



FACULTAD DE INGENIERÍA Y CIENCIAS AGROPECUARIAS

DETERMINACIÓN DE LOS NIVELES PRELIMINARES DE ANCESTRÍA
EN PACIENTES DE LA POBLACIÓN ECUATORIANA CON MUTACIÓN
EN EL GEN IT-15

AUTOR

Estefanía Nathaly Eras Muñoz

AÑO

2017



FACULTAD DE INGENIERIA Y CIENCIAS AGROPECUARIAS

DETERMINACIÓN DE LOS NIVELES PRELIMINARES DE ANCESTRÍA EN
PACIENTES DE LA POBLACIÓN ECUATORIANA CON MUTACIÓN EN EL
GEN IT-15

Trabajo de titulación presentado en conformidad con los requisitos establecidos
para optar por el título de Ingeniera en Biotecnología

Docente Guía
Ph.D. Alina Aguirre Quevedo

Autor
Estefanía Nathaly Eras Muñoz

Año
2017

DECLARACIÓN DEL DOCENTE GUIA

“Declaro haber dirigido este trabajo a través de reuniones periódicas con la estudiante, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación”.

Alina Aguirre Quevedo

Doctora en Ciencias Biológicas

C.I. 175659460-0

DECLARACIÓN DEL PROFESOR CORRECTOR

“Declaro haber revisado este trabajo, dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación”.

María Gabriela Granja Bastidas

Máster en Bioquímica, Biología Molecular y Biomedicina

especialidad: Patología Molecular

C.C. 171299514-9

DECLARACIÓN DEL DIRECTOR CIENTÍFICO

“Declaro haber dirigido científicamente al estudiante para la realización de su trabajo experimental de titulación en base al método científico, conduciéndole con coherencia en el conjunto de experimentos realizados, y orientando sus conocimientos para lograr los objetivos propuestos”.

Carolina Salazar Ruales

Máster en Biotecnología

C.C. 171883609-9

DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE

“Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.”

Estefanía Nathaly Eras Muñoz

C.C. 171914889-0

AGRADECIMIENTOS

Agradezco a Dios por haberme dado las fuerzas necesarias para culminar esta etapa de mi vida, a mi familia por ser el motor que me impulsa día a día.

A la MSc. Salazar y a la Dra. Aguirre por confiar en mis conocimientos y guiarme a través de su experiencia en mi formación como profesional.

A Giss L., Ale S., Jorge A., Freddy N. y Vale F. por compartir juntos la vida universitaria, gracias por su amistad.

DEDICATORIA

Con mucho cariño y esfuerzo dedico este trabajo a mis padres, a mis hermanas y a mi abuelita Marina “mi ángel protector”.

Por ser el pilar fundamental en mi vida, quienes me alientan a luchar por mis sueños y sobre todo me enseñaron que un tropiezo no es caída y que la palabra “no puedo” no existe.

A ustedes, que han estado conmigo en los buenos y malos momentos, gracias por su amor incondicional, consejos, cuidados y sobre todo por apoyarme siempre en mis decisiones.

“Familia, donde la vida empieza y el amor nunca termina”.

RESUMEN

La Corea de Huntington (EH) es una enfermedad neurodegenerativa hereditaria que afecta a todas las poblaciones independientemente de la edad. Su incidencia a nivel mundial se encuentra alrededor de 5-10 pacientes por cada 100.000 habitantes. A pesar de que esta enfermedad fue descrita inicialmente en 1872, en el Ecuador su prevalencia aún no ha sido establecida. Para poder ampliar el panorama sobre la enfermedad es necesario conocer la estructura genética poblacional de los pacientes. Por ello, el objetivo de este estudio fue determinar los niveles preliminares de ancestría en pacientes de la población ecuatoriana con mutación en el gen-IT15. Implementando 46 marcadores AIMS-INDELS y siguiendo el protocolo establecido por Pereira y colaboradores (2012), los resultados mostraron que esta población se agrupa en tres clústeres principales. Los que corresponden al componente ancestral africano (1,8 %), europeo (14,5 %) y en mayor proporción nativo americano (83,7 %). Los estadísticos analizados mostraron que los distintos loci se encuentran en el equilibrio de Hardy-Weinberg y no están ligados entre sí. Mientras que, el estadístico F_{ST} mostró que la población de estudio presenta una diferencia muy elevada comparada con la población africana (0,39), gran diferencia con la población europea (0,21) y una moderada y baja diferenciación en relación a la población nativa americana (0,07) y controles (0,05). En conclusión, el mayor componente ancestral de los pacientes con EH de la población ecuatoriana estudiada proviene del origen nativo americano.

Palabras clave: Corea de Huntington, gen-IT15, marcadores AIMS-InDels, población ecuatoriana, ancestría, estadístico F_{ST} .

ABSTRACT

Huntington's disease (HD) is a neurodegenerative hereditary illness which affects all the populations. HD incidence around the world is proximity to 5-10 patients for each 10 000 habitants. Even though this illness was described in 1872, in Ecuador there are few records about its incidence. In order to complete the information of the illness is necessary to know the genetic structure of the HD patients in the country. That is why the aim of this study was establish the levels of ancestry, using 46 genetic markers AIMs-INDELS, of HD patients in the Ecuadorian population. The outputs showed that all the loci analyzed were in Hardy-Weinberg equilibrium and they were not linked. HD patients can be clustered in three specific groups. The main ancestry proportions corresponded to Americans (83.7 %), European (14.5 %) and African (1.8 %) predecessor. The F_{ST} statistic showed that the studied population had the biggest difference with African population (0.39), with European (0.21) with Native American (0.07) and with controls (0.05). In conclusion, main ancestral component of the HD Ecuadorian patients comes from Native American origin.

Keywords: Huntington's disease, IT-15 gene, AIMs-InDels markers, Ecuadorian population, ancestry, F_{ST} statistics.

ÍNDICE

1.CAPITULO I. INTRODUCCIÓN

1.1 ANTECEDENTES	1
1.2 PLANTEAMIENTO DEL PROBLEMA	3
1.3 OBJETIVOS	5
1.3.1 Objetivo general	5
1.3.2 Objetivos específicos	5
1.4 JUSTIFICACIÓN.....	5

2.CAPITULO II. MARCO TEÓRICO

2.1 ENFERMEDAD DE HUNTINGTON.....	7
2.1.1 Generalidades.....	7
2.1.2 Aspectos Clínicos.....	8
2.1.3 Etapas de la Corea de Huntington	11
2.2 BASES MOLECULARES	12
2.2.1 Gen IT-15.....	12
2.2.2 Proteína Huntingtina	14
2.3 MÉTODOS DE DIAGNÓSTICO.....	16
2.3.1 Análisis según el ligamiento genético	16
2.3.2 Análisis directo mediante técnicas de Biología Molecular	17
2.4 ESTUDIO DE POBLACIONES.....	19
2.4.1 Variabilidad Genética	19
2.4.2 Marcadores autosómicos aplicados en estudios de ancestría	22
2.4.3 Programas empleados para evaluar la ancestría genética en poblaciones humanas	25

3.CAPITULO III. DISEÑO DEL PLAN EXPERIMENTAL

3.1 DISEÑO EXPERIMENTAL	28
-------------------------------	----

4.CAPITULO IV. PROCEDIMIENTOS

4.1 POBLACIÓN Y MUESTRA	30
-------------------------------	----

4.2 MATERIALES Y MÉTODOS	30
4.2.1 Extracción de ADN.....	30
4.2.2 Cuantificación de ADN	31
4.2.3 Reacción en cadena de la polimerasa multiplex (PCR-multiplex)	32
4.2.4 Electroforesis capilar y obtención de frecuencias alélicas.....	33
4.3 EVALUACIÓN ESTADÍSTICA DE LOS RESULTADOS	36
4.3.1 Determinación de la estructura poblacional	36
4.3.2 Procesamiento estadístico de los resultados	37
5. CAPITULO V. RESULTADOS Y DISCUSIÓN	
5.1 Determinación de la estructura poblacional de los pacientes afectados con la enfermedad de Huntington.....	39
5.2 Análisis de la estructura de la población de los pacientes afectados con la enfermedad de Huntington en el contexto de las poblaciones de referencia	45
5.3 Evaluación de parámetros genéticos.....	50
6. CONCLUSIONES Y RECOMENDACIONES	
6.1 Conclusiones	58
6.2 Recomendaciones	59
REFERENCIAS	60
ANEXOS	75

ÍNDICE DE FIGURAS

Figura 1. Representación gráfica de la anomalía a nivel cerebral generada en la Corea de Huntington	7
Figura 2. Contracciones musculares tipo corea	10
Figura 3. Representación de las alteraciones presentes en el gen IT-15.....	13
Figura 4. Interpretación del número de repeticiones del triplete CAG del gen IT-15.....	16
Figura 5. Diagrama de flujo del diseño experimental.	29
Figura 6. Electroferograma de LIZ 600.....	33
Figura 7. Ejemplo del panel de visualización obtenido a través del programa GeneMapper® v4.0.....	35
Figura 8. Distintas interpretaciones de las frecuencias alélicas	36
Figura 9. Determinación del valor más probable de K en la muestra de pacientes afectados por la enfermedad de Huntington	41
Figura 10. Determinación del valor más probable de K en el grupo poblacional control empleado en el estudio.....	41
Figura 11. Determinación del valor más probable para K para el grupo global conformado por la población control y en los pacientes.....	42
Figura 12. Representación gráfica de los niveles de ancestría para los individuos de las muestras poblacionales analizadas..	43
Figura 13. Representación gráfica en el orden de la matriz Q obtenida con el programa STRUCTURE de los niveles de ancestría para los pacientes con Huntington.....	44
Figura 14. Representación triangular de la distribución de los pacientes afectados con Huntington.....	44
Figura 15. Selección del número de clústeres más probable.....	46
Figura 16. Análisis de la estructura de las poblaciones “Controles” y “Pacientes” en el contexto de poblaciones de referencia..	48
Figura 17. Valores de la matriz Q para cada uno de los grupos evaluados en el contexto de las poblaciones de referencia (África, América y Europa)	49

Figura 18. Prueba de Hardy-Weinberg.....	52
Figura 19. Prueba de Student para los valores de F_{IS}	53
Figura 20. Correlación entre los F_{IS} determinados por los métodos W&R y R&H.....	54
Figura 21. Evaluación de la diferenciación entre poblaciones mediante el estadístico F_{ST} de Wright.....	55
Figura 22. Valores de 1-Qintra y 1-Qinter.....	57
Figura 23. Diferenciación génica de los individuos de las muestras poblacionales conformadas por los pacientes afectados con la enfermedad de Huntington y los controles.	57

CAPITULO I. INTRODUCCIÓN

1.1 ANTECEDENTES

En el año 1872 se describió por primera vez la enfermedad de Huntington mediante el artículo denominado “*On Chorea*” publicado por el Doctor George Huntington. En esta investigación se pudo determinar que esta es una enfermedad hereditaria, la cual puede provocar en las personas que la padecen demencia y desórdenes neurológicos. Por otra parte, esta enfermedad progresa y agudiza su sintomatología con el transcurso del tiempo (Huntington, 1872).

Con esta información base proveída por Huntington y con el avance de las investigaciones posteriores se determinó que esta enfermedad es un trastorno neurodegenerativo hereditario progresivo, autosómico dominante, que se caracteriza por alteraciones neuropsiquiátricas, cognitivas y psicomotoras que van en aumento. Una de las principales características clínicas de esta enfermedad son los movimientos anormales involuntarios también conocidos como espasmos musculares tipo corea, de ahí el nombre de Corea de Huntington (Kirkwood, Su, Conneally, y Foroud, 2001, p. 273–278).

La Corea de Huntington es una de las 30 enfermedades genéticas hereditarias que se encuentran relacionadas directamente con el incremento en el número de repeticiones de un segmento del ADN (Mirkin, 2007). Se ha observado que el gen IT-15 (de sus siglas en inglés *interesting transcript 15*), conocido en la actualidad como Huntingtina (HTT), posee una mutación dinámica basada en el aumento en las repeticiones del triplete CAG que codifican para la proteína citoplasmática del mismo nombre. La proteína Huntingtina está relacionada con el transporte de electrones a nivel mitocondrial y participa en la protección neuronal y neurotransmisión. Por otra parte, participa en la regulación

transcripcional, plegamiento y degradación proteica (Aparicio y Lucas, 2002; Tasset, Sánchez, y Túnez, 2009, p. 424–429).

El triplete expandido es inestable, razón por la cual al dividirse las células el tamaño de la secuencia varía tanto en la línea germinal como en las células somáticas (Cuenca y Morales, 1999). El gen que codifica para la Huntingtina en personas normales presenta una repetición del triplete CAG que oscila entre 9 y 36 veces, mientras que en personas diagnosticadas con la enfermedad, el triplete se puede expandir hasta 250 veces. Esto demuestra que existe una relación directamente proporcional entre la enfermedad y las repeticiones del triplete, a su vez se ha observado, que mientras mayor es el número de repeticiones, más temprana es la edad a la que inicia la sintomatología (Fernandes, 2001; Raymund, 2010; Townhill et al., 2016).

La primera manifestación de esta enfermedad se genera a una edad promedio que oscila entre los 30 y 50 años de edad. Independientemente del sexo de la persona, se ha establecido que el tiempo de vida promedio oscila entre 15 y 20 años desde el inicio de la manifestación de la sintomatología hasta el deceso; lo cual resulta alarmante (Bates, Tabrizi y Jones, 2014, p. 28).

