heterocigocidad observada para este marcador en la población fue de 0.8358.

Figura 23. Distribución de las frecuencias alélicas en la población Santandereana para el locus D21S11.

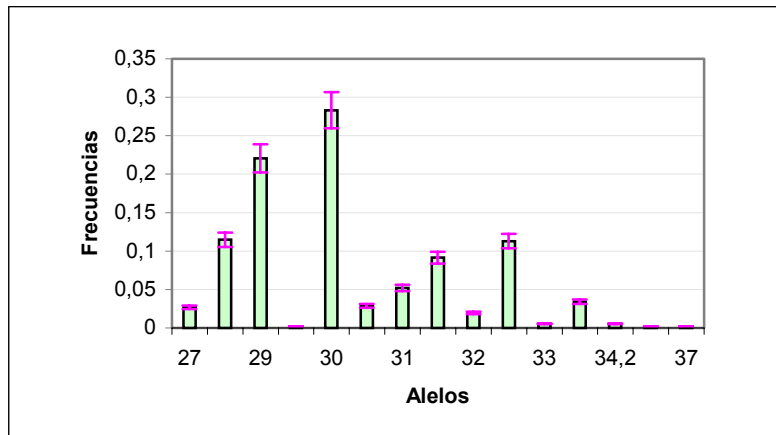


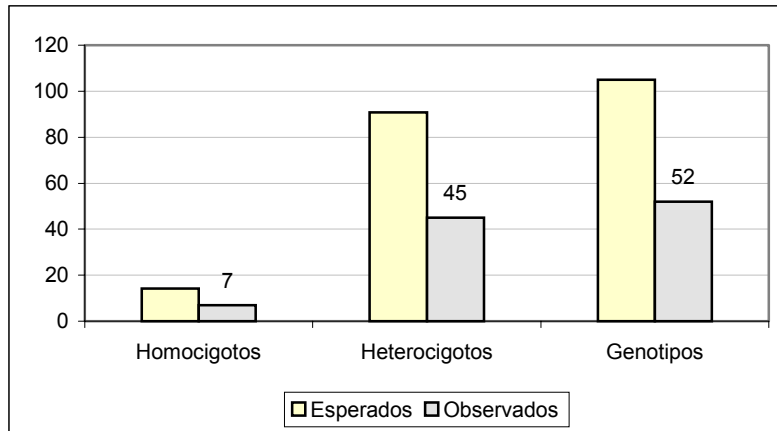
Tabla 26. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D21S11.

Genotipos	Obs	Esp	Frecuencia
27 /27	1	0.189	0,003584229
28 /27	2	1.724	0,007168459
28 /28	4	3.619	0,014336918
29 /25	2	3.312	0,007168459
29 /28	18	14.133	0,064516129
29 /29	10	13.470	0,035842294
30 /31	2	4.255	0,007168459
30 /28	16	18.154	0,05734767
30 /29	41	34.890	0,146953405
30 /30	23	22.268	0,082437276
30,2 /27	1	0.431	0,003584229
30,2 /28	2	1.838	0,007168459
30,2 /29	2	3.533	0,007168459
30,2 /30	3	4.539	0,010752688
31 /27	3	0.781	0,010752688
31 /28	1	3.332	0,003584229
31 /29	4	6.404	0,014336918
31 /30	5	8.226	0,017921147
31 /30,2	1	0.833	0,003584229
31 /31	1	0.729	0,003584229

31,2 /27	1	1.373	0,003584229
31,2 /28	8	5.860	0,028673835
31,2 /29	13	11.262	0,046594982
31,2 /30	14	14.467	0,050179211
31,2 /30,2	3	1.465	0,010752688
31,2 /31	3	2.655	0,010752688
31,2 /31,2	1	2.289	0,003584229
32 /28	3	1.379	0,010752688
32 /29	5	2.650	0,017921147
32 /30	4	3.404	0,014336918
32,2 /27	2	1.697	0,007168459
32,2 /28	3	7.239	0,010752688
32,2 /29	12	13.912	0,043010753
32,2 /30	18	17.871	0,064516129
32,2 /30,2	2	1.810	0,007168459
32,2 /31	8	3.280	0,028673835
32,2 /31,2	4	5.768	0,014336918
32,2 /32,2	6	3.506	0,021505376
33 /30	1	0.851	0,003584229
33 /31,2	2	0.275	0,007168459
33,2 /28	3	2.183	0,010752688
33,2 /29	4	4.196	0,014336918
33,2 /30	6	5.390	0,021505376
33,2 /30,2	1	0.546	0,003584229
33,2 /31	2	0.989	0,007168459
33,2 /31,2	1	1.740	0,003584229
33,2 /32,2	2	2.149	0,007168459
34,2 /29	1	0.662	0,003584229
34,2 /30	1	0.851	0,003584229
34,2 /30,2	1	0.086	0,003584229
35 /29	1	0.221	0,003584229
37 /30	1	0.284	0,003584229
52	279		1

Los genotipos observados y esperados no presentan diferencia significativa ($X^2=26.75$, 137g.l., $0.995 < p > 0.999$). No hubo desviaciones del equilibrio de Hardy-Weinberg ($P=0.4487$).

Figura 24. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D21S11.



4.1.13 Locus FGA

Para este sistema se encontraron 14 alelos (Figura 25), siendo los más segregados el 24 y 25 con frecuencias de 0.1774 para cada uno de ellos y el que menos se detectó fue el 23,2 con una frecuencia de tan solo 0.0017 (Tabla 27).

Este locus presentó un número de genotipos observados de 52 de 105 esperados. El porcentaje de homocigotos y heterocigotos fué de 15.38% y 75.86% respectivamente (Figura 26). La mayoría de los genotipos están presentes en proporciones muy bajas, sin embargo, sobresale el genotipo 25/22 y el 25/24 con frecuencias de 0.0716 y 0.06810 (Tabla 28). La Heterocigidad observada para este marcador en la población muestreada fue de 0.8996. Los genotipos observados y esperados no presentan diferencia significativa ($X^2=26.75$, 38 g.l., $0.90 < p < 0.95$). No hubo desviaciones del equilibrio de Hardy-Weinberg ($P=0.8196$).

Tabla 27. Frecuencias alélicas encontradas en la población Santandereana para el locus FGA.

Alelo	Ind.	Frecuencia	EE
18	8	0,014337	0,001189583
19	38	0,0681	0,005650519
20	45	0,080645	0,006691405
20,2	2	0,003584	0,000297396
21	52	0,09319	0,00773229
22	84	0,150538	0,012490622
23	57	0,102151	0,008475779
23,2	1	0,001792	0,000148698
24	99	0,177419	0,01472109
25	99	0,177419	0,01472109
26	59	0,105735	0,008773175
27	9	0,016129	0,001338281
28	3	0,005376	0,000446094
29	2	0,003584	0,000297396
14	558	1	

Figura 25. Distribución de las frecuencias alélicas en la población Santandereana para el locus FGA.

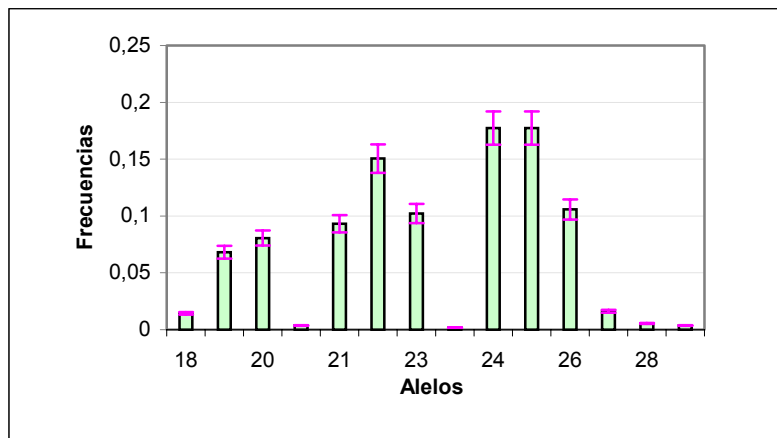
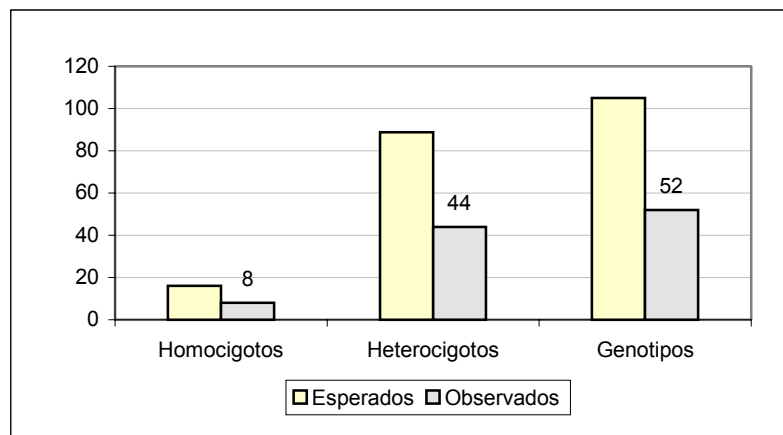


Tabla 28. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus FGA.

Genotipos	Obs.	Esp.	Frecuencia
19 /19	2	1.262	0,007168459
20 /19	7	3.070	0,025089606
20 /20	1	1.777	0,003584229
22 /19	3	3.548	0,010752688
21 /20	6	4.201	0,021505376
21 /21	2	2.381	0,007168459
22 /18	1	1.206	0,003584229
22 /19	5	5.731	0,017921147
22 /20	4	6.786	0,014336918
22 /21	6	7.842	0,021505376
22 /22	9	6.259	0,032258065
23 /19	3	3.889	0,010752688
23 /20	7	4.605	0,025089606
23 /21	5	5.321	0,017921147
23 /22	6	8.596	0,021505376
23 /23	2	2.865	0,007168459
24 /18	2	1.422	0,007168459
24 /19	5	6.754	0,017921147
24 /20	8	7.998	0,028673835
24 /20,2	1	0.355	0,003584229
24 /21	12	9.242	0,043010753
24 /22	13	14.930	0,046594982
24 /23	10	10.131	0,035842294
24 /23,2	1	0.178	0,003584229
24 /24	6	8.709	0,021505376
25 /18	3	1.422	0,010752688
25 /19	7	6.754	0,025089606
25 /20	7	7.998	0,025089606
25 /21	7	9.242	0,025089606
25 /22	20	14.930	0,071684588
25 /23	14	10.131	0,050179211
25 /24	19	17.596	0,068100358
25 /25	4	8.709	0,014336918
26 /18	2	0.847	0,007168459
26 /19	2	4.025	0,007168459
26 /20	4	4.767	0,014336918
26 /20,2	1	0.212	0,003584229
26 /21	7	5.508	0,025089606
26 /22	7	8.898	0,025089606

26 /23	8	6.038	0,028673835
26 /24	12	10.487	0,043010753
26 /25	12	10.487	0,043010753
26 /26	2	3.072	0,007168459
27 /19	1	0.614	0,003584229
27 /22	3	1.357	0,010752688
27 /24	4	1.600	0,014336918
27 /25	1	1.600	0,003584229
28 /21	1	0.280	0,003584229
28 /22	1	0.452	0,003584229
28 /25	1	0.533	0,003584229
39 /19	1	0.136	0,003584229
39 /21	1	0.187	0,003584229
52	279		1

Figura 26. Distribución de los genotipos observados y esperados en la población Santandereana para el locus FGA.



4.1.14 Locus D8S1179

Se encontraron 11 alelos (Figura 27) para una población con frecuencias alélicas que varían entre 0.3207 para el alelo 13 a 0.0017 para el alelo 19 (Tabla 29).

Figura 27. Distribución de las frecuencias alélicas en la población Santandereana para el locus D8S1179.

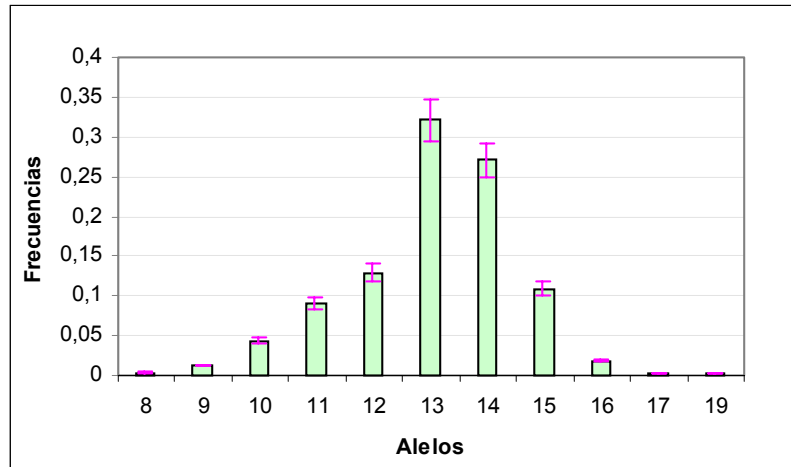


Tabla 29. Frecuencias alélicas encontradas en la población Santandereana para el locus D8S1179.

Alelo	Ind	F. alélica	EE
8	2	0,0036	0,001948409
9	7	0,0125	0,001040885
10	24	0,043	0,003568749
11	50	0,0896	0,007434894
12	72	0,129	0,010706247
13	179	0,3208	0,02661692
14	151	0,2706	0,02245338
15	61	0,1093	0,009070571
16	10	0,0179	0,001486979
17	1	0,0018	0,000148698
19	1	0,0018	0,000148698
11	558	1	

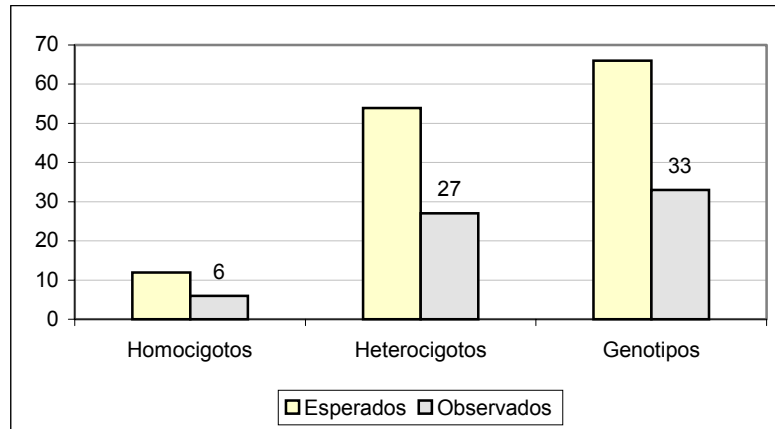
La población estudiada muestra 33 genotipos de 66 esperados, de los cuales 6 son homocigotos (18.18%) y 27 son heterocigotos (81.81%) (Figura 28). El genotipo que predomina es el 14/13 presente en 43 individuos (15.41%) (Tabla 30). La heterocigocidad observada presentó un valor de 0.7635. Los genotipos observados y esperados no

presentan diferencia significativa ($X^2=16.5$, 22 g.l., $0.75 < p < 0.90$). El locus se encuentra en equilibrio de Hardy-Weinberg ($P=0.7871$).

Tabla 30. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus D8S1179.

Genotipos	Obs.	Esp.	Frecuencia
10/10	1	0.496	0,003584229
11 /8	2	0.180	0,007168459
11/ 9	1	0.628	0,003584229
11/10	2	2.154	0,007168459
11/11	3	2.199	0,010752688
12/ 9	1	0.905	0,003584229
12/10	5	3.102	0,017921147
12 /11	4	6.463	0,014336918
12 /12	6	4.589	0,021505376
13 /9	3	2.250	0,010752688
13/10	7	7.713	0,025089606
13/11	18	16.068	0,064516129
13/12	22	23.138	0,078853047
13/13	31	28.601	0,111111111
14 /9	2	1.898	0,007168459
14 /10	6	6.506	0,021505376
14 /11	12	13.555	0,043010753
14 /12	20	19.519	0,071684588
14/13	43	48.526	0,154121864
14/14	23	20.332	0,082437276
15/10	1	2.628	0,003584229
15/11	5	5.476	0,017921147
15/12	7	7.885	0,025089606
15/13	21	19.603	0,075268817
15/14	18	16.537	0,064516129
15/15	3	3.285	0,010752688
16/12	1	1.293	0,003584229
16/13	3	3.214	0,010752688
16/14	3	2.711	0,010752688
16/15	3	1.095	0,010752688
17/10	1	0.043	0,003584229
19/14	1	0.271	0,003584229
33	279		1

Figura 28. Distribución de los genotipos observados y esperados en la población Santandereana para el locus D8S1179.



4.1.15 Locus Penta E

Fue el sistema estudiado con mayor número de alelos encontrados (Figura 29), el alelo 12 sobresale con una frecuencia de 0.1881 seguido del 14 y 15 con una frecuencia de 0.1146 para cada uno de ellos (Tabla 31).

Este marcador fue el que presento mayor número de genotipos, encontrándose un total de 101 observados contra 171 esperados (Figura 30); sus frecuencias son relativamente constantes como se indica en la Tabla 32. Los genotipos homocigotos y heterocigotos presentan valores observados de 9 y 92 respectivamente y la población tiene un valor de heterocigocidad observada de 0.8887. No hay diferencia significativa entre el número de genotipos observados y esperados ($\chi^2=28.65$, 83 g.l., $0.99 < p < 0.995$). El locus se encuentra en equilibrio de Hardy–Weinberg ($P=0.9931$).

Figura 29. Distribución de las frecuencias alélicas en la población Santandereana para el locus Penta E.

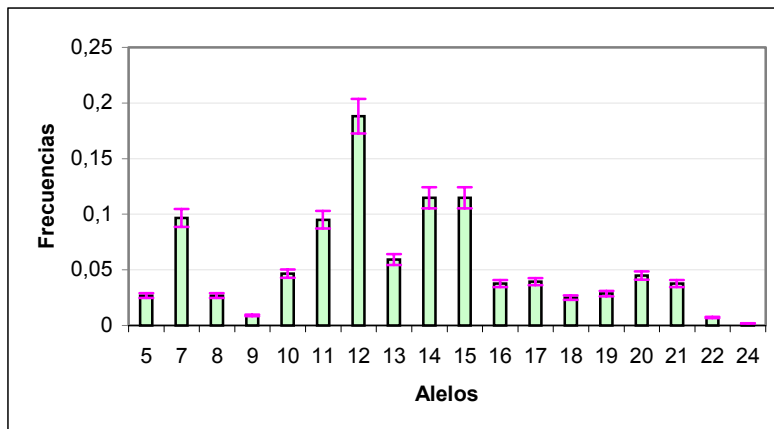


Tabla 31. Frecuencias alélicas encontradas en la población Santandereana para el locus Penta E.

Alelo	Ind	Frecuencia	EE
5	15	0,026882	0,002230468
7	54	0,096774	0,008029685
8	15	0,026882	0,002230468
9	5	0,008961	0,000743489
10	26	0,046595	0,003866145
11	53	0,094982	0,007880988
12	105	0,188172	0,015613277
13	33	0,05914	0,00490703
14	64	0,114695	0,009516664
15	64	0,114695	0,009516664
16	21	0,037634	0,003122655
17	22	0,039427	0,003271353
18	14	0,02509	0,00208177
19	16	0,028674	0,002379166
20	25	0,044803	0,003717447
21	21	0,037634	0,003122655
22	4	0,007168	0,000594792
24	1	0,001792	0,000148698
18	558	1	

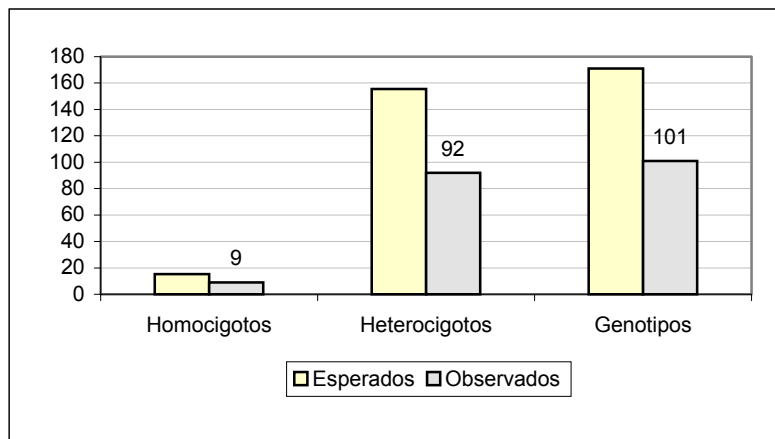
Tabla 32. Frecuencias genotípicas y valores observados y esperados en la población Santandereana para el locus Penta E.

Genotipos	Obs	Esp	Frecuencia
7/5	3	1.454	0,010752688
7/7	1	2.569	0,003584229
8/5	1	0.404	0,003584229
8/7	2	1.454	0,007168459
10/7	5	2.521	0,017921147
10/8	1	0.700	0,003584229
11/5	3	1.427	0,010752688
11/7	5	5.138	0,017921147
11/8	3	1.427	0,010752688
11/11	4	2.474	0,014336918
12/5	1	2.828	0,003584229
12/7	11	10.180	0,039426523
12/8	4	2.828	0,014336918
12/9	2	0.943	0,007168459
12/10	4	4.901	0,014336918
12/11	10	9.991	0,035842294
12/12	12	9.803	0,043010753
13/5	1	0.889	0,003584229
13/7	5	3.199	0,017921147
13/10	1	1.540	0,003584229
13/11	3	3.140	0,010752688
13/12	6	6.221	0,021505376
14/5	1	1.724	0,003584229
14/7	4	6.205	0,014336918
14/9	1	0.575	0,003584229
14/10	4	2.987	0,014336918
14/11	6	6.090	0,021505376
14/12	14	12.065	0,050179211
14/13	4	3.792	0,014336918
14/14	4	3.619	0,014336918
15/5	1	1.724	0,003584229
15/7	7	6.205	0,025089606
15/8	2	1.724	0,007168459
15/9	1	0.575	0,003584229
15/10	3	2.987	0,010752688
15/11	5	6.090	0,017921147
15/12	9	12.065	0,032258065
15/13	4	3.792	0,014336918
15/14	7	7.354	0,025089606
15/15	6	3.619	0,021505376
16/7	2	2.036	0,007168459
16/10	2	0.980	0,007168459
16/12	2	3.959	0,007168459
16/13	2	1.244	0,007168459
16/14	4	2.413	0,014336918
16/15	3	2.413	0,010752688