Estudios preliminares sobre la estructura genética poblacional de la enfermedad de Huntington ha mostrado que esta tiene múltiples orígenes ancestrales (Andrew et al., 1993). Inicialmente se estableció su origen en individuos de etnia africana, a pesar de ello, las estadísticas contradicen estos resultados ya que su incidencia en este continente es de 0,6 casos por millón de personas. Por el contrario, se ha determinado que la prevalencia más alta de la enfermedad se encuentra en Europa y en países cuyo origen es europeo con una incidencia de 40-100 pacientes por millón de personas. Basándose en estos resultados se ha establecido la hipótesis de que el origen de la

enfermedad de Huntington de África y de Europa son distintos (Krause et al., 2015; Liou, 2010; Warby et al., 2011).

A nivel de América Latina la prevalencia de la enfermedad es variable. Entre los países con mayor incidencia se encuentra Venezuela con 1 paciente por cada 23.000 habitantes y Perú con 1 paciente por cada 143 habitantes (Raymund, 2010; Warby et al., 2011). Si las tasas de incidencia de la EH a nivel mundial se extrapolan a la población ecuatoriana aproximadamente podrían existir 1.500 casos en total, dato que resulta muy preocupante (Paz-y-Miño, Pavón-Realpe, Jaramillo, López y Domínguez, 2014).

En un inicio la EH era detectada en base al ligamiento genético lo cual la convertía en una enfermedad subdiagnosticada. Sin embargo, en la actualidad existen distintos métodos moleculares como la PCR (de sus siglas en inglés *Polimerase Chain Reaction*) y qPCR (de sus siglas en inglés *Real Time Polimerase Chain Reaction*) que han sido implementados con el fin de proveer un diagnóstico temprano de la enfermedad. A pesar de estos avances, al ser esta una enfermedad hereditaria su diagnóstico contribuye únicamente al tratamiento de su sintomatología más no con la cura (Rosales & Barros, 2008).

1.2 PLANTEAMIENTO DEL PROBLEMA

En la actualidad en nuestro país se está iniciando con la investigación de enfermedades neurodegenerativas a nivel molecular, ya que como se sabe los diagnósticos clínicos generalmente se basan en la sintomatología que presenta el paciente. Actualmente, existen proyectos y datos que sientan las bases para el desarrollo de nuevas investigaciones en este campo. El grupo de investigación de Paz-y-Miño et al., (2014) ha sido el pionero en proveer datos informativos sobre la Corea de Huntington en el Ecuador; sin embargo, esta información necesita ser complementada con estudios de perfiles de expresión génica como es el análisis ancestral.

El Ecuador es un país multiétnico y pluricultural, conformado en su gran mayoría por individuos mestizos que representan el 77,4 % de la población ecuatoriana total (Ferreira, García, Macías, Pérez y Tomsich, 2013). Basándose en la reseña histórica de la época de la colonia el origen de la población ecuatoriana proviene de la mezcla entre indígenas y españoles. Por lo que, la incidencia de la enfermedad en el país podría ser aproximada a la de la población caucásica (Paz-y-Miño y Burgos, 2015, p.29). En la actualidad existen estudios sobre la enfermedad de Huntington en poblaciones aisladas de países como Brasil, Argentina, Perú, Chile, entre otros, donde se ha visto una alta incidencia de la enfermedad debido al aislamiento y al mestizaje de estos grupos sin embargo, en el Ecuador no se reportan datos globales sobre la misma.

Como se mencionó anteriormente, esta enfermedad se presenta en su etapa inicial, por lo general, entre los 30 a 40 años de edad cuando la población es aun relativamente joven. Sin embargo, se han obtenido datos estadísticos que demuestran que el 10 % de los casos inician a los 20 años, el 5 % antes de los 14 años y el 1 % en la infancia antes de los 10 años. Dentro de todo este abanico de casos se ha encontrado que entre el 70 y el 80 % de los mismos tienen un patrón de herencia paterno (Martínez y Rában, 2002; Molón, Yáñez y López-Ibor, 2010).

Aunque ya se han establecido estadísticas aproximadas sobre la incidencia de la enfermedad, los valores de edad y porcentajes pueden cambiar según el incremento de las repeticiones del triplete CAG. A pesar de que existen distintos métodos de diagnóstico para la enfermedad estos no son implementados en nuestro país debido a que son poco asequibles. Por lo que el desarrollo de esta enfermedad en el país implica una carrera contra el tiempo.

1.3 OBJETIVOS

1.3.1 Objetivo general

Determinar los niveles preliminares de ancestría en pacientes con mutación en el gen IT-15 de la población ecuatoriana.

1.3.2 Objetivos específicos

- Analizar las frecuencias alélicas en pacientes con Corea de Huntington de la población ecuatoriana mediante técnicas de biología molecular y secuenciación tipo Sanger.
- Correlacionar las frecuencias alélicas de pacientes ecuatorianos con Corea de Huntington con los datos obtenidos de los tres grupos parentales europeos, africanos y nativos americanos.

1.4 JUSTIFICACIÓN

Por todo lo antes expuesto, existe la necesidad de desarrollar nuevos métodos de diagnóstico más eficaces para la detección temprana de esta patología y complementar los existentes con información ancestral. Las tecnologías de última generación mediante el estudio de perfiles de expresión genética como el análisis de ancestría parece ser una alternativa prometedora (Fumagalli et al., 2011, p. 124–127).

La implementación de marcadores informativos de ancestría (AIMs-INDELs) permiten identificar frecuencias alélicas entre distintas poblaciones para posteriormente inferir su origen ancestral (Santos et al., 2010; Urbano et al., 2015). Los marcadores AIMs-INDELs identifican regiones genómicas que presentan una relación con la enfermedad de estudio, teniendo como fin ampliar el panorama de predicción e identificación de dicha patología (Tang,

Coram, Wang, Zhu y Risch, 2006, p. 1–12). Debido a que el Ecuador es un país multi-diverso posee frecuencias alélicas diferentes a otras poblaciones, esto justifica la caracterización de la población ecuatoriana (Paz-y-Miño et al., 2014, p.32-44).

En los últimos años se ha intentado introducir técnicas moleculares en el estudio de esta patología con el fin de establecer nuevos métodos de detección más eficaces, no invasivos, confiables y específicos para la población ecuatoriana. Es por ello que esta investigación se centra en determinar el genotipo de pacientes ecuatorianos con Corea de Huntington, con el fin de evaluar las frecuencias alélicas de esta población mediante el análisis de 46 marcadores moleculares de ancestría (AIMs-INDELS) para su futura implementación en técnicas de diagnóstico.

Este trabajo de titulación aportaría en el estudio de la Enfermedad de Huntington en el Ecuador generando información pionera sobre el análisis ancestral genético de esta patología. Los resultados aportados por este trabajo fomentan el desarrollo de la investigación científica en el campo de la genética poblacional y la biomedicina. Además, la información generada podrá ser empleada como referente para futuras investigaciones.

CAPITULO II. MARCO TEÓRICO

2.1 ENFERMEDAD DE HUNTINGTON

2.1.1 Generalidades

La enfermedad de Huntington (EH) que fue descrita por primera vez en el año 1872 es considerada una patología fatal hasta la actualidad. Esta enfermedad se encuentra dentro de las denominadas “Coreas” debido a que son patologías con sintomatologías parecidas, que se caracterizan por provocar trastornos de tipo neurológico que impiden a las personas que lo padecen desarrollar actividades cotidianas (Faria et al., 2016; Huntington, 1872, p. 317–321; López, 2015, p. 239; Yhnell, Lelos, Dunnett, y Brooks, 2016). La EH está relacionada directamente con el sistema nervioso central, afectando principalmente a determinadas zonas del cerebro como la corteza cerebral y el estriado (Figura1) (González, Sánchez, González, Gascón y Martínez, 2007; Tasset et al., 2009; Vásquez et al., 2011).



Figura 1. Representación gráfica de la anomalía a nivel cerebral.
Derecha: control de una persona sana.
Izquierda: paciente diagnosticado con Huntington.
Tomado de Martínez et al., 2002, p.39.

Al complementar los estudios iniciales reportados por el Doctor George Huntington (1872), se estableció que la EH es un trastorno neurodegenerativo hereditario, autosómico dominante y progresivo. Esta enfermedad es causada por una mutación dinámica cerca del extremo 5' en el exón 1 del cromosoma 4 del gen *IT-15* también conocido como *Huntingtina* (HTT), cuya ubicación en el genoma es 4p16.3 (Paz-y-Miño et al., 2014; Tasset et al., 2009). Esta mutación comprende la expansión del triplete citosina-adenina-guanina (CAG) en el gen que codifica para la proteína citoplasmática del mismo nombre, que se expresa a altos niveles en el cerebro, aunque también se encuentra en otros tejidos como testículos, ovarios, pulmón, hígado e intestino (Aparicio y Lucas, 2002).

La EH al ser una enfermedad hereditaria es conocida como "*La enfermedad de la familia*" esto se debe a que las estadísticas muestran que cada hijo de un padre con EH tiene la posibilidad del 50 % de adquirir el gen defectuoso o mutado por lo que es considerada una enfermedad de herencia paterna (HDSA, 2016).

2.1.2 Aspectos Clínicos

Debido a que el diagnóstico de los pacientes con Corea de Huntington se ha basado en los síntomas, se han identificado tres principales tipos de trastornos: psicomotores, cognitivos y emocionales que agudizan sus manifestaciones gradualmente. A pesar de que los tres trastornos generan una gran afectación en las personas que padecen esta enfermedad los trastornos cognitivos y emocionales son los más debilitantes (Townhill et al., 2016).

En etapas avanzadas de la enfermedad, las personas pierden la habilidad de caminar, hablar y comer requiriendo personal especializado para desarrollar sus actividades. Muchos científicos describen a la Corea de Huntington como una enfermedad que agrupa los síntomas del Parkinson, Alzheimer y la

esclerosis lateral amiotrofia (ELA) simultáneamente (HDSA, 2016; Salamanca, 2003).

La EH es una enfermedad cuya edad media de inicio de la sintomatología se encuentra alrededor de los 35 años, sin embargo, se han establecido tres tipologías clínicas. La primera es la clásica e inicia alrededor de los 35 y 40 años de edad (Vásquez y Montero, 2008). La segunda es conocida como la juvenil, esta marca el inicio de la sintomatología alrededor de los 21 años de edad. Finalmente, la tipología senil, que se manifiesta después de los 55 años de edad y posee una evolución de la sintomatología alarmantemente rápida (Molón et al., 2010; Sobrino et al., 2007, p. 1–8).

2.1.2.1 Trastornos Psicomotores

Estas alteraciones se caracterizan por movimientos anormales llamativos que en la mayoría de los casos son involuntarios, tales como los de tipo corea y distonía que son clásicos de la EH. Los movimientos tipo corea son rápidos, provienen de la palabra griega que significa “*baile o coreografía*” debido al efecto de movimiento que causan en las personas (Figura 2). Se ha observado que estos movimientos pueden afectar a cualquier parte del cuerpo siendo las zonas más afectadas los miembros inferiores y el rostro (Serna, 2014, p. 31–35; Vásquez et al., 2011).

Los espasmos musculares o distonía, al contrario de los movimientos tipos corea son contracciones involuntarias lentas que generan en la persona torpeza, lentitud en los reflejos, movimientos contraídos, movimientos anormales oculares y rigidez. Ambas alteraciones son visibles por lo general al andar (Townhill et al., 2016).



Figura 2. Contracciones musculares tipo corea generadas en la Corea de Huntington.
Tomado de Serna, 2014.

2.1.2.2 Desórdenes Cognitivos

Los desórdenes de este tipo han sido ampliamente estudiados, la primera publicación sobre este tema fue realizada en 1974 por el equipo de Boll y sus colaboradores. En este estudio se identificaron una serie de cambios en el rendimiento cognitivo al comparar personas sanas y pacientes. Este trastorno va de la mano con el deterioro psicomotor. A pesar de ello, su relación es variable, un paciente puede presentar un considerable daño motor con un bajo daño cognitivo y viceversa (Boll, Heaton y Reitan, 1974, p. 61–69).

Debido a la gran cantidad de procesos cognitivos propios del ser humano es difícil caracterizar el deterioro cognitivo específico en la EH (Stout, Glikmann-Johnston y Andrews, 2016, p. 286–294). Síntomas como la pérdida de peso y sobre todo la alteración del sueño son características que se ha observado, fomentan el deterioro cognitivo, el que puede ser difícil de controlar y manejar.

En la EH el trastorno del sueño ocurre en una fase temprana y en enfermedades neurodegenerativas como Parkinson y Alzheimer está directamente relacionado con este tipo de deterioro. Por lo que podría ser característico de estas enfermedades. A diferencia del Alzheimer las personas con EH no presentan problemas en la acumulación de recuerdos, más bien tienen dificultad para recordar (Lazar et al., 2015, p. 635; Martínez y Rában, 2002).

2.1.2.3 Trastornos Emocionales

Los trastornos emocionales son los más debilitantes tanto para el paciente como para su familia. Este tipo de alteración se presenta frecuentemente en pacientes adultos que refieren problemas como irritabilidad, agresividad, depresión, acciones compulsivas entre otras perturbaciones de tipo psiquiátricas además de los problemas psicomotores y cognitivos (Tasset et al., 2009).

El Doctor Huntington en su publicación pionera sobre la EH llegó a la conclusión de que este trastorno era uno de los que más afectaban a los pacientes y según investigaciones realizadas posteriormente, la tasa de suicidios en personas con Corea de Huntington es elevada, lo que confirma la conclusión a la que se arribó en el año 1872 (Hughes y Lo, 2010, p. 1-3).

2.1.3 Etapas de la Corea de Huntington

Según la Huntington's Disease Society of América (HDSA, 2016), a pesar de que la sintomatología y las manifestaciones de la enfermedad varían tanto de persona a persona, se ha logrado establecer tres etapas de la enfermedad:

2.1.3.1 Etapa Inicial

En esta etapa se empiezan a evidenciar cambios leves en la coordinación, los movimientos involuntarios son escasos y se pueden percibir dificultades para

razonar y resolver problemas. El estado de ánimo de los pacientes es por lo general irritable o deprimido pero puede ser controlado con medicamentos. A nivel social, la persona disminuye su eficiencia convirtiéndose en menos funcional.