16/16	1	0.377	0,003584229
17 / 5	1	0.592	0,003584229
17 / 7	3	2.133	0,010752688
17 / 8	1	0.592	0,003584229
17 / 9	1	0.197	0,003584229
17/11	2	2.093	0,007168459
17/12	3	4.147	0,010752688
17 /13	3	1.303	0,010752688
17 /14	1	2.528	0,003584229
17 /15	2	2.528	0,007168459
17 /17	1	0.415	0,003584229
18 / 7	2	1.357	0,007168459
18 /10	1	0.654	0,003584229
18 /11	1	1.332	0,003584229
18 /12	2	2.639	0,007168459
18/14	1	1.609	0,003584229
18/15	1	1.609	0,003584229
18 /17	2	0.553	0,007168459
18/18	1	0.163	0,003584229
19 /5	1	0.431	0,003584229
19 /7	1	1.551	0,003584229
19 /10	1	0.747	0,003584229
19/11	1	1.522	0,003584229
19/12	4	3.016	0,014336918
19 /13	2	0.948	0,007168459
19 /14	4	1.838	0,014336918
19 /16	1	0.603	0,003584229
20 /8	1	2.424	0,003584229
20/10	2	1.167	0,007168459
20/11	4	2.379	0,014336918
20 /12	4	4.713	0,014336918
20 /13	1	1.481	0,003584229
20 /14	3	2.873	0,010752688
20 /15	5	2.873	0,017921147
20 /16	1	0.943	0,003584229
20 /18	1	0.628	0,003584229
20 /20	1	0.539	0,003584229
21 /5	2	0.566	0,007168459
21/ 7	1	2.036	0,003584229
21 /8	1	0.566	0,003584229
21/10	2	0.980	0,007168459
21/11	2	1.998	0,007168459
21/12	3	3.959	0,010752688
21 /14	2	2.413	0,007168459
21 /15	1	2.413	0,003584229
21 /16	2	0.792	0,007168459
21 /17	1	0.829	0,003584229
21 /18	1	0.528	0,003584229
21 /19	1	0.603	0,003584229
21 /20	1	0.943	0,003584229
22/12	1	0.754	0,003584229
22 /13	1	0.237	0,003584229

22 /15	1	0.460	0,003584229
22 /21	1	0.151	0,003584229
24/12	1	0.189	0,003584229
101	279		1

Figura 30. Distribución de los genotipos observados y esperados en la población Santandereana para el locus Penta E.



Como se resume en la tabla 33 se encontraron para los 15 marcadores estudiados un total de 42 alelos, siendo el marcador Penta E y FGA los que más aporte de alelos presentaron. Las frecuencias fueron muy variables y están en un rango de 0.0018 a 0.4838.

Tabla 33. Frecuencias alélicas para los 15 STR's estudiados en la población Santandereana.

Alelo	FGA n=279	TPOX n=278	VWA n=279	D18S51 n=279	TH01 n=279	D3S1358 n=279	Penta D n=279	D7S820 n=279
2,2							0,0125	
5							0,0018	
6		0,0072			0,3925			
6,3					0,0018			
7					0,1989		0,0125	0,0215
8		0,4838			0,0663		0,0179	0,0986
9		0,0522			0,1380		0,1541	0,0753
9,3					0,2007			
10		0,0432		0,0108	0,0018		0,2186	0,2706
11		0,3345		0,0090			0,1720	0,2652
12		0,0791		0,1093			0,1667	0,2222
13			0,0036	0,1219		0,0054	0,1470	0,0412
14			0,0699	0,1487		0,0699	0,0735	0,0054
15			0,1057	0,1416		0,4032	0,0179	
16			0,3674	0,1254		0,2724	0,0036	
17			0,2599	0,1667		0,1505	0,0018	
18	0,0143		0,1505	0,0771		0,0950		
19	0,0681		0,0341	0,0341		0,0036		
20	0,0806		0,0090	0,0251				
20,2	0,0036							
21	0,0932			0,0215				
22	0,1505			0,0036				
23	0,1022			0,0036				
23,2	0,0018							
24	0,1774			0,0018				
25	0,1774							
26	0,1057							
27	0,0161							
28	0,0054							
29	0,0036							
29,2								
30								
30,2								
31								
31,2								
32								
32,2								
33								
33,2								

Continuación

Alelo	D13S317 n=279	D5S818 n=279	D16S539 n=279	CSF1PO n=279	D21S11 n=279	D8S1179 n=279	Penta E n=279
2,2							
5							0,0269
6				0,0018			
6,3							
7		0,0358					0,0968
8	0,1057	0,0018	0,0143	0,0072		0,0036	0,0269
9	0,1308	0,0520	0,1470	0,0108		0,0125	0,0090
9,3							
10	0,0735	0,0627	0,1505	0,2778		0,0430	0,0466
11	0,2330	0,3853	0,2867	0,2796		0,0896	0,0950
12	0,2832	0,3172	0,2491	0,3387		0,1290	0,1882
13	0,1004	0,1308	0,1272	0,0735		0,3208	0,0591
14	0,0699	0,0143	0,0233	0,0090		0,2706	0,1147
15	0,0036		0,0018	0,0018		0,1093	0,1147
16						0,0179	0,0376
17						0,0018	0,0394
18							0,0251
19						0,0018	0,0287
20							0,0448
20,2							
21							0,0376
22							0,0072
23,2							
24							0,0018
25							
26							
27					0,0269		
28					0,1147		
29					0,2204		
29,2					0,0018		
30					0,2832		
30,2					0,0287		
31					0,0520		
31,2					0,0914		
32					0,0197		
32,2					0,1129		
33					0,0054		
33,2					0,0341		
34,2					0,0054		
35					0,0018		
37					0,0018		

En la Tabla 2 se encuentran los valores de heterocigocidad observada y esperada para cada uno de los 15 marcadores analizados en la población santandereana. Todos los loci tienen un alto grado de heterocigocidad que varía entre 0.9 y 0.65% para FGA y TPOX respectivamente. En tres marcadores, D5S818, CSF1PO y TPOX se observa un decremento de heterocigocidad lo que se ve reflejado en un aumento de homocigocidad. La frecuencia observada de heterocigotos muestra que los mayores valores corresponden a su vez con los más polimórficos, principalmente FGA y Penta E, subrayados con negrilla.

4.2 DESEQUILIBRIO DE LIGAMIENTO

De los 15 marcadores utilizados en el estudio solo los loci D5S818 y el CSF1PO se encuentran localizados en el cromosoma 5 (Tabla1) y no se presenta desequilibrio de ligamiento entre estos sistemas ($p > 0.268$). De las 105 combinaciones posibles entre los 15 marcadores estudiados, 4 (3.8%) presentan un comportamiento de desequilibrio de asociación, estos sistemas son: D18S51 y D21S11, TPOX y D3S1358, Penta E y D16S539 y Penta D y D13S317 (tabla 34).

4.3 COMPARACIÓN ENTRE POBLACIONES

Los 15 marcadores de la población santandereana fueron comparados con 21 poblaciones a nivel mundial, 5 Colombianas (Yunis *et al*, 2000; Benítez-Paéz *et al*, 2003; Rey *et al*, 2003 y Gómez *et al*, 2003), 4 Centro y Suramericanas (Luna-vázquez *et al*, 2003; Bydlowski *et al*, 2003; Andrade *et al*, 2003 y Chiurillo *et al*, 2003), 4 Europeas (Gehrig *et al*, 1999; Biondo *et al*, 2001; Anorin *et al*, 2001 y Paredes *et al*, 2003) y 8 Asiáticas (Yamamoto *et al*, 1999; Ahmed *et al*, 2001; Chantratita *et al*, 2001; Zhang *et al*, 2003; Seah *et al*, 2003 y Cho *et al*, 2003).

Tabla 34. Desequilibrio de ligamiento presente en la población Santandereana. La negrilla indica desequilibrio de ligamiento.

	FGA	TPOX	D8S1179	vWA	Penta E	D18S51	D21S11	TH01	D3S1358	Penta D	CSF1PO	D16S539	D7S820	D13S317	D5S818
FGA	**	0,333	0,515	0,98	1	0,069	0,932	0,918	0,208	0,517	0,955	0,111	0,902	0,517	0,57
TPOX		**	0,228	0,768	0,35	0,244	0,809	0,119	0,011	0,43	0,39	0,303	0,055	0,37	0,28
D8S1179			**	0,748	0,824	0,928	0,963	0,653	0,079	0,399	0,768	0,475	0,769	0,315	0,962
Vwa				**	0,577	0,94	0,428	0,578	0,557	0,394	0,933	0,549	0,287	0,321	0,211
Penta E					**	1	0,475	0,156	0,13	0,218	0,09	0,049	0,901	0,697	0,379
D18S51						**	0,036	0,888	0,009	0,432	0,091	0,447	0,512	0,324	0,443
D21S11							**	0,436	0,107	0,317	0,548	0,762	0,359	0,644	0,963
TH01								**	0,775	0,413	0,924	0,336	0,866	0,521	0,111
D3S1358									**	0,383	0,507	0,591	0,234	0,756	0,443
Penta D										**	0,574	0,081	0,309	0,046	0,334
CSF1PO											**	0,899	0,718	0,269	0,268
D16S539												**	0,063	0,017	0,41
D7S820													**	0,233	0,784
D13S317														**	0,064
D5S818															**

Se encontró que nueve alelos presentes en las poblaciones Colombianas nos distinguen del resto de poblaciones reportadas por la literatura comparadas en el presente estudio, 4 de ellos presentes únicamente en la población santandereana (Alelo 16 del locus Penta D, Alelo 6 del locus CSF1PO, Alelo 37 del locus D21S11 y Alelo 24 del locus Penta E) y 6 que se comparten con las demás poblaciones colombianas (Alelo 15 del locus TPOX, Alelo 4 del locus TH01, el alelo 16 del locus Penta D, el alelo 6 del locus CSF1PO, Alelo37 del locus D21S11 y Alelo 24 del locus Penta E); mientras que 11 diferencian a las poblaciones asiáticas de las Europeas y Americanas (Tabla 35).

Todos los alelos únicos que forman parte de las poblaciones comparadas presentan frecuencias en el rango de 0.001 a 0.065. En total se encontraron 223 alelos para las 22 poblaciones, de los cuales el 69.5% (153 alelos) están presentes en la población Santandereana.

4.3.1 Locus TPOX

Se encontró que para el marcador TPOX el alelo 15 hace parte solo de la población colombiana (Benítez-Paéz *et al*, 2003), mientras que el alelo 7 es compartido entre las poblaciones colombianas (Yunis *et al*, 2000; Benítez-Paéz *et al*, 2003; y Gómez *et al*, 2003), venezolana (Chiurrillo *et al*, 2003) y brasileña (Bydlowski *et al*, 2003). El alelo 14 es exclusivo del continente Asiático de la población japonesa (Yamamoto *et al*, 1999). Se encontraron 6 alelos de los 10 reportados por la literatura (tabla 35).

Los resultados del chi-cuadrado global demuestran que la población santandereana presenta diferencias estadísticamente diferentes con las poblaciones de Brasil (Bydlowski *et al*, 2003), Italia (Biondo *et al*, 2001)

y todas las poblaciones Asiáticas (Japón, Yamamoto, *et al*, 1999; Egipto, Ahmed, *et al*, 2001; China e India, Seah, *et al*, 2003; y Corea, Cho, *et al*. 2003), excluyendo a Tailandia (Chantratita, *et al*, 2001) (Tabla 36).

El alelo que presenta mayor frecuencia en nuestra población de estudio es el 8, el cual muestra diferencias estadísticamente significativas con las poblaciones de Brasil ($p = 0.0084$) e India ($p = 0.0006$).

4.3.2 Locus vWA

El alelo vWA en la población Santandereana estudiada presenta 8 alelos de los 14 reportados por la literatura. El alelo 10 es el único presente en la población ecuatoriana (Andrade *et al*, 2003) y el 11 de Sudamérica (Gómez *et al*, 2003; Bydlowski *et al*, 2003; Andrade *et al*, 2003 y Chiurrillo *et al*, 2003). Aunque en frecuencias muy bajas el alelo 12 está presente en poblaciones Suramericanas (Gómez *et al*, 2003; Andrade *et al*, 2003) y asiáticas (Yamamoto, *et al*, 2003) pero ausente en las 4 europeas (Tabla 35).

Las poblaciones de Brasil (Bydlowski *et al*, 2003), Ecuador (Andrade *et al*, 2003), Suiza (Gehrig *et al*, 1999), Italia (Biondo *et al*, 2001) y las 7 Asiáticas (Japón, Yamamoto *et al*, 1999; Egipto, Ahmed *et al*, 2001; Tailandia, Chantratita *et al*, 2001; China y India, Seah *et al*, 2003; China, Zhang *et al*, 2003 y Corea, Cho *et al*, 2003), excluyendo a Malasia (Seah *et al*, 2003), presentan diferencias estadísticamente diferentes con respecto a la población santandereana (Tabla 36).

Tabla 36. Resultados de la prueba de Chi-cuadrado Global para las 21 poblaciones comparadas teniendo como referencia la población Santandereana. La negrilla indica las diferencias estadísticamente significativas.

	FGA	TPOX	D8S1179	vWA	PentaE	D18S51	D21S11	TH01	D3S1358	PentaD	CSF1PO	D16S539	D7S820	D13S317	D5S818
Santander	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Colombia	**	0,5709	0,7742	0,4	**	**	**	0,756	**	**	0,5242	0,9676	0,8347	0,6469	0,5686
Bogotá	0,83	0,5258	0,747	**	**	0,8141	0,4873	0,621	0,7693	**	0,817	0,9054	0,9094	0,355	0,398
Boyacá	0,869	0,1759	0,9148	**	**	0,6414	0,0034	0,453	0,7307	**	0,7227	0,6461	0,9808	0,8254	0,6141
V. Del Cauca	**	0,2007	0,8523	0,57	**	0,5594	0,5921	0,153	0,1747	**	0,0529	**	0,0596	0,3846	0,7835
Venezuela	**	0,1771	**	0,42	**	**	**	0,021	0,4071	**	0,1356	0,9483	0,6253	0,4386	0,8128
Brazil	**	0	**	0	**	**	**	0	**	**	0,0002	0,0809	0,0778	0,0339	0,1697
Mexico	0,448	**	0,7166	0,2	**	0	0	**	0,2023	**	**	**	0,3574	0,0132	0,1567
Ecuador	0	0,0501	0,0523	0	0,0106	**	0,4413	0,052	0,2074	0,3762	0,9194	0,2617	0,0976	0,3256	0,0001
Suiza	0,279	0,1063	0,1079	0,02	**	0,4865	0,7561	0,003	0,0007	**	0,9051	0,1564	0,006	0,0534	0,3273
España	**	0,1457	0,0444	0,07	0,377	0,5207	0,6786	0	0,0019	0,2048	0,523	0,0213	0,0021	0,0207	0
N. Portugal	0,076	**	0,1561	0,02	**	0,5722	0,5461	**	0,0229	**	**	**	0,011	0,0064	0,0058
Italia	0,01	0,0024	**	0	**	**	**	5E-04	0	**	0,4519	**	0,0207	0,0022	0,003
China	0,318	**	0,0014	0	**	0,0042	0,0268	**	0,1246	0,0006	**	**	0,3116	0	0,0046
China 1	0	0,0002	0,0001	0	**	0,6148	0,4514	0	0,011	**	0,2361	0,2167	0,1351	0	0,0001
Corea	**	0,0006	**	0	**	**	**	0	0,0004	**	0,1153	0	0	0	**
Egipto	0,007	0	**	0,01	**	**	**	**	0,0004	**	0,2656	0,4227	0,0247	0,0979	0
Japon	0,469	0	**	0	**	**	**	0	0,1936	**	0,0301	**	0	0	0
India	0,027	0	**	0	**	0,0024	0,2074	0	0,4928	**	0,0399	0,0172	0,0027	0,0097	0,0117
Malasia	0,028	0,0035	0	**	**	0,0005	0,0315	0	0,014	**	0,4723	0,9624	0,002	0	0
Tailandia	0	0,0823	**	0	**	**	**	0	0,0118	**	0,5016	**	0,0651	0	0,0045

4.3.3 Locus D18S51

Este marcador presenta 2 alelos únicos para dos de las 21 poblaciones comparadas, el alelo 15,2 esta presente únicamente en la población Venezolana (Chiurrillo *et al*, 2003) y el alelo 10/2 en la de Malasia (Seah *et al*, 2003). El alelo 13,2 se encuentra presente en las poblaciones de valle del Cauca (Gómez *et al*, 2003), Venezuela (Chiurrillo *et al*, 2003) y España (Paredes *et al*, 2003) (Tabla 35).

El alelo que presenta mayor frecuencia en nuestra población de estudio es el 17, el cual muestra diferencias significativamente estadísticas con las poblaciones de China ($p = 0.0132$), India ($p = 0.0030$) y Malasia (0.0028).

4.3.4 Locus TH01

Aunque con frecuencias muy bajas este locus presenta el alelo 4 y 6/3 únicamente para poblaciones colombianas (Benítez-Paéz *et al*, 2003), mientras que el alelo 11 es encontrado en el 75% de las poblaciones asiáticas únicamente (Yamamoto *et al*, 1999; Ahmed *et al*, 2001; Chantratita *et al*, 2001; Seah *et al*, 2003 y Cho *et al*, 2003). En este marcador se encontraron 7 de los 13 alelos reportados en la literatura (Tabla 35).

El 59% de las poblaciones que reportan datos para este marcador muestran diferencias estadísticamente significativas con respecto la población santandereana (Tabla 36).

El alelo de menor frecuencia es el 6/3 presente solo en la población santandereana y la del Valle del Cauca, las cuales no muestran diferencias estadísticamente significativas ($p= 0.4198$).

4.3.5 Locus DS3S1358

El 85.71% (18) de las poblaciones comparadas presentan reporte para este locus. Se encontró que el alelo 11, el 9 y el 15,2 están solamente en las poblaciones americanas (Luna-Vázquez *et al*, 2003; Chiurrillo, *et al*, 2003; Gómez, *et al*, 2003 y Rey, *et al*, 2003) mientras que el 9 y el 15,2 hacen parte de las poblaciones colombianas (Gómez, *et al*, 2003 y Rey, *et al*, 2003) (Tabla 35).

Todas las poblaciones Europeas (Gehrig *et al*, 1999; Biondo *et al*, 2001; Anorin *et al*, 2001 y Paredes *et al*, 2003) y el 65% de las poblaciones Asiáticas presentan diferencias estadísticamente significativas con respecto a la población de referencia (Tabla 36).

4.3.6 Locus Penta D

Solo tres de los 21 estudios revisados en la literatura (Andrade *et al*, 2003; Zhang *et al*, 2003 y Paredes *et al*, 2003) reportan frecuencias para este marcador, tal vez es por esta razón que el 37.5% del total de alelos reportados en estas poblaciones son únicos para ellas (Tabla 35).

Ningún alelo reportado para este marcador presenta diferencias significativas, lo que se ve reflejado en las frecuencias para cada uno de ellos presentes en las 4 poblaciones que lo reportan (Tabla 36).

4.3.7 Locus D7S820

Los Individuos de la población santandereana presentan 8 de los 12 alelos encontrados en la literatura para este marcador; el alelo 6/3 solo se encuentra en la población de Norte de Portugal (Anorin *et al*, 2001), mientras que el alelo 6 forma parte de las poblaciones americanas (Gómez *et al*, 2003 y Bydlowski *et al*, 2003) y asiáticas (Seah *et al*, 2003), pero no europeas (Gehrig *et al*, 1999; Biondo *et al*, 2001; Anorin *et al*, 2001 y Paredes *et al*, 2003). El alelo 5 es único en las poblaciones de china, India y Malasia (Seah *et al*, 2003) (tabla 35).

Este marcador analizado en el 100% de las poblaciones no presenta diferencias significativas entre las poblaciones Americanas con respecto a la de referencia, sin embargo todas las poblaciones Europeas y el 62% de las Asiáticas si muestran diferencias (Tabla 36).

4.3.8 Locus D13S317

Este marcador presentó 2 alelos únicos para las poblaciones revisadas, el 6 para la población de Corea (Cho *et al*, 2003) y el 16 para la población de España (Paredes *et al*, 2003) (Tabla 35).

El alelo que presenta mayor frecuencia en nuestra población de estudio es el 12, el cual muestra diferencias estadísticamente significativas con las poblaciones de China ($p=0.0088$), China1 ($p=0.0011$) y Corea ($p=0.0002$). Aunque el alelo 15 es el de menor frecuencia, solo se encontró en las poblaciones americanas revisadas (Tabla 36).

4.3.9 Locus D5S818

Se encontró en la población santandereana 8 (72.72%) de los 11 alelos reportados por la literatura y 9 (84.81%) en la población colombiana (Benítez-Páez *et al*, 2003). Los alelos 16 y 17 identifican a la población de Malasia de las otras comparadas, ya que no se encuentran en ninguna otra.

La población de Ecuador (Andrade, *et al*. 2003), tres Europeas y todas las Asiáticas son estadísticamente diferentes a la población santandereana (tabla 36).

4.3.10 Locus D16S539

En el caso D16S539 Santander tiene el 88.88% de los alelos reportados, El alelo 15 aunque en frecuencias muy bajas está presente en poblaciones de los tres continentes comparados (Rey *et al*, 2003; Bydlowski *et al*, 2003; Andrade *et al*, 2003 y Chiurrillo *et al*, 2003; Gehrig *et al*, 1999; Anorin *et al*, 2001; Paredes *et al*, 2003; Seah *et al*, 2003; Cho *et al*, 2003; Ahmed *et al*, 2001).

El alelo que presenta mayor frecuencia en nuestra población de estudio es el 11, el cual no muestra diferencias estadísticamente significativas con ninguna de las poblaciones comparadas. De igual forma el alelo 15 segregado con menor frecuencia dentro de la población santandereana, tampoco presenta diferencias con las poblaciones que también lo reportan. Sin embargo el Chi-cuadrado global indica diferencias entre las poblaciones de España (Paredes *et al*, 2003), Corea (Cho *et al*, 2003) e India (Seah *et al*, 2003), con respecto a la santandereana (Tabla 36).

4.3.11 Locus CSF1PO

Aunque en frecuencia muy baja el alelo 6 es único para la población santandereana que tiene el 82% de los alelos reportados por la literatura (Tabla 34). Tanto el alelo de mayor frecuencia como el de menor frecuencia para la población santandereana están presentes en las poblaciones de los tres continentes (Yunis *et al*, 2000; Benítez-Paéz *et al*, 2003; Rey *et al*, 2003; Gómez *et al*, 2003; Bydlowski *et al*, 2003; Andrade *et al*, 2003 y Chiurrillo *et al*, 2003; Gehrig *et al*, 1999; Biondo *et al*, 2001; Paredes *et al*, 2003; Yamamoto *et al*, 1999; Ahmed *et al*, 2001; Chantratita *et al*, 2001; Seah *et al*, 2003 y Cho *et al*, 2003).