2.1.3.2 Etapa Intermedia

El trastorno psicomotor es el más evidente en esta etapa, los medicamentos utilizados logran aliviar y disminuir los movimientos involuntarios, sin embargo, estos incrementan su intensidad. Los problemas cognitivos generan la dificultad del habla y del deglutir. Además, la pérdida de las habilidades propias del ser humano como el razonar y el pensar impiden casi en su totalidad desarrollar las actividades cotidianas.

2.1.3.3 Etapa Tardía

En esta etapa la persona que padece Huntington depende completamente de otros ya que los trastornos y alteraciones antes mencionadas se agudizan. Existe la imposibilidad de caminar, hablar y comer, lo cual afecta drásticamente y aumenta la manifestación de los trastornos emocionales. Los movimientos tipo corea en algunos casos aumenta su intensidad, mientras que en otros cesa completamente. Los pacientes en esta etapa presentan complicaciones como asfixia e infecciones que son preocupantes ya que muchas veces las personas con EH mueren por complicaciones asociadas a la enfermedad más no por la enfermedad en sí misma.

2.2 BASES MOLECULARES

2.2.1 Gen IT-15

El gen IT-15 o HTT tiene una longitud de 210 kb., con un ARNm de 10 366 bp. que presenta un marco de lectura abierto de 9 432 bp. Este gen codifica para la proteína Huntingtina, conformada por 3 144 aminoácidos cuyo peso molecular aproximado es de 348 kDa. (Peñuelasa, Miguela y Gullón, 1997; Tasset et al., 2009). La proteína a partir del residuo 17 posee un segmento “*poliQ*”, es decir

un segmento de poliglutaminas seguidas de varias repeticiones del aminoácido prolina cercanas al extremo N-terminal (Cattaneo, Zuccato y Tartari, 2005; Hughes y Lo, 2010, p.55-57).

Dentro del fragmento génico que codifica para el segmento poliQ de la proteína se encuentra el triplete CAG. Este triplete tiene un número variable de repeticiones siendo entre 9 y 36 el rango normal en personas sanas, mientras que en pacientes puede llegar hasta a 250 veces; generando problemas neurológicos principalmente por la elevada expresión de la proteína a nivel cerebral (Figura 3) (Tasset et al., 2009; Vázquez et al., 2013).

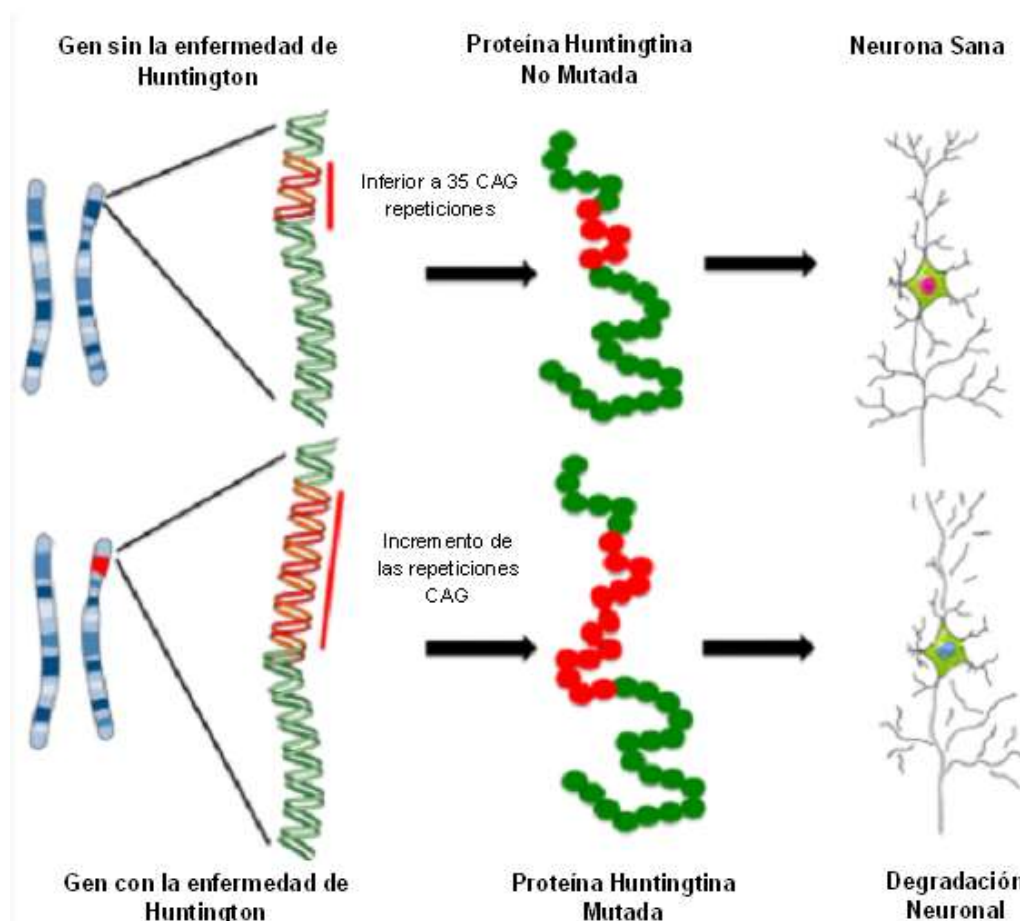


Figura 3. Representación de las alteraciones generadas en el gen IT-15 en la Enfermedad de Huntington. Al incrementarse las repeticiones del triplete CAG en el exón 1 del gen IT-15 se obtiene una proteína mutada que produce alteraciones a nivel neurológico. Adaptado de Eurostemcell, 2013.

2.2.2 Proteína Huntingtina

La Huntingtina (HTT) es una proteína que se expresa en los mamíferos y principalmente en los humanos. Se encuentra en varios tipos celulares presentando altas concentraciones en los testículos y cerebro y concentraciones moderadas en órganos como el hígado, pulmón y corazón. Existen varias proteínas ortólogas presentes en organismos como *Drosophila Melanogaster*, el Pez cebra (*Danio rerio*) y *Caenorhabditis elegans* donde se ha determinado que la HTT se expresa constitutivamente y forma parte del complejo dinactina encargado del transporte intracelular (Rosales y Barros, 2008).

Las primeras investigaciones realizadas en murinos, mostraron que la Huntingtina es una proteína crítica durante el desarrollo embrionario. Estos estudios también permitieron determinar que la mutación en homocigosis en los primeros estadios del desarrollo es letal y en heterocigosis genera trastornos cognitivos principalmente (Peñuelasa et al., 1997).

Cuando la proteína Huntingtina esta mutada se incrementa su expresión a nivel nuclear, lo que aumenta su toxicidad como resultado de las alteraciones generadas en el proceso de transcripción. Estudios realizados por Luthi-Carter et al., (2002), en ratones con EH mostraron cambios a nivel del cerebelo que aparentemente eran dependientes de la expansión del segmento poliQ.

2.2.2.1 Funciones de la Huntingtina

Su función específica aún no ha sido comprendida en su totalidad ya que presenta escasa homología con otras proteínas, sin embargo se conoce que se localiza en todos los compartimientos a nivel subcelular donde se asocia a organelos como el retículo endoplasmático, el aparato de Golgi y la mitocondria (Hayden y Fremer, 2000, p. 5678). Entre sus principales funciones se

encuentra el transporte vesicular mediante la interacción con la tubulina- β , participa también en el proceso de endocitosis mediada por clatrina. Pero sobre todo tiene un papel fundamental en la sinapsis y el transporte neuronal ya que protege a las neuronas de la apoptosis causada por el estrés oxidativo (Landles y Bates, 2004).

Otras funciones de la HTT están relacionadas con la regulación transcripcional, degradación y plegamiento proteico. Principalmente, se cree que al interactuar la HTT mutada (mhtt) con las chaperonas Hsp 70 y Hsp 40 genera una proteína mal plegada que podría desencadenar la enfermedad de Huntington (Bates, 2003, p.1642–1644).

2.2.3 Triplete CAG

Como ya se mencionó anteriormente, existe una relación directamente proporcional entre el número de tripletes CAG y la enfermedad de Huntington. Por lo tanto, el incremento en el número de repeticiones determina la presencia o ausencia de la enfermedad y a su vez la gravedad de la misma.

2.2.3.1 Interpretación de las repeticiones del triplete CAG

Basándose en el número de repeticiones de CAG se han realizado estudios para establecer un rango aproximado que permita diagnosticar la enfermedad empleando métodos moleculares. Los alelos que poseen hasta 28 repeticiones son considerados normales, de 29 a 35 repeticiones son alelos mutantes que no producen sintomatología; sin embargo, existe la posibilidad de expansión del triplete en la descendencia (Paz-y-Miño et al., 2014).

De 36 a 39 repeticiones de CAG se considera que existe una penetrancia incompleta, la cual provoca una inestabilidad meiótica que puede reflejarse en la progenie, es decir, existe la posibilidad de que la descendencia presente la enfermedad con una sintomatología tardía. Si las repeticiones superan estos

valores oscilando entre 40 o más, existe la certeza total del padecimiento de la enfermedad con transmisión a la descendencia (Figura 4) (Aronin et al., 1995, p. 1193–1195; Margolis, 2003, p. 1726–1732).

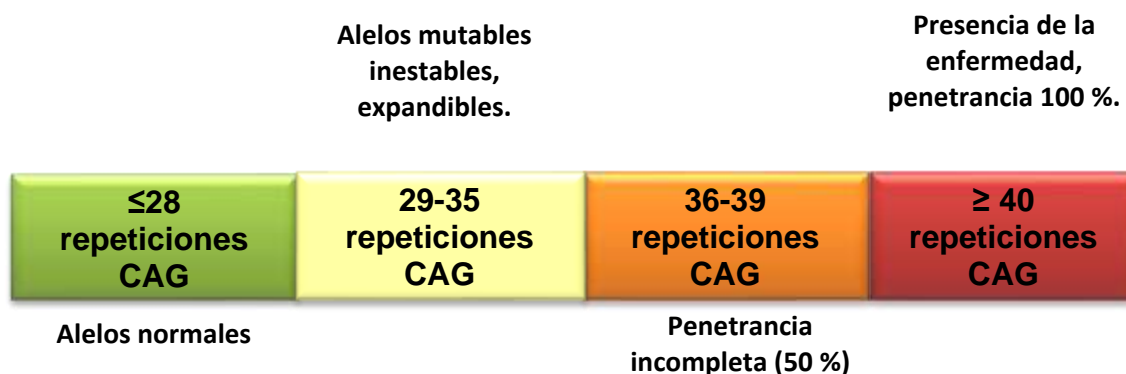


Figura 4. Interpretación del número de repeticiones del triplete CAG del gen IT-15.

Adaptado de Ramos et al., 2012,p.585.

2.3 MÉTODOS DE DIAGNÓSTICO

La EH no tiene cura, por lo que únicamente se pueden utilizar medicamentos para disminuir los síntomas generados por la misma. Hace un tiempo atrás la determinación de esta enfermedad era realizada únicamente cuando se expresaban los síntomas.

2.3.1 Análisis según el ligamiento genético

2.3.1.1 Test según el Historial Familiar

Un método que se utilizaba es el del árbol familiar, este debía contener toda la información detallada del paciente y sus familiares, más aún si existía uno de los padres con Huntington. A pesar de que esta era una técnica muy útil hace mucho tiempo atrás, tenía un rango de error elevado; ya que existen enfermedades denominadas “*fenocopias*” que presentan similar sintomatología pero su origen genético es distinto, lo cual puede generar un falso positivo (Rosales y Barros, 2008).

2.3.1.2 Test de exclusión

Básicamente este análisis se centra en un seguimiento familiar, principalmente se lo realiza para diagnósticos de tipo prenatal o preimplantacional cuando el progenitor desea saber si su descendencia será sana. Consiste en determinar si el feto o embrión son portadores de la enfermedad o si ha heredado el haplotipo del abuelo lo que permite considerarlo como paciente sano. Hay que tener en cuenta que este análisis posee varias implicaciones a nivel bioético, ya que según datos bibliográficos existe la posibilidad del 50 % de heredar la enfermedad (Ramos et al., 2012).

2.3.2 Análisis directo mediante técnicas de Biología Molecular

Con el avance de la ciencia y gracias a la Biología Molecular, hoy en día se pueden llevar a cabo análisis de tipo genético únicamente para diagnóstico. La denominada "*prueba de oro*" consiste en la extracción de sangre periférica para posteriormente obtener el ADN de la persona y analizarlo. El fin de este avance es detectar la enfermedad antes de que se inicie la sintomatología y avance a etapas de mayor gravedad. Sin embargo, es una técnica poco accesible ya que posteriormente se debe realizar un proceso de secuenciación para validar los resultados.

2.3.2.1 Reacción en Cadena de la Polimerasa (PCR)

Una técnica muy empleada para el diagnóstico es la reacción en cadena de la polimerasa (PCR), esta fue utilizada inicialmente para determinar el número de repeticiones del triplete CAG. Aunque existen varias publicaciones que respaldan su aplicación resulta poco eficiente cuando existe un número de repeticiones elevado. Esto se debe a que la región mutada es altamente rica en citosina y guanina las cuales se aparean dificultando la amplificación del fragmento por este método (Panagopoulos, Lassen, Kristoffersson y Aman, 1999, p.232–236).

Existen compuestos químicos como el dimetilsulfoxido, el glicerol o la formamida que son utilizados para incrementar la eficiencia de este proceso. Estos impiden la formación de estructuras secundarias formadas por el apareamiento de C+G durante la amplificación. Sin embargo, estos dificultan la visualización de los amplicones. Por esta razón, algunos autores recomiendan utilizar iniciadores marcados radiactivamente o acoplar iniciadores fluorescentes (dNTPs marcados) a la técnica con el fin de mejorar la amplificación. Para obtener el producto final y visualizarlo posteriormente por electroforesis capilar en el secuenciador (Rosales y Barros, 2008).

2.3.2.2 Real Time PCR (qPCR)

Debido a la poca sensibilidad de la PCR convencional se ha establecido el uso de su variante la qPCR (de sus siglas en ingles *Real Time Polimerase Chain Reaction*). Esta es mucho más sensible, pues es una técnica cuantitativa que registrar la amplificación de cada ciclo asegurando la especificidad del producto al emplear la temperatura de fusión específica (T_m). Finalmente, la qPCR permite diferenciar cada producto de la reacción en función de la temperatura específica de fusión (T_m) por lo tanto, mientras aumenta el número de repeticiones CAG aumenta el valor de T_m convirtiéndola en una técnica útil para el diagnóstico de EH (Mahjoubi, Montazeri, Zare-kahrizi, Nafisi y Engineering, 2013).

2.3.2.3 Southern Blot

A pesar de que esta técnica es poco empleada para el análisis de esta enfermedad es útil cuando existe un gran número de repeticiones del triplete CAG. Sin embargo, es más una herramienta que puede ser utilizada como confirmatoria de diagnóstico cuando se presenta un genotipo normal en homocigosis (Ramos et al., 2012).

2.4 ESTUDIO DE POBLACIONES

Al identificar la variabilidad genética que existe entre individuos y a su vez entre poblaciones se puede determinar la mayor o menor susceptibilidad de un grupo étnico a una enfermedad dada. Por otra parte, se pueden realizar estudios a nivel farmacológico con el fin de entender la variabilidad en el efecto de los fármacos y su relación con factores genéticos.

2.4.1 Variabilidad Genética

Una población de interés en el estudio de la variabilidad genética humana es la africana. Esta población se ha visto expuesta a cambios extremos de tipo climático, alimenticio y sobre todo de procesos infecciosos que traen como consecuencia variaciones tanto genéticas como fenotípicas en sus individuos. Según estudios realizados, las características anatómicas de la humanidad moderna se originan en África, a consecuencia de la migración que existió hace aproximadamente 100.000 años atrás (Campbell y Tishkoff, 2008).

Las otras poblaciones a nivel mundial se encuentran distribuidas alrededor del continente americano, europeo, asiático, antártico y oceánico. Estas son poblaciones de procedencia “*no-africana*” cuyos individuos en su mayoría tienen como origen una mezcla de grupos ancestrales de nativos americanos, caucásicos, indígenas y afrodescendientes. Al analizar a la población humana dentro del viejo continente se ha determinado que existe una variabilidad genética que oscila entre el 85 al 90 %. Sin embargo, al analizar entre los continentes África, Asia y Europa la variabilidad que existe entre ellos es entre el 10 al 15 %. Esta variación puede ser consecuencia de distintos tipos de marcadores autosómicos expresados en las mismas (Jorde, 2001).

Los niveles de variación genética expresados mediante heterocigosidad indican que mientras más distancia geográfica exista de la población Africana esta disminuye notablemente; convirtiéndose en lineales las frecuencias alélicas si aumenta la distancia geográfica (DeGiorgio, Jakobsson y Rosenberg, 2009, p. 16057–16062). Esto se debe a que los alelos derivados son más frecuentes en poblaciones que presentan una distancia considerable de África, lo que indica que existe una relación proporcional entre la ubicación geográfica y las frecuencias alélicas (Ramachandran et al., 2005, p. 15942–15947; Rosenberg, 2002).

Centrándonos en las poblaciones latinoamericanas se puede evidenciar que sus individuos son genéticamente heterogéneos a causa de la mezcla de las etnias caucásica-indígena y caucásica-africana. Esta diversidad genética fomenta la pérdida de la población nativa debido a la integración de nuevas etnias como la caucásica y africana que junto con su origen nativo americano marcan la descendencia ancestral de estas poblaciones (Mulligan, Hunley, Cole, y Long, 2004; Salzano y Cátira, 2003).

En el Ecuador el grupo étnico más representativo es el mestizo con un 77.4 %. Esta etnia proviene de la mezcla generada en la época colonial entre caucásicos con indígenas nativos (Aguirre, 2012, p. 165). Sin embargo, debido a la esclavitud individuos nativos africanos también aportaron acervo genético. Por lo tanto, los tres grupos étnicos principales presentes en el país son de ascendencia caucásica, indígena y africana (Ayala, 2008, p.16-20; Li et al., 2008, p. 1100–1104).

2.4.1.1 Marcadores Uniparentales

Los marcadores genéticos utilizados para este tipo de análisis son el ADN mitocondrial (ADNmt) y el cromosoma sexual Y (Martínez, Sánchez y Martínez, 2010, p. 520–525) Estos marcadores permiten determinar la descendencia

ancestral de un individuo o población ya que al ser de herencia materna y paterna respectivamente permiten rastrear linajes susceptibles a cambios en su estructura (Rondón y Barreto, 2013).

El estudio del cromosoma sexual Y provee información útil cuando se desea realizar un estudio hereditario de padre a hijo, ya que este es propio del sexo masculino. Existe una región específica en el cromosoma Y que no se recombina, dando como resultado la transmisión sin mutaciones de sus loci a la descendencia de generación en generación (Martínez, 2011; Penacino, 2004).

Por otra parte, el ADNmt provee una visión desde la herencia materna. Esto significa que al no existir procesos de recombinación en este ADN y siempre que no exista una mutación que lo altere, el ADNmt de los hijos va a ser idéntico al de la madre. Convirtiéndolo en una herramienta potencial para estudios de identificación, filogenética de poblaciones entre otros, ya que se puede determinar directamente la ancestría materna de un individuo (Crespillo, Bañón y Valverde, 2011, p. 17–21; Martínez, 2011).

Según Merriwether (2003), las frecuencias alélicas más altas del cromosoma Y corresponden a una herencia caucásica y se encuentran en la población mestiza, mientras que las poblaciones africanas y amerindias presentan una frecuencia alélica alta de ADNmt si se comparan con las frecuencias alélicas del cromosoma Y. A pesar de que estos marcadores proveen información útil para los estudios de filogenia son muy generales y proporcionan una visión sesgada de la realidad biológica, por lo que se han desarrollado otros métodos de análisis.

2.4.1.2 Marcadores autosómicos

Estos marcadores permiten determinar la composición genética de una población mediante el análisis de regiones específicas del ADN autosómico (Crespillo et al., 2011). Principalmente, facilitan información ancestral de una población tomando en cuenta la ubicación y adaptación demográfica. Mediante su aplicación se pueden analizar genes defectuosos responsables de una u otra patología (Jarve, 2012, p.13-14).

El análisis consiste en seleccionar un tipo determinado de marcadores que se adapten a la población objeto de investigación estableciendo parámetros obtenidos a partir de la información evolutiva ancestral. El fin es obtener varias sub-estructuras étnicas dentro de una población cuyos caracteres aparentemente son homogéneos (Campbell y Tishkoff, 2008; Plagnol y Wall, 2006; Rourke y Raff, 2010; Stefflova et al., 2011).

Estudios realizados por el grupo de Martínez y colaboradores (2002), donde se analizaron 8 marcadores autosómicos bialélicos en la población colombiana concluyeron que estos marcadores son ideales para identificar la estructura genética de las poblaciones de Sudamérica, principalmente de poblaciones afroamericanas asentadas en estas regiones.

2.4.2 Marcadores autosómicos aplicados en estudios de ancestría

2.4.2.1 Ancestría

El estudio de la estructura genética permite determinar el origen ancestral de una determinada población o individuo mediante la obtención de frecuencias alélicas de la muestra. Los estudios genéticos proporcionan patrones de mezcla de una población lo cual puede ayudar a comprender su historia biológica (Enoch, 2006, p. 19–26; Rojas et al., 2010).

Además de ello, este tipo de análisis facilita la identificación de genes asociados a enfermedades lo que permite predecir la incidencia de una determinada enfermedad. Aportando a su vez, información para el desarrollo y mejora de las técnicas de diagnóstico y tratamiento de las mismas (Mulligan et al., 2004; Väli, Brandström, Johansson y Ellegren, 2008).

2.4.2.2 Polimorfismos de nucleótidos simples (SNPs)

Los polimorfismos de un nucleótido son consecuencia de una mutación puntual generada en la secuencia de ADN, su ubicación puede estar tanto en regiones codificantes como en no-codificantes. Estos cambios son muy frecuentes y aparecen aproximadamente cada 600-1 000 pb., además presentan una tasa de incidencia del 90 % en el genoma humano (Kwok y Chen, 2003). Debido a que estos SNPs no cambian de generación en generación son apropiados para el estudio de análisis de variabilidad genética, niveles ancestrales y presencia de enfermedades genéticas (Guerra y Yu, 2006; Rondón y Barreto, 2013).

Con respecto a la población ecuatoriana, el grupo de Cabrera et al., (2016), llevo a cabo un estudio sobre la intolerancia a la lactosa (LP). El objetivo de esta investigación fue relacionar los grados de intolerancia con los niveles ancestrales de la población estudiada. Los resultados mostraron que el polimorfismo que está directamente relacionado con la sobreexpresión del gen de la lactosa es predominante en el alelo *T-13910*. Mientras que, el análisis ancestral mostró que este polimorfismo presente en los ecuatorianos tiene sus raíces en la población europea (caucásica) como consecuencia de la migración y la colonización española en América del Sur.

2.4.2.3 Polimorfismos en la longitud de fragmentos amplificados (AFLP)

Estos marcadores genéticos son obtenidos mediante la amplificación por PCR de un fragmento de ADN donde se utilizan enzimas de restricción aprovechando los patrones de restricción del ADN genómico. Este método permite generar varios cientos de marcadores altamente reproducibles a partir del material genético de cualquier organismo sin requerir información previa del genoma de estudio (Vekemans, Beauwens, Lemaire, y Roldan-Ruiz, 2002).

Debido a la eficiencia, resolución y a la alta tasa de reproducibilidad que este método posee, se han convertido en un importante tipo de marcador genético ampliamente empleado para el estudio de genética de poblaciones y para el mapeo de loci de rasgos cuantitativos (QTL) (Ferriol, Picó, y Nuez, 2003, p. 275; Mueller y Wolfenbarger, 1999). Su amplia aplicación se debe fundamentalmente, a la rápida y eficiente generación de datos por la alta capacidad de la técnica para localizar secuencias funcionales. Específicamente han sido empleados para detectar la pérdida de la diversidad genética en distintas especies (Song, Zhang, Han, y Gao, 2016, p. 230–235).

2.4.2.4 Marcadores informativos de ancestría (AIMs-INDELs)

Los marcadores AIMs permiten diferenciar entre dos o más poblaciones ancestrales mediante la obtención de frecuencias alélicas. El fin es identificar los grupos nativos de los cuales proviene una determinada población o inferir la biogeografía de un individuo (Campbell y Tishkoff, 2008; Halder, Shriver, Thomas, Fernandez y Frudakis, 2008, p. 648–650).

Para poder excluir poblaciones se utiliza un panel de varios juegos de marcadores. Los marcadores AIMs idóneos son aquellos que son monomórficos en una población pero se encuentran en abundancia en otra (Enoch, 2006; Miller et al., 2005). Estos marcadores son útiles para la

reconstrucción de poblaciones, para probar la asociación de mezcla con fenotipos y en la ciencia forense (Aquino, Jannuzzi, Carvalho, y Gusmão, 2015; Halder et al., 2008).

Una variante son los AIMS-INDELS, estos permiten identificar mediante inserciones y deleciones las diferencias generadas en las frecuencias alélicas entre una o más poblaciones. Estos marcadores tienen una gran versatilidad, baja tasa de mutación y son rápidamente identificables, lo que los convierte en una alternativa idónea para el estudio de polimorfismos a nivel del genoma humano. Principalmente, son útiles para identificar subestructuras poblacionales y determinar el grado de mestizaje (Pereira et al., 2012; Santos et al., 2010; Urbano et al., 2015).

2.4.3 Programas empleados para evaluar la ancestría genética en poblaciones humanas

2.4.3.1 STRUCTURE v2.3.4

Este es uno de los softwares más utilizados para estimar la ancestría global y local de una determinada población. Fue implementado inicialmente por Pritchard y sus colaboradores en el año 2000 sin embargo, posteriormente fue perfeccionado por Falush et al., (2003). El modelo que utiliza este programa se basa en formar agrupaciones o clústeres, dentro de los cuales se infiere la estructura de la población de estudio mediante el análisis del genotipo obtenido a través de marcadores moleculares no ligados (Pritchard, Wen y Falush, 2010).

En este programa lo que se hace es asumir un número de K poblaciones cada una caracterizada por poseer en cada locus un determinado grupo de frecuencias alélicas. Mediante una asignación probabilística, los individuos de estudio son asignados a una o más poblaciones dependiendo de sus

frecuencias alélicas. Cabe recalcar que un requerimiento en este programa son las réplicas, las que proveen confiabilidad a los resultados (Pritchard, Stephens y Donnelly, 2000).

2.4.3.2 STRUCTURE HARVESTER

El Structure Harvester es un programa que permite analizar, condensar y resumir eficazmente los datos obtenidos a partir del programa Structure. Entre sus principales funciones esta estimar el valor de K de una determinada población mediante la implementación del método Evanno. Finalmente, mediante el análisis de algoritmos matemáticos se obtiene un grupo de valores K cuyas características se ajustan de mejor manera a los datos analizados (Earl y VonHoldt, 2012, p. 359–361).