Este es uno de los marcadores que presento diferencias con menos poblaciones a nivel mundial, con solo tres: Brasil (Bydlowski *et al*, 2003), Japón (Yamamoto *et al*, 1999) e India (Seah *et al*, 2003) (Tabla 36).

El alelo que presenta mayor frecuencia en nuestra población de estudio es el 12, el cual muestra diferencias significativamente estadísticas con las poblaciones de Brasil (Bydlowski *et al*, 2003) y Corea (Cho *et al*, 2003).

4.3.12 Locus D21S11

El Locus D21S11 es uno de los que más presenta alelos únicos para las poblaciones con un total de 6, de los cuales dos, el 25 y el 29,3 solo se encontraron en la población portuguesa, uno en la mexicana (27,2), uno en la ecuatoriana (35,2), uno en la santandereana (37) y uno en la población India (23,2). El alelo 24,2 esta presente en la población colombiana (Gómez *et al*, 2003) y en la de India (Seah *et al*, 2003),

mientras que los alelos 36 y 37 están ausentes en las poblaciones Europeas (Gehrig *et al*, 1999; Biondo *et al*, 2001; Anorin *et al*, 2001 y Paredes *et al*, 2003) y Asiáticas (Yamamoto *et al*, 1999; Ahmed *et al*, 2001; Chantratita *et al*, 2001; Zhang *et al*, 2003; Seah *et al*, 2003 y Cho *et al*, 2003).

Este marcador es el único que presenta diferencias estadísticamente significativas con una población colombiana (Tabla 36) a causa de las frecuencias del alelo 33.2 ausente en la población de Boyacá.

4.3.13 Locus FGA

Este marcador fue el más polimórfico en la población santandereana, se encontraron 14 (51.85%) de los 32 alelos reportados para este sistema; la población venezolana (Chiurrillo *et al*, 2003) presenta dos alelos que están solamente en esta población con respecto al total de poblaciones comparadas, el 31,2 y el 43,2 con una frecuencia de 0.002 para cada uno de ellos; al igual que la población anterior. La China (Seah *et al*, 2003) presenta dos alelos únicos que la diferencian del resto de poblaciones comparadas, el 11 y el 26/2 con frecuencias de 0.002 y 0.005 respectivamente.

El alelo 20/2 es único de poblaciones americanas (Rey *et al*, 2003; Chiurrillo *et al*, 2003; Luna-Vásquez *et al*, 2003).

El 46% de las poblaciones que reportan frecuencias para este marcador presentan diferencias con respecto a la población santandereana (Tabla 36), al igual que con la mayoría de los marcadores anteriores las diferencias más grandes se dan con respecto a las poblaciones

Asiáticas (Yamamoto *et al*, 1999; Ahmed *et al*, 2001; Chantratita *et al*, 2001; Zhang *et al*, 2003; Seah *et al*, 2003 y Cho *et al*, 2003).

4.3.14 Locus D8S1179

El 91% de los alelos reportados por la literatura están presentes en la población santandereana. El alelo 19 solo esta reportado para la población de Boyacá (Rey *et al*, 2003) y Colombia (Benítez-Páez, *et al*, 2003); el 18 esta presente en la población del Valle del Cauca (Gómez *et al*, 2003), india (Seah *et al*, 2003) y Malasia (Seah *et al*, 2003), pero no en poblaciones Europeas, mientras que el alelo 8 esta en Boyacá (Rey *et al*, 2003), Bogotá (Rey *et al*, 2003), Valle del Cauca (Gómez *et al*, 2003), Ecuador (Andrade *et al*, 2003), México (Luna-Vásquez *et al*, 2003), Venezuela (Chiurrillo *et al*, 2003), España (Paredes *et al*, 2003) y Norte de Portugal (Anorin *et al*, 2001), pero no en Asia (Tabla 35).

Solo tres poblaciones de las 14 que presentan frecuencias para dicho marcador presentan diferencias, estas son: España (Paredes *et al*, 2003), China (Zhang, *et al*. 2003) y China (Seah *et al*, 2003).

El alelo que presenta mayor frecuencia en nuestra población de estudio es el 13, el cual muestra diferencias estadísticamente significativas con las poblaciones de China (Zhang *et al*, 2003) y China (Seah *et al*, 2003). El alelo de menor frecuencia es el 17 el cual no presenta diferencias con otras poblaciones (Tabla 36).

4.3.15 Locus Penta E

Solo el 10% de las poblaciones comparadas (2) reportan frecuencias para este marcador (Andrade *et al*, 2003; Paredes *et al*, 2003). El alelo

24 esta presente solo en la población santandereana, mientras que el alelo 23 se segrega en la población Ecuatoriana (Andrade, *et al.* 2003) y no en la santandereana (Tabla 35).

Todas las poblaciones comparadas presentan heterocigocidades en un rango de 0.592 (para el marcador D3S1358) a 0.913 (para el FGA), para las poblaciones de Venezuela (Chiurrillo, *et al.* 2003) y México (Luna-Vasquéz, *et al.* 2003) respectivamente (Tabla 37).

En cuanto a la comparación de las heterocigocidades, basada en una prueba de Chi-cuadrado, la población de Brasil (Bydlowski *et al.*, 2003) y la población de China (Zhang, *et al.* 2003) son las que presentan una mayor diferencia con respecto a la santandereana. Solo la población de Boyacá presenta diferencias con las demás poblaciones colombianas (Tabla 38). Con respecto a la relación de la heterocigocidad entre la población Santandereana y las comparadas según lo reportado por la literatura, estas difieren posiblemente por los diferentes aportes de mestizaje que se han dado en las 2 poblaciones.

Tabla 37. Heterocigocidad observada en las poblaciones comparadas. En negrilla se resalta la heterocigocidad media para cada población.

	FGA	TPOX	D8S1179	vWA	PentaE	D18S51	D21S11	TH01	D3S1358	PentaD	CSF1PO	D16S539	D7S820	D13S317	D5S818	
Santander	0,9	0,6511	0,7635	0,753	0,8887	0,8678	0,8358	0,746	0,75992	0,839	0,7069	0,8319	0,7744	0,8468	0,7027	0,7911
Colombia		0,6612	0,7965	0,784				0,745			0,7278	0,8027	0,7967	0,8184	0,7167	0,761
Bogotá	0,838	0,621	0,797			0,891	0,789	0,802	0,758		0,681	0,798	0,732	0,863	0,704	0,7728
Boyacá	0,822	0,756	0,761			0,882	0,806	0,715	0,794		0,805	0,752	0,818	0,832	0,712	0,7879
V. Del Cauca	0,875	0,6703	0,764	0,769	*	0,8754	0,839	0,744	0,7436	*	0,7176	*	0,7857	0,8129	0,7415	0,7179
Venezuela	0,914	0,702	0,7412	0,769	*	0,8549	0,8392	0,769	0,7804	*	0,7333	0,7529	0,8118	0,7843	0,749	0,8499
Ecuador**	*	*	*	*		*	*	*	*	*	*	*	*	*	*	*
Brazil	*	0,716	*	0,764	*	*	*	0,786	*	*	0,786	0,799	0,825	0,779	0,735	0,7737
Mexico	0,876	*	0,8142	0,673	*	0,8761	0,8319	*	0,5929	*	*	*	0,823	0,7788	0,6549	0,7666
Suiza	0,893	0,607	0,811	0,791	*	0,854	0,869	0,767	0,777	*	0,699	0,767	0,791	0,723	0,68	0,7714
España	0,912	0,5882	0,8235	0,833	0,8284	0,8818	0,7892	0,804	0,8186	0,7843	0,6912	0,7794	0,8522	0,798	0,6814	0,9123
N. Portugal	0,862	*	0,813	0,813		0,885	0,842	*	0,776	*	*	*	0,784	0,79	0,711	0,8084
Italia**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
China	0,842	*	0,867	0,85	0,875	0,833	0,808		0,69	0,783	*	*	0,783	0,741	0,823	0,7408
China 1	0,854	0,557	0,852	0,766	*	0,872	0,817	0,706	0,794	*	0,699	0,781	0,781	0,817	0,836	0,7792
Korea	*	0,613	*	0,799	*	*	*	0,641	*	*	0,716	0,76	0,753	0,778	*	0,6325
Egipto	0,89	0,628	*	0,857	*	*	*	0,736	0,823	*	0,689	0,75	0,742	0,81	0,736	0,8511
Japon	0,82	0,617	*	0,767	*	*	*	0,612	0,665	*	0,728	*	0,743	0,806	0,782	0,7266
India	0,866	0,699	0,823	0,809	*	0,833	0,823	0,799	0,775	*	0,679	0,818	0,828	0,78	0,665	0,7843
Malasia	0,824	0,619	0,852	0,748	*	0,848	0,871	0,743	0,757	*	0,767	0,752	0,752	0,781	0,79	0,842
Tailandia	0,83	0,75	*	0,81	*	*	*	0,73	0,71	*	0,7	*	0,78	0,78	0,76	0,7611

**Poblaciones que no reportan heterocigocidad observada

*Marcadores no analizados para dicha población

Tabla 38. Comparación de las heterocigocidades para todas las poblaciones basada en una prueba de Chi-cuadrado.

	FGA	TPOX	D8S1179	vWA	PentaE	D18S51	D21S11	TH01	D3S1358	PentaD	CSF1PO	D16S539	D7S820	D13S317	D5S818
Colombia	**	0,789	0,337	0,348	**	**	**	0,974	**	**	0,536	0,264	0,484	0,335	0,716
Bogotá	0,060	0,561	0,419	**	**	0,617	0,181	0,226	0,849	**	0,113	0,423	0,260	0,637	0,593
Boyacá	0,022	0,032	0,800	**	**	0,673	0,359	0,512	0,423	**	0,722	0,056	0,283	0,586	0,782
V. Del Cauca	0,245	0,442	0,987	0,553	**	0,724	0,878	**	0,548	**	0,713	**	0,662	0,183	0,173
Venezuela	0,690	0,211	0,552	0,667		0,586	0,996	0,607	0,574	**	0,550	0,019	0,285	0,067	0,230
Brazil	**	0,000	**	0,003	**	**	**	0,205	**	**	0,017	0,223	0,097	0,030	0,353
México	0,490	**	0,238	0,078	**	0,789	0,912	**	0,0005	**	**	**	0,259	0,089	0,324
Suiza	0,818	0,318	0,212	0,320	**	0,682	0,304	0,587	0,665	**	0,866	0,769	0,653	0,001	0,589
España	0,654	0,160	0,110	0,033	0,074	0,625	0,199	0,132	0,121	0,128	0,724	0,150	0,043	0,140	0,545
N. Portugal	0,142	**	0,120	0,056	**	0,493	0,821	**	0,813	**	**	**	0,776	0,072	0,640
China	0,100	**	0,019	0,031	0,770	0,373	0,217	**	0,155	0,185	**	**	0,254	0,014	0,018
China 1	0,083	0,035	**	0,007	**	0,896	0,548	0,317	0,424	**	0,884	0,159	0,859	0,355	0,001
Korea	**	0,487	**	0,039	**	**	**	0,012	**	**	0,477	0,595	0,909	0,097	**
Egipto	0,706	0,578	**	0,002	**	**	**	0,799	0,060		0,735	0,059	0,483	0,355	0,534
Japón	0,044	0,419	**	0,757	**	**	**	0,002	0,023	**	0,652	**	0,388	0,242	0,058
India	0,198	0,305	0,137	0,135	**	0,293	0,641	0,199	0,000	**	0,478	0,620	0,179	0,049	0,340
Malasia	0,015	0,467	0,015	0,898	**	0,535	0,336	0,947	0,849	**	0,135	0,031	0,574	0,066	0,028
Tailandia	0,065	0,070	**	0,245	**	**	**	0,761	0,326	**	0,909	**	0,905	0,134	0,274

4.4 ANÁLISIS DE AGRUPAMIENTO

Debido a la insuficiencia de datos publicados para todos los marcadores utilizados en este estudio, fue posible comparar solo 10 marcadores con un total de 135 alelos (TPOX, D8S1179, D18S51, D21S11, TH01, D3S1358, CSF1PO, D7S820, D13S317 y D5S818) para 10 poblaciones (Bogotá, Rey, *et al.* 2003; Boyacá, Rey, *et al.* 2003; Valle del Cauca, Gómez, *et al.* 2003; Venezuela, Chiurrillo, *et al.* 2003; Suiza, Gehrig, *et al.* 1999; España, Paredes, *et al.* 2003; China, Zhang, *et al.* 2003, Malasia, Seah, *et al.* 2003; India, Seah, *et al.* 2003).