2.4.3.3 CLUMPP y DISRUPT

A partir de los resultados obtenidos mediante los análisis tipo clúster se obtiene una matriz denominada como “*matriz Q*” la que contiene la fracción o probabilidad de pertenencia también conocido como “*coeficiente de pertenencia*” de cada individuo estudiado a un determinado clúster analizado. La sumatoria de estos coeficientes es igual a 1 independientemente del número de clústeres. Sin embargo, debido al uso de repeticiones se pueden obtener resultados distintos de un mismo individuo (Jakobsson y Rosenberg, 2007).

Debido a ello se ha generado la necesidad de condensar los resultados de las repeticiones. Estos softwares son utilizados en el procesamiento de resultados, lo que hacen los programas CLUMPP y DISRUPT es proporcionar una nueva matriz Q donde se encuentra la media de las proporciones de cada clúster obtenidas anteriormente en las réplicas. De tal manera que todas las réplicas de un mismo individuo tengan la coincidencia más cercana posible y sean expresadas como un solo resultado (Jakobsson y Rosenberg, 2007; Rosenberg, 2002).

2.4.3.4 GENEPOP

Este software es utilizado como una herramienta estadística para las ciencias exactas. Entre las pruebas que se pueden llevar a cabo está el equilibrio de Hardy-Weinberg para el estudio de poblaciones. Además, permite obtener estadísticos F, distancias poblacionales entre otras. Básicamente es una consola que es controlada mediante un grupo de opciones y comandos (Rousset, 2008).

CAPITULO III. DISEÑO DEL PLAN EXPERIMENTAL

3.1 DISEÑO EXPERIMENTAL

En este trabajo se buscó determinar los niveles de ancestría en pacientes con Huntington del Ecuador. Para ello, como variable independiente se cuantificaron los niveles de ancestría a partir de las frecuencias alélicas y la variable dependiente fue la incidencia de la enfermedad de Huntington. Para determinar la relación existente entre estas variables primero se extrajo el ADN a partir de muestras de sangre periférica mediante el juego de reactivos PureLink Genomic DNA Mini Kit (Invitrogen, Alemania). Posteriormente, se cuantificó el ADN obtenido empleando el equipo NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific, USA). Continuando con la PCR multiplex, de sus siglas en inglés *Multiplex Polymerase Chain Reaction*, en la que se utilizaron las condiciones de reacción descritas por Pereira et al., (2012). Posteriormente, se analizaron los fragmentos amplificados utilizando el marcador fluorescente Liz 600TM (Applied Biosystems, USA). Este proceso se llevó a cabo en el equipo Genetic Analyzer 3130 utilizando el software GeneMapper® v4.0 (Applied Biosystems, USA) donde se determinaron las frecuencias alélicas. Finalmente, se analizaron los resultados en el programa STRUCTURE 2.3.4 (Pritchard et al., 2000) y se realizó el análisis estadístico mediante los softwares GENEPOP (Raymond y Rousset, 2016; Rousset, 2008) y Prisma v6.0 (GraphPad Software, San Diego, CA) (Figura 5).

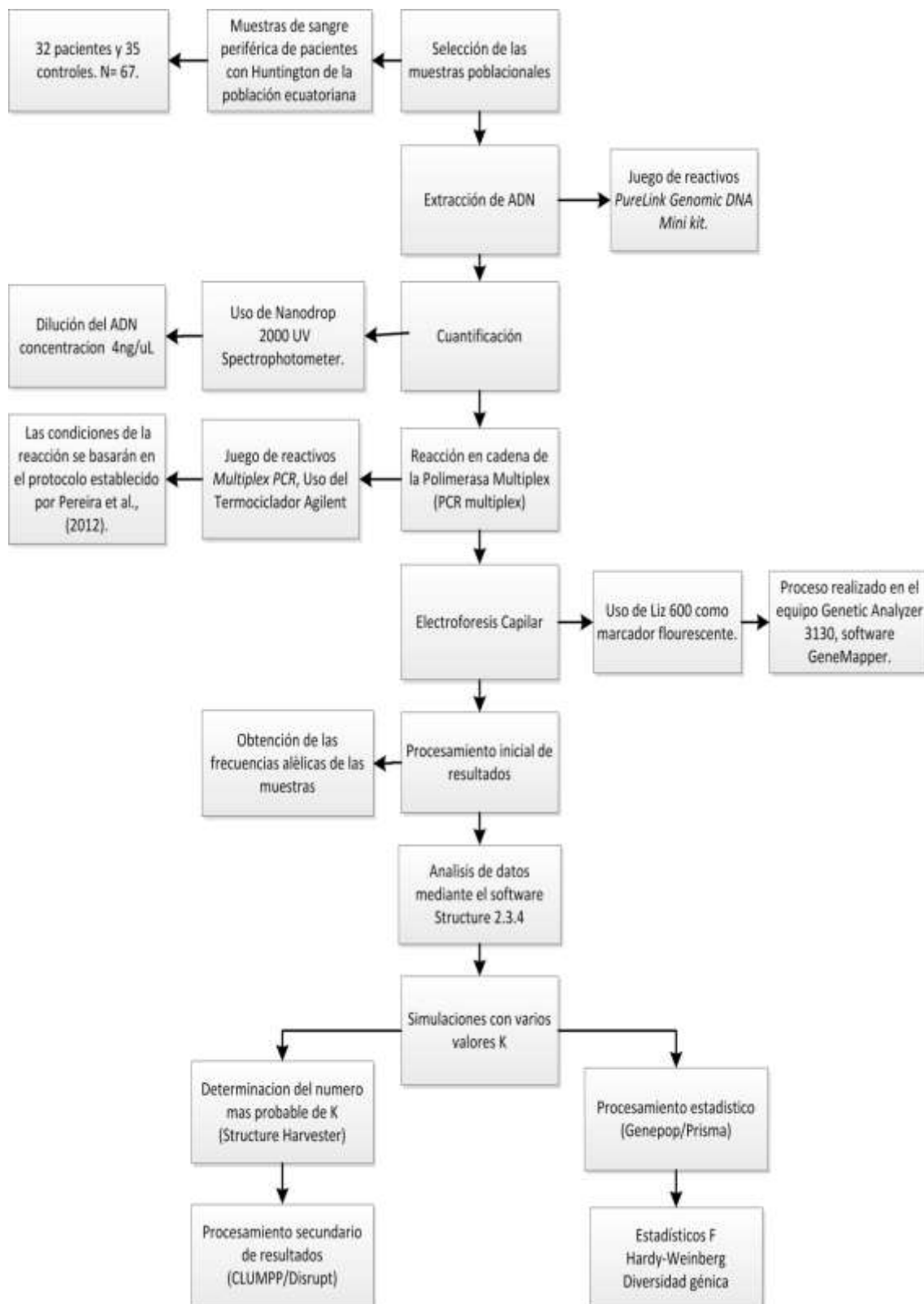


Figura 5.Diagrama de flujo del diseño experimental para la determinación de los niveles preliminares de ancestría en la población ecuatoriana con mutación en el Gen IT-15.

CAPITULO IV. PROCEDIMIENTOS

4.1 POBLACIÓN Y MUESTRA

La población utilizada para este trabajo experimental se centró en una colección de muestras de sangre periférica de pacientes con Huntington de la población ecuatoriana. El muestreo que se realizó fue por conveniencia con una n=32. Para realizar el estudio se utilizaron controles n=35, es decir, muestras de sangre periférica de personas sanas que no presentan ninguna relación con la enfermedad de Huntington. Todas las muestras y controles se encontraban disponibles en el Banco de muestras del Instituto de Investigaciones Biomédicas de la UDLA (IIB).

Este trabajo de titulación forma parte del proyecto denominado “Estudio del Gen IT-15 en el Ecuador” y se llevó a cabo bajo la dirección científica de Carolina Salazar M.Sc. Además, este trabajo de titulación cuenta con los permisos necesarios otorgados por el Comité de Ética y Bioética de la Universidad de las Américas (CEBE) con número de aprobación 2016-1206.

4.2 MATERIALES Y MÉTODOS

4.2.1 Extracción de ADN

El proceso de extracción de ADN se realizó a partir de sangre periférica de los individuos que participaron en el estudio, para ello se utilizó los juegos de reactivos de Invitrogen PureLink Genomic DNA Mini Kit, siguiendo las instrucciones del fabricante Life Technologies (2012).

Inicialmente se colocaron 200 μ l de sangre periférica en un tubo eppendorf estéril de 1,5 mL se añadieron 20 μ l de Proteinasa K y se mezcló mediante vortex durante 10 segundos (Invitrogen, Alemania). Continuando con el proceso, se añadieron a la mezcla 20 μ L de ARNasa se agitó nuevamente en el

vortex. Continuando con el proceso, se incubó a una temperatura de 55°C durante 10 minutos en un bloque térmico Accublock Digital Dry Bath (Labnet, USA). Posteriormente, se añadieron 200 µL de etanol absoluto al 95 % y 200 µL del buffer de lisis y se homogenizó la mezcla en el vortex durante 10 segundos.

Para realizar el proceso de purificación, con el uso de micropipetas se procedió a trasvasar la mezcla a columnas PureLink™ y se centrifugó a 10 000 revoluciones por minuto (RPM). Culminado este proceso se tomaron las columnas y se cambió el tubo colector por uno nuevo, se añadieron 500 µL de Wash Buffer1 y se centrifugó durante un minuto a una velocidad de 10 000 RPM. Se repitió el proceso anterior pero ahora añadiendo 500 µL de Wash Buffer2 y centrifugando a 15 000 revoluciones durante tres minutos (Rotofix 32 A-Hettich, Alemania).

Finalmente, se desechó el tubo de colección y se colocó la columna en un tubo cónico. Se añadieron 25 µL del buffer de elución (PureLink™) y se dejó incubar durante 2 minutos a temperatura ambiente (23°C). Tras la incubación, se centrifugó a 15 000 RPM; este proceso se realizó dos veces con el fin de obtener la mayor cantidad de ADN posible. El tubo eppendorff de 1,5 mL con el ADN purificado fue almacenado a -20°C hasta su posterior uso.

4.2.2 Cuantificación de ADN

Una vez extraído en ADN se determinó su concentración y pureza. Para ello se utilizó el equipo NanoDrop 2000 (ThermoScientific, USA). La concentración del ADN se determinó por su absorbancia a 260 nm y la calidad del mismo fue establecida mediante la relación entre las absorbancias a 260/280 nm.

4.2.3 Reacción en cadena de la polimerasa multiplex (PCR-multiplex)

Para el proceso de amplificación de los genes relacionados con el estudio de ancestría se utilizó el juego de reactivos Multiplex PCR (QUIAGEN, Alemania). El volumen final de la reacción fue de 5 μL y se utilizó el protocolo descrito por Pereira y sus colaboradores (2012), con ciertas modificaciones. Los cebadores utilizados fueron los 46 marcadores AIMS-InDels y se utilizó 1 μL del cebador 196, 0,5 μL del cebador 659 y 2,5 μL de la master mix Quiagen Multiplex PCR.

A esta reacción se añadió un volumen de 0,5 μL del ADN purificado, a una concentración de 4 ng/ μL . Continuando con la amplificación de los fragmentos se utilizaron las condiciones del programa descritas por Pereira et al., 2012 (Tabla 1).

Tabla 1.

Condiciones del programa utilizado para la amplificación de los fragmentos
Adaptado de Pereira et al., 2012.

Protocolo de PCR		
Calentamiento inicial	95°C	15 min
Desnaturalización inicial	28-30 ciclos*	94°C, 30seg
Hibridación		60 °C, 1min 30 seg
Elongación		72°C, 45 seg
Elongación Final	72 °C	Min

*Dependiendo de la presencia de pull-ups.

4.2.4 Electroforesis capilar y obtención de frecuencias alélicas

Este proceso se realizó en el equipo Genetic Analyzer 3130 (Applied Biosystems, USA). Para ello se utilizó como marcador fluorescente Liz 600. Se emplearon 1 μL de amplificado y 9 μL del marcador, para un volumen final de 10 μL (Pereira et al., 2012). Este marcador de fluorescencia está diseñado para marcar fragmentos de 20, 40, 60, 80, 100, 114, 120, 140, 160, 180, 200, 214, 220, 240, 250, 260, 280, 300, 314, 320, 340, 360, 380, 400, 414, 420, 440, 460, 480, 500, 514, 520, 540, 560, 580 y 600 bases (Figura 6A). La finalidad de su uso es que cada uno de los fragmentos estén marcados con un fluoróforo que será visualizado como una sola señal en forma de pico (Figura 6B) (Applied Biosystems, 2005).