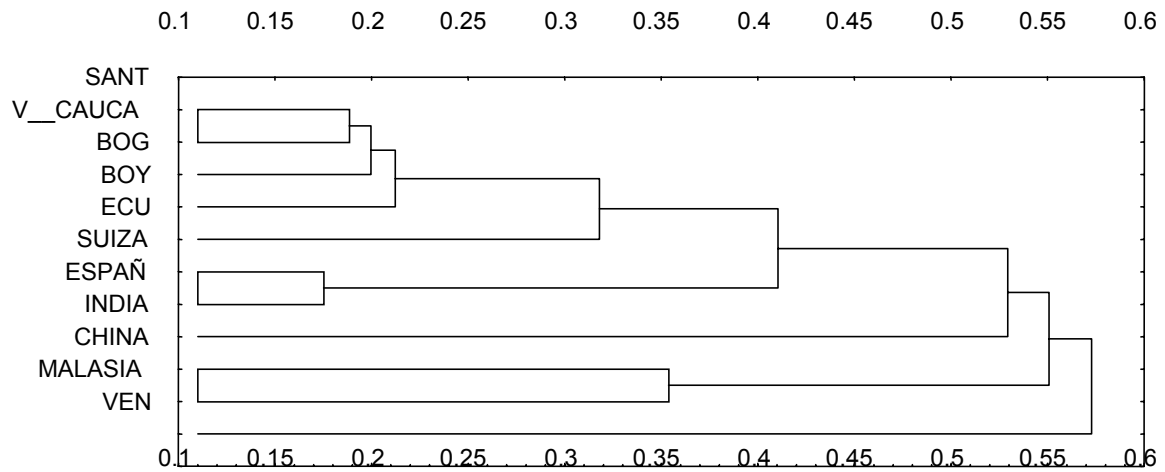
Las distancias genéticas entre la población santandereana y las otras poblaciones a nivel mundial fueron obtenidas por medio del programa Statistica 4.0. Estas distancias varían de una mínima de 0.1750 (España-Suiza) a una máxima de 0.8230 (China - Venezuela). Santander, en particular, muestra una distancia pequeña con la población de Valle del Cauca y una mayor con la población de Malasia (Tabla 39).

El dendograma construido con base en las distancias genéticas euclidianas de las poblaciones comparadas (Figura 31) muestra un grupo de poblaciones Colombianas seguido de la población Ecuatoriana, de igual forma se agrupan las poblaciones Europeas y Asiáticas con Venezuela en medio de estas. Es clara la ancestralidad por parte de las poblaciones Europeas con respecto a las Americanas, Sin embargo existe algún proceso a nivel genético que nos aleja de la población venezolana.

Tabla 39. Matriz de distancias genéticas para la población Santandereana y otras poblaciones a nivel mundial.

	SANT	BOG	BOY	V__CAUCA	VEN	ECU	SUIZA	ESPAÑ	CHINA	INDIA
SANT	0									
BOG	0,2523	0								
BOY	0,2123	0,2400	0							
V__CAUCA	0,1885	0,1997	0,2433	0						
VEN	0,5826	0,5798	0,5901	0,5729	0					
ECU	0,3425	0,3180	0,3667	0,3635	0,6623	0				
SUIZA	0,4104	0,4794	0,4790	0,4177	0,6158	0,5922	0			
ESPAÑ	0,4367	0,5053	0,5005	0,4427	0,6205	0,6372	0,1751	0		
CHINA	0,6640	0,6409	0,6804	0,6567	0,8230	0,7039	0,6815	0,7105	0	
INDIA	0,5385	0,5295	0,5682	0,5363	0,6849	0,6651	0,5714	0,5490	0,5861	0
MALASIA	0,6658	0,6184	0,6749	0,6300	0,7553	0,7107	0,6281	0,6545	0,3538	0,5508

Figura 31. Dendrograma basado en el método UPGMA con distancias euclidianas para la población santandereana y otras reportadas a escala mundial.



DESEQUILIBRIO DE LIGAMIENTO

5. DISCUSIÓN DE RESULTADOS

Usando las frecuencias alélicas de los 15 STR analizados en este estudio fué posible inferir nuevos datos que ayudan en el análisis de la estructura genética de la población santandereana.

5.1 FRECUENCIAS ALÉLICAS, GENOTÍPICAS, EQUILIBRIO DE HARDY-WEINBERG Y HETEROCIGOCIDAD

Pocos trabajos con marcadores de tipo STR's existen en poblaciones suramericanas (Chiurrillo, *et al.*2003; Andrade, *et al.*2003 y Bydlowski, *et al.* 2003) y colombianas (Benítez-Páez *et al.* 2003; Yunis, *et al.* 2003; Rey, *et al.* 2003; Gómez, *et al.* 2003; Vargas, *et al.* 2002; Bravo, *et al.* 2001 y Lemus-López, 2000). A pesar de esto, se hacen necesarios trabajos extensivos en este campo, debido a los diferentes procesos de mestizaje que se han dado en los países de Sur América desde que se dio la conquista de América en el año de 1492 (Rey, *et al.* 2003).

Es indudable que el mestizaje, desde el punto de vista étnico, es un proceso lento. A partir de los procesos de la conquista y la colonia, seguidos de los movimientos demográficos sucedidos en el momento de la independencia se empezó a configurar nuestra estructura genética actual, Es importante resaltar que trabajos previamente publicados (Yunis, *et al.* 2001) para poblaciones negras colombianas indican la presencia de alelos, como el 5 para el marcador TH01, el 7 para el TPOX, el 7 para el CSF1PO y el 11 y el 21 para vWA, que se encuentran ausentes en la población santandereana, pero se segregan en la población del Valle del Cauca; si se tiene en cuenta que el aporte negro para esta zona del país es del 21% (Sandoval, 1993) se podría

inferir que estos alelos son propios de poblaciones negras o de descendencia africana y que ciertos componentes genéticos de la raza negroide en esta zona del país aun se mantienen. De igual forma, trabajos para la población colombiana mestiza-caucásica (Yunis, *et al.* 2000) reportan el 74% de los alelos segregados en la población Santandereana, lo que indicaría que nuestra población es básicamente de este tipo, y refuerza la hipótesis de Sandoval (1993) en la que se afirma que el 65% de la población para esta parte del país esta formada por individuos de origen mestizo-caucásico.

Los resultados de la presente investigación demuestran que todos los marcadores estudiados en la población santandereana se encuentran en equilibrio de Hardy-Weinberg (Tabla 2). Este equilibrio esta dado posiblemente por el apareamiento al azar que se da entre los individuos, además de la no selectividad por parte de los diferentes genotipos, lo que quiere decir que todos son igualmente viables (Klug, 1999); adicionalmente, nos lleva a pensar que no existen factores de mutación o selección propias de poblaciones naturales, dinámicas que cambian constantemente de tamaño y estructura como parte de sus ciclos biológicos.

El equilibrio de Hardy-Weinberg ocurre en poblaciones humanas bajo ciertas condiciones; una de ellas, que el cruce entre los individuos sea libre y no se presente selección, mutación o migración en forma significativa, como se menciono anteriormente. La desviación con base en cualquiera de estas variantes puede alterar la frecuencia génica en la población, y desequilibrarla de una generación a otra (Yunis, *et al.* 2003).

Una de las aplicaciones prácticas de la Ley de Hardy-Weinberg es la estimación de las frecuencias de los heterocigotos en una población ya que es su mayoría presentan un rápido aumento de las frecuencias, además de mantener la variabilidad en las poblaciones. Los 15 marcadores analizados mostraron heterocigocidades mayores al 65% (Tabla 2). Debido al gran número de alelos detectados para cada marcador (Tabla 33) y el cálculo de la heterocigocidad se pudo observar la variación genética en la población santandereana, la alta heterocigocidad media observada (Tabla 37) puede ser producto de la gran diversidad alélica, además de considerar que el alto número de alelos obtenidos provee un alto nivel de información genética comparado con los marcadores clásicos.

La relativa alta frecuencia de alelos raros (Tabla 35), ausentes en el resto de poblaciones, parece dar una cierta idea de la diferenciación de la población santandereana con respecto a las otras poblaciones a nivel mundial, aunque no puede descartarse la idea de que dichos alelos sean producto de mutaciones.

5.2 DESEQUILIBRIO DE LIGAMIENTO

La necesidad de considerar las asociaciones entre genes de distintos loci para entender la dinámica de las poblaciones se aprecia inmediatamente en el estudio del desequilibrio gamético. No es suficiente que los genes de cada loci estén en equilibrio de Hardy-Weinberg para que la población lo esté en su conjunto. Las asociaciones gaméticas pueden coexistir perfectamente con estos equilibrios de loci individuales y tardar muchas generaciones en desaparecer, aunque el apareamiento sea al azar y no existan mecanismos mutacionales o selectivos (Fontdevilla *et al*, 1999).

Los marcadores ubicados en el cromosoma 5 se encuentran en equilibrio (Tabla 34), tal vez porque presentan frecuencias de recombinación alta y se segregan de acuerdo a lo esperado bajo el supuesto de Hardy-Weinberg (Futuyma, 1998); o porque los marcadores están ubicados en el mismo cromosoma pero la segregación de los alelos se da de manera independiente.

Es evidente que si los apareamientos son al azar y no existen mecanismos de selección a favor de la asociación gamética, la recombinación ira destruyendo dicha asociación hasta llegar al equilibrio, pero este proceso puede ser muy lento y depende del grado de ligamiento entre los loci (Fontdevilla, *et al.* 1996). Otro mecanismo que puede estar generando el desequilibrio puede ser la mezcla de poblaciones ya que el sistema de apareamiento o la subdivisión poblacional conducen a una disminución de la heterocigocidad afectando las probabilidades de recombinación al interior de un grupo poblacional.

5.3 COMPARACIÓN ENTRE POBLACIONES

La tabla 36 resume los resultados de la prueba de Chi-cuadrado global en la que se reflejan las diferencias entre las diferentes poblaciones comparadas a nivel mundial. Es claro que para las poblaciones colombianas no existe diferencias en ningún marcador a excepción del locus D21S11 en la población de Boyacá (Rey, *et al.* 2003) lo que indicaría un posible aislamiento entre dichas poblaciones en cuanto a caracteres genéticos se refiere. De igual forma se observa que a medida que las poblaciones se van alejando geográficamente de la de referencia, las diferencias se hacen más frecuentes y más significativas, siendo las poblaciones asiáticas las que presentan una mayor

disimilitud con respecto a la nuestra y en general a las colombianas; esto es confirmado por el porcentaje de probabilidades con valor significativo entre las 8 poblaciones asiáticas, ya que alcanza un valor del 76% frente a un 43% de las poblaciones Americanas. Aunque se podría pensar que el flujo génico que se da normalmente en poblaciones humanas disminuiría estas diferencias, es claro que existe un intercambio genético que no es lo suficientemente fuerte como para reducir el aislamiento que de una u otra forma está influenciado también por las grandes distancias geográficas que hay entre las diferentes poblaciones de estudio, además que se debe tener en cuenta también que los miembros de una población particular prefieren cruzarse con los del mismo grupo cuando median criterios raciales, étnicos, religiosos, o de las mismas subculturas, por posición socioeconómica, o condición médica, o por vínculos familiares (Yunis, *et al.* 2002). De la misma manera la tabla 36 resume las probabilidades de las heterocigocidades promedio entre poblaciones para cada marcador, aquí se observan las diferencias de nuevo en la población de Boyacá para el locus FGA y TPOX; las poblaciones de China y Brasil son las que presentan una mayor diferencia con cuatro marcadores, D8S1179, vWA, D13S317 y D5S818 para la primera y TPOX, vWA, CSF1PO y D13S317 para la segunda. Poblaciones como Tailandia, Egipto, Portugal, Suiza, y las demás colombianas a excepción de Boyacá no presentan diferencias con nuestra población. Estas discordancias pueden ser producto de un desequilibrio en las frecuencias de las poblaciones que afectan la heterocigidad poblacional y, aunque esto no es tan trascendental, ya que los valores de heterocigidad media son relativamente altos (Tabla 37), si deben ser tenidos en cuenta con cierta importancia ya que se puede reflejar en un aumento de la homocigidad afectando así la variabilidad de las poblaciones afectadas. En general todas las poblaciones reportan una heterocigidad media mayor al 76% lo que

significa una gran diversidad alélica de los marcadores y un gran proceso de mestizaje a lo largo del tiempo.

5.4 ANÁLISIS DE AGRUPAMIENTO

Debido a que muy poco se puede aprender de un locus que es igual en cada individuo, el estudio de la filogenética depende de los loci polimorficos. Durante las ultimas décadas, han sido desarrollados métodos que permiten que tipos diferentes de estos "marcadores" polimorficos sean examinados en los individuos. Antes de la década de los 90, la diversidad genética usualmente se deducía a través de los polimorfismos clásicos (no-ADN), como los tipos sanguíneos, proteínas del suero, aloenzimas e inmunoglobina. Luego, fueron empleadas enzimas de restricción para producir una clase útil de marcador al nivel del ADN, los polimorfismos de restricción de fragmento de longitud (RFLP). También han sido utilizados en estudios genéticos de poblaciones otros loci como el ADN mitocondria (ADNmt), inserciones Alu, minisatélites, polimorfismos de nucleótido único (SNP) y microsatélites (STRP). Los microsatélites probablemente son la más importante clase de marcador utilizado en la actualidad debido a sus altas tasas de polimorfismos, mutaciones rápidas y distribución aleatoria a través del genoma (Cooper, *et al.* 1999). Los loci con altas variables son ventajosos en la filogenética debido a que pueden proporcionar la resolución necesaria para distinguir a poblaciones bastante relacionadas.

El uso de frecuencias alélicas en la construcción de dendogramas, dan la idea general de que las subpoblaciones se organizan de manera compacta representando así al país. Esto indicaría que la estimación de las frecuencias génicas de estas subpoblaciones podrían ser

utilizadas para resolver casos de paternidad y forenses a nivel nacional (Lemus-López, 2000).

Las pequeñas distancias genéticas existentes entre las poblaciones colombianas, indican cierta afinidad genética entre estos grupos, al igual que al interior de las poblaciones Europeas y Asiáticas (Tabla 39). El agrupamiento es aparente y consistente en los múltiples loci genéticos; y las importantes distancias genéticas intergrupales indican historias naturales únicas en la escala de tiempo evolutiva.

Los resultados de la comparación de frecuencias entre poblaciones por medio de la prueba de Chi-cuadrado global pueden ser extrapolados para discutir el análisis de agrupamiento ya que se observa el mismo comportamiento de agrupación para las diferentes poblaciones de los 3 continentes (Figura 31); la longitud de las ramas y la posición de cada población dentro del cladograma formando grupos de acuerdo a su ubicación geográfica reflejan la diferenciación genética existente entre los continentes. Como se mencionó anteriormente estas agrupaciones pueden estar influenciadas por las preferencias de entrecruzamiento entre individuos de una misma población debido a costumbres, raza, creencias, religión, entre otras. El patrón observado en este análisis de agrupamiento ha sido encontrado en trabajos anteriores (Pérez-Lezaun, *et al.* 1997) con marcadores de tipo STR, con polimorfismos de Proteínas (Cavalli-Sforza, *et al.* 1994) y con fragmentos de restricción (Bowcock, *et al.* 1991).

6. CONCLUSIONES

- Los resultados de las frecuencias alélicas proveen datos para la construcción de una base de datos genética, indicando las peculiaridades de algunos marcadores como el Penta D, CSF1PO, D21S11 y Penta E, que pueden ser importantes para discriminar la población de Santander otras, ya que presenta alelos reportados, hasta el momento, únicamente es esta población.
- El equilibrio de ligamiento para los locus CSF1PO y D5S818 ubicados en el cromosoma 5 puede ser debida a las altas frecuencias de recombinación y a la segregación de los alelos que se da independientemente.
- El presente estudio es una revisión de la variación de loci microsatélite entre diferentes poblaciones humanas y proporciona un marco de información de la distribución de las frecuencias alélicas de aquellos marcadores que han sido estudiados en tres continentes: América, Asia y Europa. Esta información puede ser usada para estudiar la variación entre diferentes grupos humanos y ayudar en la inferencias de la historia poblacional de dichos grupos.
- Las pequeñas distancias genéticas presentes entre las poblaciones colombianas indican cierta afinidad entre ellas, al igual que para algunos marcadores, lo que permite a su vez algún tipo de discriminación entre los diferentes grupos.

- Comparaciones de la población Santandereana con otras en el ámbito mundial, reflejan la presencia de un límite genético que separa a las poblaciones colombianas de las europeas y asiáticas.

RECOMENDACIONES

- ☆ Se debe tener especial cuidado con aquellos alelos que presentan frecuencias menores al 1% ya que pueden ser producto de mutaciones lo cual podría interferir en la interpretación de los resultados de las pruebas de paternidad.

- ☆ Este estudio presenta una revisión de la variación de 15 marcadores de tipo STR entre diferentes poblaciones humanas, la cual puede ser usada para inferir la historia de los diferentes grupos humanos, sin embargo se hace necesario implementar más registros de los cinco continentes para predecir de manera mas certera la diferenciación genética y la ubicación de dichas poblaciones dentro de un gran análisis de agrupamiento, además de incorporar al análisis marcadores tales como DNA_{mit} y marcadores del cromosoma Y, para estar seguros de la posición de las poblaciones dentro del análisis.

- ☆ El analizar la población santandereana mediante el cálculo de las frecuencias alélicas actuales y compararla con otras poblaciones, nos permite acercarnos al entendimiento de cómo ha sido el proceso de mestizaje de nuestra región, que sumado a la evidencia antropológica aportaría al conocimiento de las distancias relativas

☆ entre diferentes culturas y ayudaría a determinar los orígenes ancestrales de nuestro grupo poblacional contemporáneo.

BIBLIOGRAFIA

ACOSTA, M. A., BRION, M., LAREU, M. V. & A. CARRACEDO. Genetic data on eight STRs (D5S818, D7S820, F13B, LPL, TH01, vWA31, CSF1PO) from a Colombian population. 2002. *Forensic Sci Int.* 129(3):216-218.

AHMED, A., LINACRE, A. M., MOHAMMED, A. A., VANEZIS, P. & W. GOODWIN. STR population data for 10 STR loci including the GenePrint™ PowerPlex™ 1.2 kit from El.-Mina (Central Egypt). 2001. *Forensic Science International* 117:233-234

AMORIM , A., GUSMAO, L. & C. ALVES. STR data (AmpF/STR profiler plus) from north Portugal. 2001. *Forensic Science International* 115:119-121.

ANDRADE, F.G., SANCHEZ-Q, D. & MARTINEZ-JARRETA, B. 2003. *Forensic Sci Int.* 135:64-66.

BENITEZ-PAEZ, A. & OSSA, R. H. Allelic frequencies at 12 STR loci in Colombian population. 2003. *Forensic Sci Int.* 136:86-88.

BERNAL, J. La diversidad étnica en Colombia. 2000. <http://javeriana.edu.co/humana/terrenos2A.html>

BIONDO, R., SPINELLA, A., MONTAGNA, P., WALSH, P. S., HOLT, C. & B. BUDOWLE. Regional italian allele frequencies at nine short tandem repeat loci. 2001. *Forensic Science International* 115:95-98

BOWCOCK, A. M., A. RUIZ-LINARES., J. TOMFOHRDE., et al. High Resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 1994. 368: 455-457.

----- & L.L CAVALLI-SFORZA. The study variation in the human genome. 1991. *Genomics* 11: 491-498.

BUSQUE, L., DESAMARIS, D., PROVOST, S., SCHUMM J.W., ZHONG, Y. & CHAKRABORTY, R. Analysis of Allele Distribution for Six Short Tandem Repeat Loci in a French Canadian Population of Québec. *J Forensic Sci* 1997; 42(6):1147-1153.

BYDLOWSKI, S. P., MOURA-NETO,R. S.,SOARES, R., SILVA, R., DEBES-BRAVO, A. A. & L. MORGANTI. Genetic data on 12 STR's (F13A01, F13B, FESFPS, LPL, CSF1PO,TPOX, TH01, vWA, D16S539, D7S820, D13S317, D5S818) from four ethnic groups of Sao Paulo, Brazil. 2003. *Forensic Sci Int.* 135:67-71.

CALO C. M., L. GAROFANO., A. MAMELI., M. PIZZAMIGLIO & G. VONA. Genetic Analysis of a Sicilian Population Using 15 Short Tandem Repeat. *Human Biology* 2003 75 (2): 16-178.

CHANTRATTITA, W., RERKAMNUAYCHOKE, B., JOMSAWAT, U., THANAKITGOSATE, J., RUANGVITHAYANON, T. & P. ROJANASUNAN. Thai population data on nine tetrameric STR loci – D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820. 2001. *Forensic Science International* 115:113-115.

CHIURRILLO, M. A., MORALES, A., MENDES, A. M., LANDER, N., TOVAR, F., FUENTES, A. & j. L. RAMIREZ. Genetic profiling of a central venezuelan population using 15 STR markers that may be of forensic importance. 2003. *Forensic Sci Int.* 139:99-101.

CHO, N. S., HWANG, J. H., LEE Y. A. & H. PARK. Population genetics of nine STR loci: TH01, TPOX, CSF1PO, vWA, FESFPS, F13A01, D13S317, D7S820

and D16S539 in a corean population. 2003. *Forensic Science International* 137:97-99.

COMAS, J. 1974. *Antropología de los pueblos americanos*. Editorial Labor. Barcelona.

COOPER, G., AMOS, W., BELLAMY, R., SIDDIQUI, M.R., FRODSHAM, A., HILL, A., RUBINSZTEIN, D. An Empirical Exploration of the (dm)² Genetic Distance for 213 Human Microsatellite Markers. *Am. J. Hum. Genet.* 1999; 65:1125-1133.