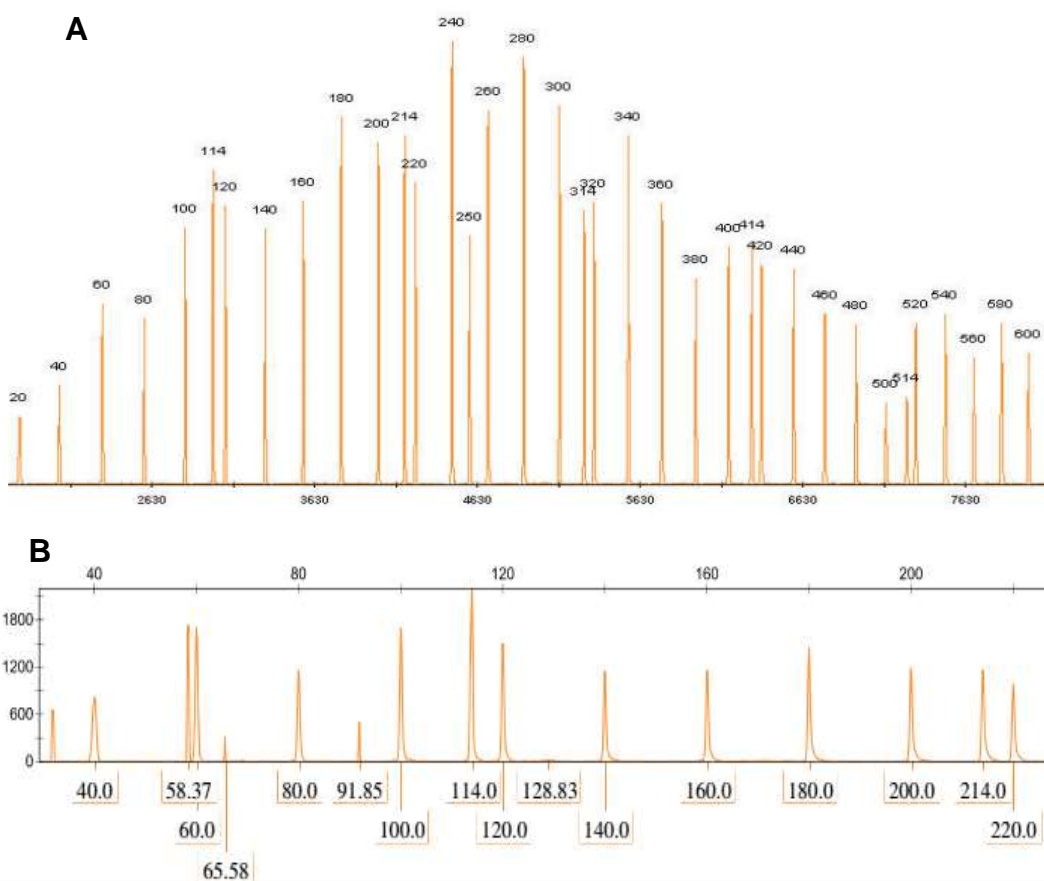


Figura 6. Electroferograma de LIZ 600.

A) Liz 600TM Gene-ScanTM ubicación de cada uno de los fragmentos, (cada pico indica el peso de los marcadores de estudio).
 B) Ejemplo de la lectura obtenida empleando el marcador de fluorescencia en una de las muestras durante el proceso de secuenciación.

Continuando con el proceso se realizó la visualización de los resultados mediante el programa GeneMapper® v4.0 (Applied Biosystems, USA), donde se obtuvieron las repeticiones y las frecuencias alélicas de cada individuo en forma de panel.

Este panel está compuesto por: los marcadores estudiados cuyos nombres se encuentran en la parte superior del panel en forma horizontal (color gris), por las barras verticales que simbolizan los picos correspondientes a cada marcador y por dos franjas denominadas "*bines*" donde se localizan las señales en forma de picos. Estos picos en la parte posterior contienen el número del alelo que le corresponde. El número 1 representa al alelo corto que es propio de las deleciones mientras que el alelo 2 también llamado alelo largo es el correspondiente a las inserciones (Figura 7) (Pereira et al., 2012).

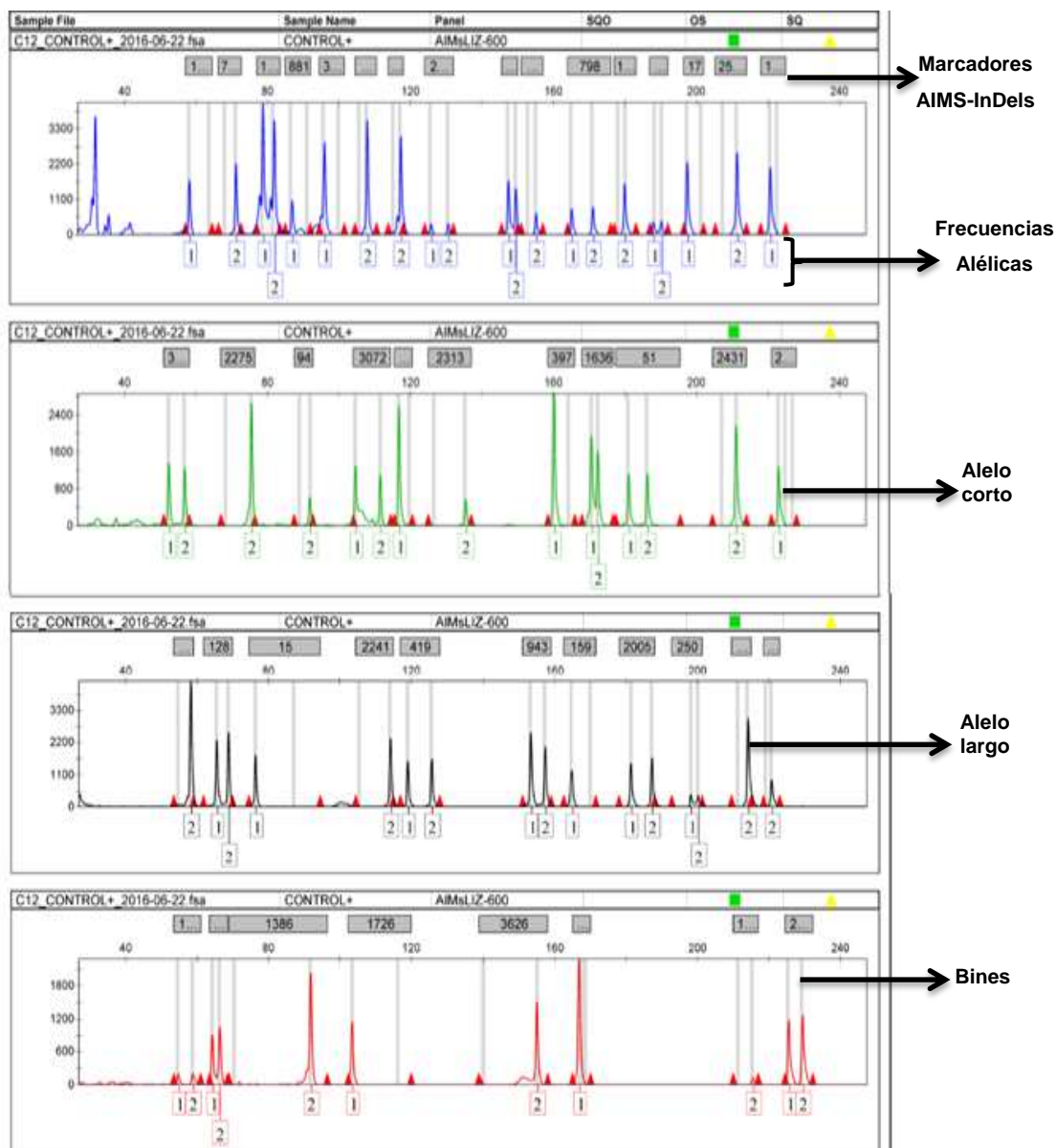


Figura 7. Ejemplo del panel de visualización obtenido a través del programa GeneMapper® v4.0. Presencia de los alelos 1 y 2 respectivamente en cada uno de los 46 marcadores estudiados.

Dependiendo de los alelos que se encontraban presentes en el panel de visualización o electroferograma se determinó que si existían picos en los dos bins del marcador, el individuo era heterocigoto para dicho marcador lo que significaba que existía la presencia de deleciones e inserciones en los alelos del cromosoma (Figura 8A). Por otra parte, si existía únicamente un pico el individuo era homocigoto para el marcador. Si este pico correspondía al alelo 1

existía una delección mientras que, si correspondía al alelo 2 se mostraba que existía una inserción en el respectivo “*bine*” (Figura 8B y 8C) (Väli et al., 2008).

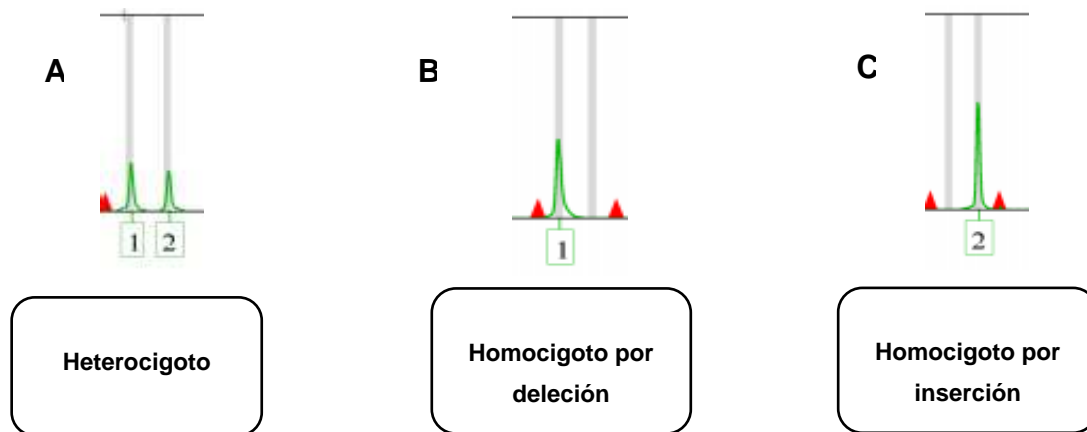


Figura 8. Distintas interpretaciones de las frecuencias alélicas según los picos obtenidos.

- A) Individuo heterocigoto presencia de inserción y delección.
 B) Individuo homocigoto para el alelo 1 presencia de delección.
 C) Individuo homocigoto para el alelo 2 presencia de inserción.

4.3 EVALUACIÓN ESTADÍSTICA DE LOS RESULTADOS

4.3.1 Determinación de la estructura poblacional

El programa empleado para la determinación de la estructura poblacional fue el STRUCTURE v2.3.4 (Falush et al., 2003; Pritchard et al., 2000). Para ello se realizaron simulaciones con los valores de $K=1-10$ con 10 réplicas en cada caso. Las corridas se realizaron con 100.000 réplicas de calentamiento (*burn-in*) y la aplicación de 100.000 MCMC (Monte Carlo de cadenas de Markov).

En el modelo empleado para el análisis se asumió la correlación entre la frecuencia alélica y el grado de mezcla de la población fue inferido de los datos. El análisis se realizó en diferentes muestras poblacionales: (1) pacientes afectados con la enfermedad de Huntington, (2) controles: grupo conformado por individuos sanos, (3) controles y pacientes, en conjunto y (4) controles y

pacientes en el contexto de poblaciones de referencia de África, América y Europa con datos obtenidos a partir de trabajos previos (Pereira et al., 2012). El valor más probable del número de *clústeres* (K) fue analizado en base a la probabilidad (Ln) estimada de los datos (Rosenberg et al., 2001) o según el valor de ΔK (Evanno, Regnaut y Goudet, 2005). Estos dos algoritmos se encuentran implementados en el programa STRUCTURE HARVESTER (Earl y VonHoldt, 2012).

El análisis de los resultados de cada una de las réplicas de las simulaciones realizadas con el programa STRUCTURE v2.3.4, se realizó con la aplicación CLUMPP (Jakobsson y Rosenberg, 2007). Para cada una de las corridas se presentó el valor del número de clústeres más probables (K), el número de individuos de la población y el número de réplicas. Además, se tuvo en consideración el tamaño de la muestra de cada una de las poblaciones evaluadas en el análisis realizado.

El procesamiento de los resultados obtenidos del análisis de la estructura poblacional se realizó con el programa DISRUPT (Rosenberg, 2002). Los parámetros empleados fueron el número de clústeres más probables (K), el número de poblaciones pre-definidas y el número de individuos. Los resultados obtenidos fueron visualizados con el programa GSView 6.0 (Artifex Software, Inc., USA).

4.3.2 Procesamiento estadístico de los resultados

Se calcularon los estadísticos F de Wright (F_{IS} y F_{ST}) (Wright, 1931; WRIGHT, 1949) y el desequilibrio de Hardy-Weinberg (Hardy, 1908; Weinberg, 1912; Weinberg, 1908). Estos algoritmos fueron usados en la aplicación GENEPOP (Raymond y Rousset, 1995; Rousset, 2008).

El análisis estadístico de los resultados se ejecutó en el programa Prisma v6.0 (GraphPad Software, San Diego, CA). Las pruebas de normalidad de los datos se realizaron con las pruebas de D'Agostino & Pearson, Shapiro-Wilk y Kolmogorov–Smirnov. En los casos correspondientes los resultados fueron expresados como media \pm desviación estándar. La diferencia entre las medias se evaluó mediante la prueba de *Student*. El análisis de varianza (ANOVA) se realizó de acuerdo a la prueba de Dunnet. El nivel de confianza empleado en todos los casos fue de un 95 %.

CAPITULO V. RESULTADOS Y DISCUSIÓN

Una vez realizados los procesos antes mencionados se obtuvieron las frecuencias alélicas exportadas en una tabla (Anexo 1) para posteriormente mediante los distintos softwares analizar la diversidad genética ancestral de la población de estudio.

5.1 Determinación de la estructura poblacional de los pacientes afectados con la enfermedad de Huntington

El análisis realizado para los pacientes en el programa Structure 2.3.4 mostró que el valor más probable de los “clústeres” o subpoblaciones fue $K=3$ (Figura 9). Sin embargo, los resultados para el análisis tanto de los controles (Figura 10) como de los controles & pacientes en conjunto (Figura 11) están asociados a una $K=2$ como el número de “clústeres” más probable. Lo que indica que estos grupos muestrales están subdivididos en tres y dos subpoblaciones respectivamente.

Estos resultados se muestran en la matriz Q (Anexo 2) obtenida a partir del programa STRUCTURE. Esta matriz contiene la proporción que cada individuo tiene en relación a un clúster determinado. La proporción de individuos en cada uno de los clústeres indicó que en los pacientes, los elementos genéticos se dividen en fracciones semejantes. Mientras que para las otras dos simulaciones (controles & pacientes) ($K=2$) los individuos se reparten en una proporción aproximada del 50 % (Tabla 2).