DEKA, R., JIN. L., SHRIVER, M. D. YU, L. M., DECROO, S., HUNDRIESER, J., BUNKER, C. H., FERRELL, R. E. & R. CHAKRABORTY. Intra- and interpopulation diversity at short tandem repeat loci in diverse populations of the world. *Electrophoresis.* 16:1659-1664.

DROBNIC, K. & BUDOWLE, B. The Analysis of Three Short Tandem Repeat (STR) Loci in Slovene Population by Multiplex PCR. *J Forensic Sci.* 2000; 45(4):893-895.

EDWARDS, A., CIVITELLO. A., HAMMOND. H. A. , CASKEY, C. T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* 1991; 49:746-756 in MARTINEZ, 1999. *La prueba del ADN en medicina Forense*.

EDWARDS, A., HAMMOND H. A., LIN, J., CASKEY, C. T. CHAKRABORTY, R., Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992 12:241-53.

FONTDEVILA, A. & A. MOYA. *Introducción a la genética de poblaciones*. 1999. Editorial Sintesis. Madrid- España.

FUREDI, S., WOLLER, J. & Z. PADAR. Hungarian population data for the STR systems TH01 and VWA. *Int. J. Leg. Med.* 108:48-49.

FUTUYMA, D. Evolutionary Biology. Population structure and genetic drift.. 1998. Sinauer associates, Inc.

GEHRIG., C., HOCHMEISTER, M., BORER, U. V., DIRNHOFER, R. & BUDOWLE, B. Swiss Caucasian Population Data for 13 STR Loci Using AmpFISTR Profiler Plus and Cofiler PCR Amplification Kits. J Forensic Sci. 1999; 44(5):1035-1038.

GOMEZ, M.V., REYES, M. E., CARDENAS, H. & GARCIA, O. Genetic variation for 12 STR's loci in a Colombian population (Department of Valle del Cauca). 2003. Forensic Sci Int. 137:235-237.

HARTL, D., 2000. A primer of population genetics. Third edition. Sinauer Associates, Sunderland.

HAUGE X. Y. & M. LITT. A study of the origin shadow bands seen when typing dinucleotide repeat polymorphism by the PCR. Hum Mol Genet 1993; 2: 411-415 In MARTINEZ, J. B. 1999. La prueba de ADN en medicina Forense. Editorial Masson S.A. Barcelona- España.

HEARNE, C., GHOSH, S. & J. TODD. Microsatellites for linkage analysis og genetics traits. Tr. Genet. 1992; 8:288-294

HERNANDEZ, M. Migración, Colonización y Flujo Génico. Rev. Arg. Antrop. Biol. 2001. 3 (2): 19.

JORDE, L. B., BAMSHAD, M.J., WATKINS, W. S., ZENGER, R., FRALEY, A. E., KRAKOVIAK, P.A., CARPENTER, K. D., SOODYAL, H., JEMKINS, T. & A. R. ROGERS. Origins and affinities of modern humans: A comparasion of mitochondrial and nuclear genetics data. Am. J. Hum. Genet. 1995. 57: 523-538.

KLITZ, W., REYNOLDS, R., CHEN, J. & H. A. ERLICH. ANALYSIS OF GENOTYPE FRECUENCIES AND Interlocus Associations for the PM, DQA1, and D1S80 Loci In Four Populations. J Forensic Sci. 2000; 45(5) 1009-1015.

KLUG, W.S. & CUMMINGS, M. R.1999. Conceptos de genética. Quinta Edición. Prentice Hall. Madrid.

KRIEGER, et al. Racial admixture in North-eastern Brazil. 1965. In: SANDOVAL, C. E. 1993. Estructura genética de la población Colombiana. Tesis de grado. Universidad Nacional de Colombia.

LEMUS-LOPEZ, N.M. 2000 Caracterización genética de la población humana de los departamentos de Boyacá, Norte de Santander y Santander, mediante loci STR. Tesis de grado. Universidad Nacional de Colombia.

LISKER, R. et al. Gene frequencies and admixture estimates in four Mexican urban center. 1988. In: SANDOVAL, C. E. 1993. Estructura genetica de la población Colombiana. Tesis de grado. Universidad Nacional de Colombia.

LUNA-VAZQUEZ, A., VILCHIS-DORANTES, G., PAEZ-RIBEROS, L. A., MUÑOZ-VALLE, F., GONZALEZ-MARTIN, A. & H. RANGEL-VILLALOBOS. Population data on nine STRs of mexican-mestizos from Mexico City. 2003. Forensic Sci Int. 136:96-98.

MARTIN, P., ALONSO, B., & B. BUDOWLE. Spanish population data on 7 tetrameric short tandem repeat loci. Int. J leg Med. 1995. 108:145-149.

MARTINEZ, J. B. 1999. La prueba de ADN en medicina Forense. Editorial Masson S.A. Barcelona- España

NEI, M. & K. ROYCHOUDHURY. Evolutionary relationships of human population on a global scale.. Mol Biol Ecol. 1993; 10(5):927-943.

----- & A. ROYCHOUDHURY. Samples Variances of heterozigosity and genetic distance. *Genetics*. 1974; 89:583-590.

OLDROYD, N., URQUHART, A., & C. KIMPTON. A highly discriminating octoplex short tandem repeat polymerase chain reaction system suitable for human individual identification. *Electrophoresis*. 1995. 16:334-337.

PAREDES, M., CRESPILO, M., LUQUE, J. A. & J. L. VALVERDE. STR frequencies for the powerPlex 16 system kit in a population from Northeast Spain. 2003. *Forensic Science International* 135:75-78.

PARSON, J. La Colonización Antioqueña del Occidente Colombiano. 1950. In: SANDOVAL, C. E. 1993. Estructura genética de la población Colombiana. Tesis de grado. Universidad Nacional de Colombia.

PEREZ-LEZAUN, A., CALAFELL, F., MATEU, E., COMAS, D., BOSCH, E. & J. BERTRANPETIT. Allele frequencies for 20 microsatellites in a worldwide population survey. *Hum Hered*. 1997. 47:189-196.

PIAZZA, P. M. & I. L. Cavalli-Sforza. 1981. The making and testing of geographic gene frequency maps. *Biometrics*. 37: 635-659.

PINHEIRO, F., PONTES, L., PINTO DA COSTA, J., HUGUET, E., MORNO P. & m GENE. Allelic distribution of four tetranucleotide repeat loci (D3S1358, D18S51, D19S253, and FGA) in a population from Porto (North Portugal). *J Forensic Sci*. 2000. 45(4):891-892.

PUERTAS, M. J. *Genética fundamentos y perspectivas*. 1992. Mc Graw-Hill. Madrid.

REED, T. E. Caucasian genes in American negroes. 1969. In: SANDOVAL, C. E. 1993. Estructura genética de la población Colombiana. Tesis de grado. Universidad Nacional de Colombia.

REY, M., GUTIERREZ, A., SCHROEDER, B., USAQUEN, W., CARRACEDO, A., BUSTOS, I. & A. GIRALDO. Allele frequencies for 13 STR's from two Colombian populations: Bogotá And Boyacá. 20003. Forensic Sci Int. 136:83-85.

RODRIGUEZ PLATA, H. La Migración Alemana al Estado Soberano de Santander en el Siglo XIX. Ed. Kelly. Bogotá. 1968.

ROSE, G., DE LUCCA, M. & E. FALCONE. Allele frequency distributions at seven DNS hypervariable loci in a population sample from Calabria (Southern Italy). Gene Geogr. 1996; 10:135-145.

SALDANHA, P. H. Race mixture among northeastern Brazilian population. 1962 In: SANDOVAL, C. E. 1993. Estructura genética de la población Colombiana. Tesis de grado. Universidad Nacional de Colombia.

SANDOVAL, C. E. 1993. Estructura genética de la población Colombiana. Tesis de grado. Universidad Nacional de Colombia.

SEAH, L. H., JEEVAN, N. H., OTHMAN, M. I., JAYA, P., OOI, Y. S., WONG, P. C. & S. S. KEE. STR data for the AmpF/STR identifier loci in three ethnic groups (Malay, Chinese, Indian) of the Malaysian population. 2003. Forensic Science International 138:134-137.

SUZUKY, D., GRIFFITHS, A., MILLER, J. & R. LEWONTIN. Introduction Al Análisis Genético. 1996. 4^a ed. España.

TOURRET, N., LOPEZ-CAMELO, J. & VIDAL-RIOJA, L., Allele frequencies of six STR loci in Argentine populations. J. Forensic Sci. 1999; 44(6): 1265-1269

VARGAS, C. Allele frequencies for the AMPFistr profiler loci in a Colombian population (Department of Boyaca). J Forensic Sci. 2002. 47(2) 406-407.

URQUHART, A., OLDROYD, N. J., KIMPTON C. P. & P. GILL. *Biotechniques* 1995; 18: 116-121. In MARTINEZ, J. B. 1999. *La prueba de ADN en medicina Forense*. Editorial Masson S.A. Barcelona- España

YAMAMOTO, T., UCHIHI, R., NOZAWA, H., HUANG, X-L., LEONG Y-K., TAMAKA, N., MIZUTAMI, M., TAMAKI, K. & KATSUMATA, Y., Allele Distribution at Nine STR Loci –D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S1317, and D7S820_ in the Japanese Population by Multiplex PCR and Capillary Electrophoresis. *J. Forensic. Sci.* 1999; 44(1):167-170.

YUNIS, E. J. & J. J. YUNIS. *El ADN en la identificación Humana*. 2001. Editorial Temis S. A. Bogotá – Colombia.

-----, GARCÍA, O., MORENO, S., Uriarte, I. & Yunis, E. J. Population data on D7S820, FGA, D1S533 and D9S304 in a sample of Caucasian-Mestizos from Colombia International Congress Series. *Forensic in Genetics*. 2002. 201-205.

-----, GARCÍA, O., URIARTE, I. & YUNIS, E. J. Population data on D16S539, D7S820, D13S317, LPL, F13B, and D1S80 in a sample of caucasian- mestizos from Colombia. *Forensic Science International* 2001. 115:117-118.

-----, GARCÍA, O., MORENO, S., PINEDA, C., RODRIGUEZ, C., URIARTE, I. & YUNIS, E. J. Population data on Poweplex 2.1 (FGA, vWA, TPOX, TH01, Penta E, D18S51, D21S11, D3S1358, D8S1179) and Gamma Star (D16S539, D7S820, D13S317, D5S818) in a sample in caucasian-mestizos from Colombia. International Congress series. *Progress in Forensic Genetics*. ICS 1239, 2002, pp 207-212.

-----, GARCÍA, O., URIARTE, I. & YUNIS, E. J. Population data on 6 short tandem repeat loci in a sample of Caucasian-mestizos from Colombia. *Int. J. Legal Med.* 2000. 113(3):175-178.

-----, GARCIA, O., BAENA, A., ARBOLEDA, G., URIARTE, I. & E. YUNIS. Population frequency for the short tandem repeat loci D18S849, D3S1744, and

D12S1090 in Caucasian-mestizo and African descent population of Colombia.2000. J. Forensic Sci. 45(2):429-431

ZHANG, Y. J., XU, Q. S. & J. B. LEE. Population data for 11 STR loci in northeast China Han. 2003. Forensic Science International 138:116-118.

ZHIVOTOVSKY, L. A., BENNETT, L., BOWCOCK, A. M. & FELDMAN, M. W. Human Population Expansion And Microsatellite Variation. Mol. Biol. Evol. 2000. 17(5): 757-767.