Los valores de subpoblaciones posibles (K) fueron obtenidos ya que este programa se basa en un algoritmo de tipo Bayesiano, que subdivide a la muestra poblacional analizada en varias subpoblaciones o clústeres (valor K) que satisfacen el equilibrio de Hardy-Weinberg (Hardy, 1908; Wagner, 1999; McKeigue, Carpenter, Parra y Shriver, 2000). Aunque la estructura de una

población puede ser inferida en base a su historia ancestral, el número de subpoblaciones podría obtenerse con mayor exactitud mediante el análisis de marcadores genéticos.

Al emplear el análisis gráfico mediante el programa STRUCTURE HARVESTER se corroboraron los resultados obtenidos inicialmente. Esto se debe a que este programa emplea un algoritmo que se basa en el método propuesto por Evanno (Evanno et al., 2005) donde se obtienen valores de probabilidad $Ln(K)$ para posteriormente obtener el ΔK y establecer el número máximo que indica el número de subpoblaciones en las que se puede agrupar la población de estudio (Earl y VonHoldt, 2012; Evanno, Goudet, 2005). De esta manera, mediante la implementación de algoritmos, el estimado del posible número de clústeres en los que se divide la población original se puede establecer de modo menos subjetivo. (Figura 9, 10 y 11 respectivamente).

Un estudio de ancestría realizado en la población Colombiana utilizando los mismos marcadores AIMs-InDels indicó que el número de subpoblaciones fue igual al obtenido en el análisis de los pacientes de la población ecuatoriana ($K=3$). Esto se puede deber a que estas poblaciones geográficamente cercanas poseen un alto grado de mezcla interétnica, lo que proporciona estimaciones confiables de la descendencia para investigar su asociación a trastornos genéticos complejos asociados a una determinada patogenia (Garavito et al., 2015).

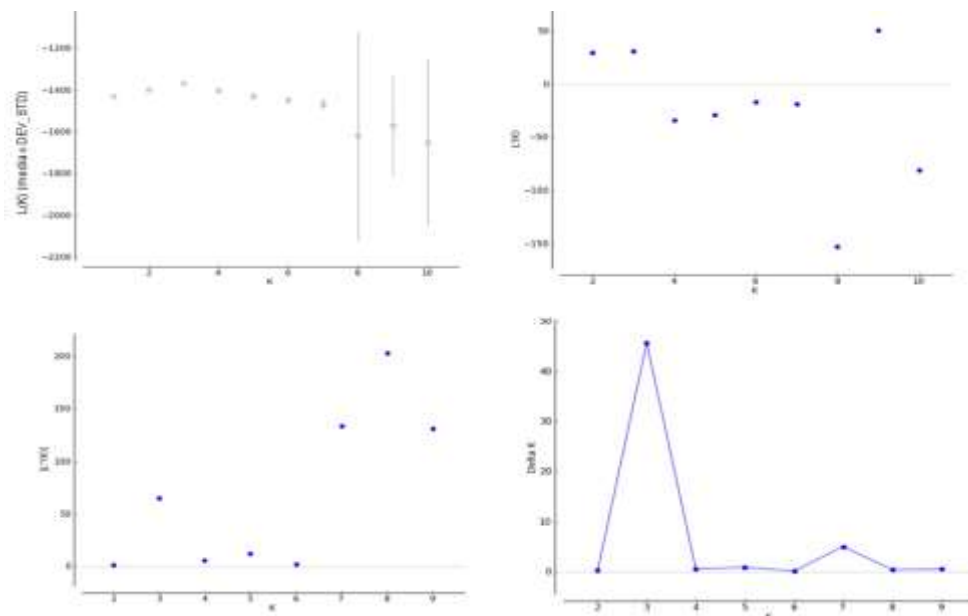


Figura 9. Determinación del valor más probable de K en la muestra de pacientes afectados por la enfermedad de Huntington (N=32) mediante el método de Evanno et al., (Evanno, Regnaut y Goudet, 2005) evaluado en la aplicación STRUCTURE HARVESTER. En los gráficos se muestra la media del valor estimado del Ln(K), la primera y la segunda derivada de la función de probabilidad y el delta K ($\text{mean}(L''(K)/\text{sd}(L(K)))$) para los valores del número de clústeres (K) evaluados en la simulación.

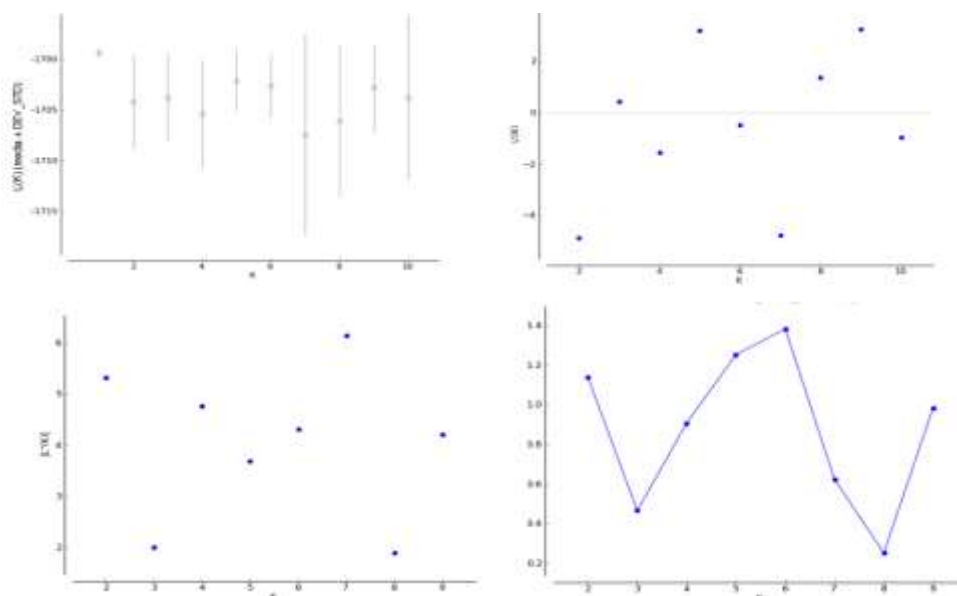


Figura 10. Determinación del valor más probable de K en el grupo poblacional control empleado en el estudio (N=35) mediante el método de Evanno et al., (Evanno, Regnaut y Goudet, 2005) evaluado en la aplicación STRUCTURE HARVESTER. En los gráficos se muestra la media del valor estimado de Ln(K), la primera y la segunda derivada de la función de probabilidad y el delta K ($\text{mean}(L''(K)/\text{sd}(L(K)))$) para los valores del número de clústeres (K) evaluados en la simulación.

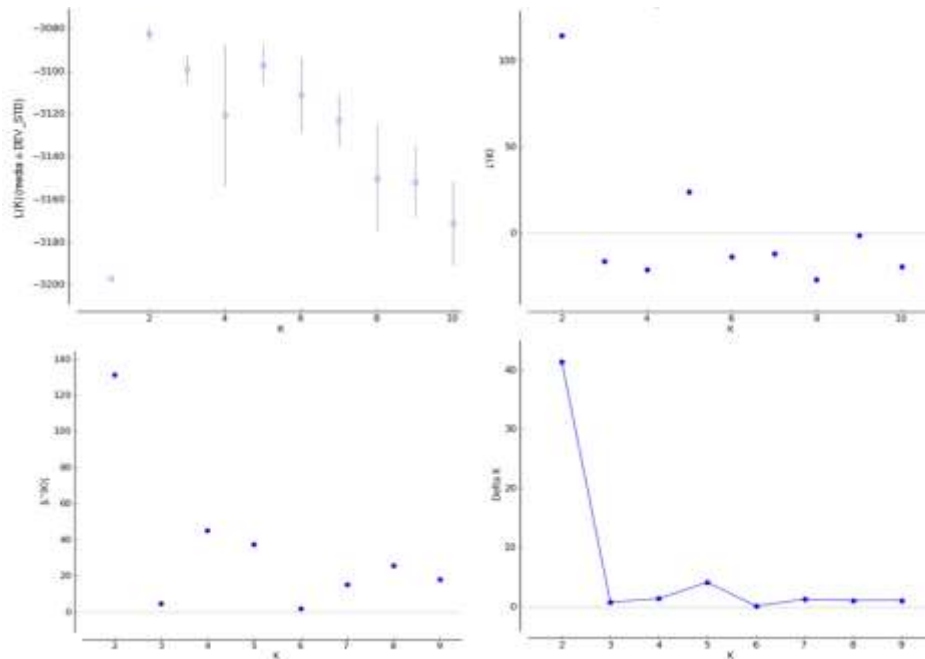


Figura 11. Determinación del valor más probable para K para el grupo global conformado por la población control empleado en el estudio (N=35) y en los pacientes (N=32) mediante el método de Evanno et al., (Evanno, Regnaut y Goudet, 2005) evaluado en la aplicación STRUCTURE HARVESTER. En los gráficos se muestra la media del valor estimado de Ln(K), la primera y la segunda derivada de la función de probabilidad y el delta K ($\text{mean}(L''(K)/\text{sd}(L(K)))$) para los valores del número de clústeres (K) evaluados en la simulación.

Tabla 2.

Proporción de los individuos en cada una de las simulaciones en relación a los clústeres inferidos.

Simulaciones	Clústeres inferidos			Número de individuos
	1	2	3	
Pacientes	0,312	0,371	0,317	32
Controles	0,500	0,500	-	35
Controles y pacientes	0,520	0,480	-	67

Las distribuciones correspondientes a los niveles de ancestría para la muestra de los pacientes y los controles reflejan una distribución diferente en la matriz Q (**¡Error! No se encuentra el origen de la referencia.** 12). Esto se debe a que el valor K varía entre las tres poblaciones analizadas, lo cual puede tener relación con el origen ancestral de la enfermedad.

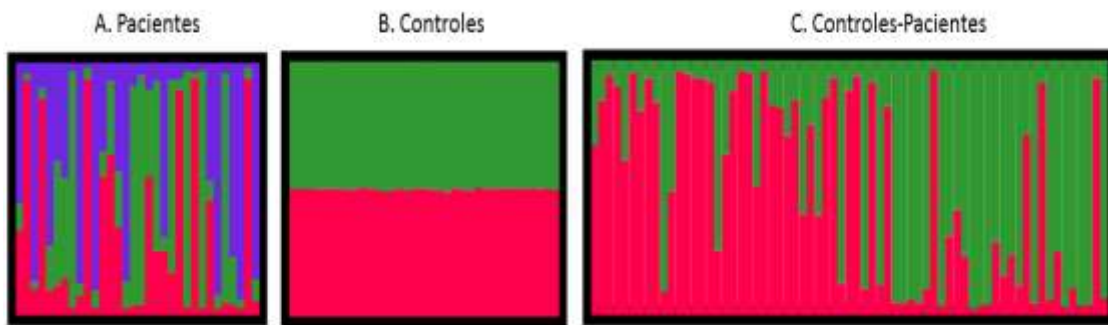


Figura 12. Representación gráfica de los niveles de ancestría para los individuos de las muestras poblacionales analizadas. A) Pacientes B) Controles C) Controles y Pacientes en conjunto. Los resultados se expresaron como los promedios correspondientes a las 10 réplicas realizadas en las simulaciones con el programa STRUCTURE. El procesamiento se realizó con las aplicaciones CLUMPP y DISRUPT. Los clústeres 1, 2 y 3 se muestran en los colores rojo, verde y azul, respectivamente.

La organización de los pacientes de acuerdo al valor de la matriz Q (Figura 13) mostró como esta población se encuentra dividida en tres subestructuras. Estos resultados también fueron visualizados en una “*representación triangular*” donde se evidenció la distribución de los pacientes en el contexto de tres subpoblaciones (Figura 14). Aunque existe una distribución que tiende hacia los vértices de la representación (Clúster 1-3), también se evidencia una dispersión de los individuos. Lo que significa que los individuos que se encuentran en los vértices están definidos genéticamente por una determinada subpoblación mientras que, los individuos que se encuentran dispersos son aquellos que están conformados en diferentes proporciones por más de una subpoblación contribuyendo así a su origen genético ancestral (Bryc et al., 2010).

De acuerdo al análisis realizado con el programa STRUCTURE v2.3.4 y STRUCTURE HARVESTER se observó que las poblaciones evaluadas tienen elementos ancestrales que se diferencian entre sí ($K=3$ y 2) como una posible consecuencia del origen de la enfermedad. De tal modo, parece ser probable la distribución ancestral de la población de los pacientes en tres subpoblaciones.

Como ya lo han establecido otros grupos estas subpoblaciones son llamadas grupos parentales (europea, africana y nativa americana) debido a que se cree que son el origen de la población actual. Además de ello, estas poblaciones han sido establecidas como poblaciones de referencia debido a que son la base para realizar estudios de ancestría gracias a su variabilidad genética y aporte hacia la generación de nuevas sub-etnias (Mulligan et al., 2004; Sánchez, González, & Martínez, 2003).

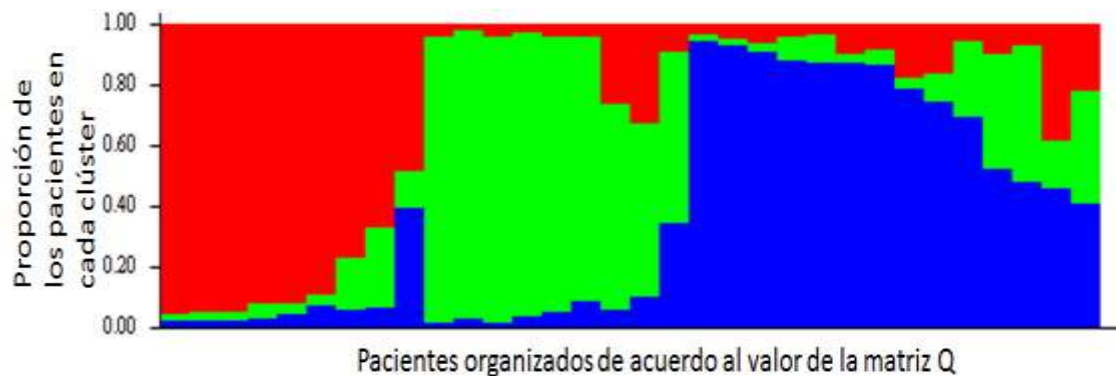


Figura 13. Representación gráfica en el orden de la matriz Q obtenida con el programa STRUCTURE de los niveles de ancestría para los pacientes con Huntington.

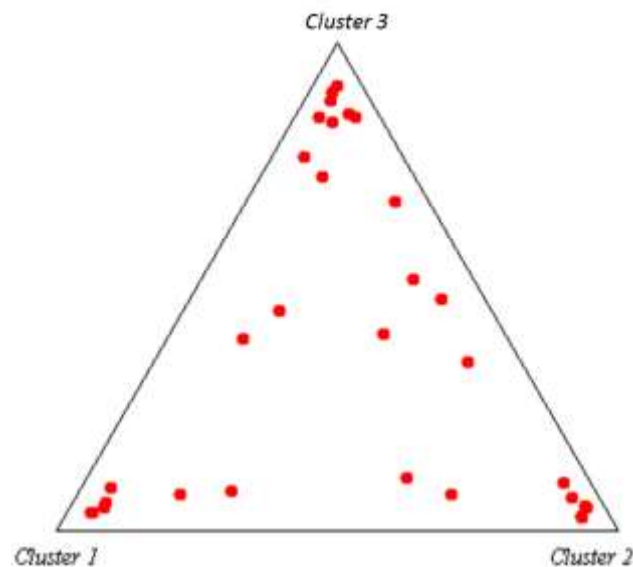


Figura 14. Representación triangular de la distribución de los pacientes afectados con Huntington en los tres clústeres determinados en los que se distribuye de modo más probable esta población desde el punto de vista ancestral.

5.2 Análisis de la estructura de la población de los pacientes afectados con la enfermedad de Huntington en el contexto de las poblaciones de referencia

Debido a que el resultado obtenido del valor de K para los pacientes con Huntington ($K=3$) indicó que su estructura poblacional está relacionada con tres grupos ancestrales se obtuvo el análisis en el contexto de las tres poblaciones de referencia (americana, europea y africana) que según González et al., (2007) y Baeta et al., (2012), corresponden a los orígenes de la población ecuatoriana. A pesar de que el análisis implicó a los tres grupos establecidos como referencia (Pereira et al., 2012), el número de clústeres también fue estimado mediante el método de Rosenberg (2001). Estos resultados del análisis de la función de probabilidad $L(K)$ obtenidos de la simulación con varios valores de posibles subpoblaciones indicaron que la población original puede ser agrupada en tres fracciones (Figura 15) confirmando que la población ecuatoriana tiene sus bases ancestrales en las poblaciones ya mencionadas.

Los valores de la función de probabilidad alcanzaron un valor máximo alrededor de $K=3$. Estos resultados indican que la presencia de otros grupos poblacionales en el contexto de las poblaciones de referencia no modifica la estructura ancestral original de los pacientes con Huntington ($K=3$). Esto está dado porque los individuos que conforman el grupo de los controles y pacientes comparten elementos ancestrales con las poblaciones de referencia analizadas (Cárdenas et al., 2015; González et al., 2007; Pereira et al., 2012; Phillips et al., 2007, p. 273–280). En estudios realizados en poblaciones mixtas como Latinos, Afroamericanos, poblaciones de Centroamérica y Sudamérica se ha visto que su origen ancestral provienen de una mezcla interétnica generada a partir de los tres grupos parentales (europeos, africanos y nativo americanos) lo que confirma que existen elementos genéticos en común (Santos et al., 2010).

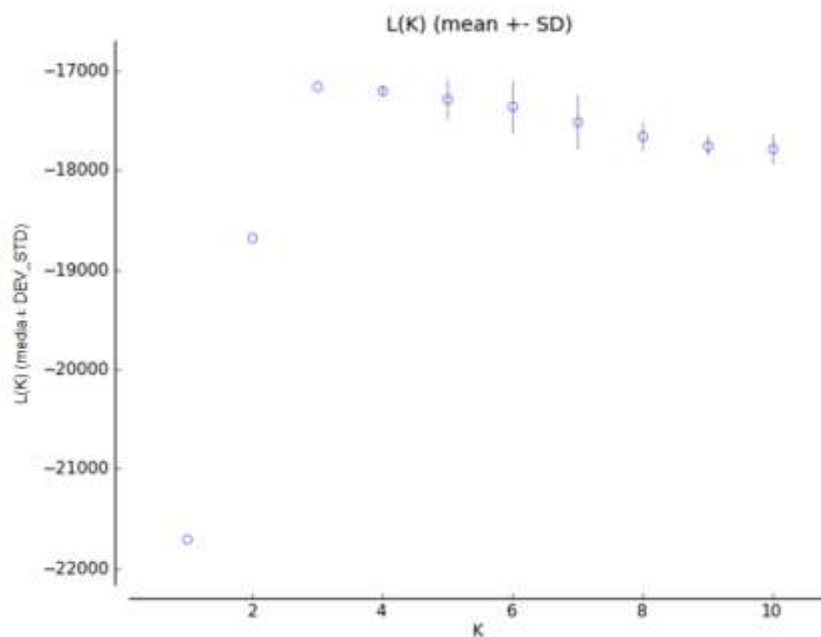


Figura 15. Selección del número de clústeres más probable de acuerdo con los datos obtenidos del análisis bayesiano realizado con el programa STRUCTURE v2.3.4 empleando la aplicación STRUCTURE HARVESTER. La muestra analizada estuvo conformada por: 32 pacientes afectados con la enfermedad de Huntington, 35 individuos sanos (controles) y las muestras de referencia de África (105 individuos), América (64 individuos) y Europa (158 individuos).

Entre los tres clústeres relacionados con las poblaciones de referencia, el primero está asociado a la población de origen africano (97,2 %), el segundo clúster lo constituyen los individuos nativos americanos (95,8 %) y el tercero de los grupos identificados corresponde a la población europea (96,8 %). La estructura de la población de pacientes afectados con Huntington se encuentra enmarcada fundamentalmente en el clúster 2 (83,7 %) y en el clúster 3 (14,5 %). En los controles el 51 % se encuentra en el clúster 2 y el 45 % en el clúster 3. Los niveles de ancestría de la población africana son más notables en el grupo de los controles (4,1 %) en relación con el grupo de los pacientes (1,8 %) (Tabla 3).

Tabla 3.

Proporción de cada una de las poblaciones predeterminadas en cada uno de los clústeres obtenidos.

Poblaciones predefinidas	Clústeres inferidos			N
	1	2	3	
África (Pop 1)	0,972	0,013	0,015	105
América (Pop 2)	0,010	0,958	0,032	64
Europa (Pop 3)	0,011	0,021	0,968	158
Controles (Pop 4)	0,041	0,510	0,449	35
Pacientes-Huntington (Pop 5)	0,018	0,837	0,145	32

* *N=tamaño de la muestra poblacional.*

La representación de los resultados ordenados de acuerdo con las matrices Q de los pacientes (Figura 16A) corroboró la estructura genética en tres clústeres. En el ploteo triangular (Figura 16B) se evidenció que cada una de las poblaciones establecidas se agrupaban en un vértice determinado. Además, se observó como los controles están distribuidos entre los clústeres 1 y 2 correspondientes a las poblaciones de referencia nativo americanas y europeas. Sin embargo, los pacientes se encuentran mayormente enmarcados en el clúster número 2 correspondiente a la población de referencia con origen nativo americano, indicando que este componente tiene un gran aporte en esta población (89,7 %). A pesar de que se cree que el origen de la enfermedad es africano, en las dos poblaciones analizadas este componente tiene escasa contribución (Krause et al., 2015; Liou, 2010; Warby et al., 2011). Estos resultados también fueron visualizados cuando se realizó el tratamiento de los datos con CLUMPP/DISRUPT (Figura 16C).

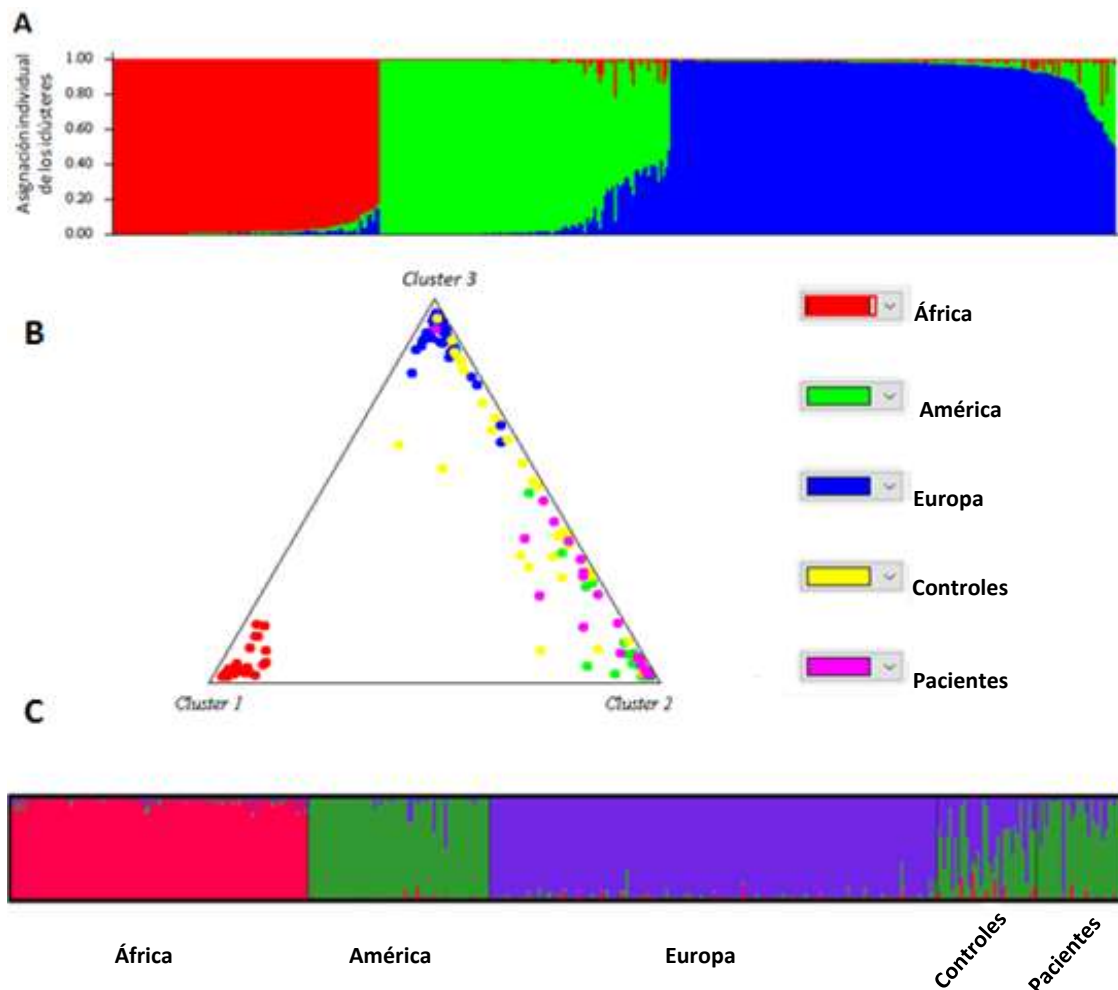


Figura 16. Análisis de la estructura de las poblaciones Controles y Pacientes en el contexto de poblaciones de referencia. A) Representación de los niveles de ancestría individuales en el orden de la matriz Q.

B) Ploteo triangular para las tres subpoblaciones asumidas en el análisis con el programa STRUCTURE.

C) Análisis con el programa CLUMPP y visualización con el DISRUPT de los resultados obtenidos para la simulación con 10 réplicas.

La distribución de los coeficientes ancestrales en cada uno de los clústeres indicó que el primer clúster está conformado por la población de África (Pop1). El segundo clúster está conformado por la población de referencia de América (Pop2) y el tercero por la población de origen europeo (Pop3). Por otra parte, los controles (Pop4) tienen una contribución de los elementos asociados a las poblaciones de referencia americana y europea, mientras que la población conformada por los pacientes afectados con la enfermedad de Huntington (Pop5) está relacionada fundamentalmente con los elementos ancestrales

americanos (Figura 17). Este resultado pudiera estar en contradicción con los obtenidos por otros grupos, que indican que esta enfermedad afecta mayoritariamente a individuos de origen africano y caucásico (Bardien et al., 2007, p. 2083–2089; Cazeneuve y Durr, 2014, p. 168-170; Krause et al., 2015; Moily et al., 2014). Esto podría deberse a que la muestra es originaria de una región geográfica que posee altas proporciones de individuos de origen nativo americano y que el componente ancestral africano se encuentra en bajas proporciones.

Según la investigación realizada por Bryc et al.,(2010), la población ecuatoriana presenta los niveles más bajos de ancestría africana en relación a otras poblaciones Latino Americanas y a su vez, esta presenta los niveles más altos de ancestría nativa-americana. Lo que podría deberse a la presencia de grupos nativos que han permitido mantener su estructura genética.

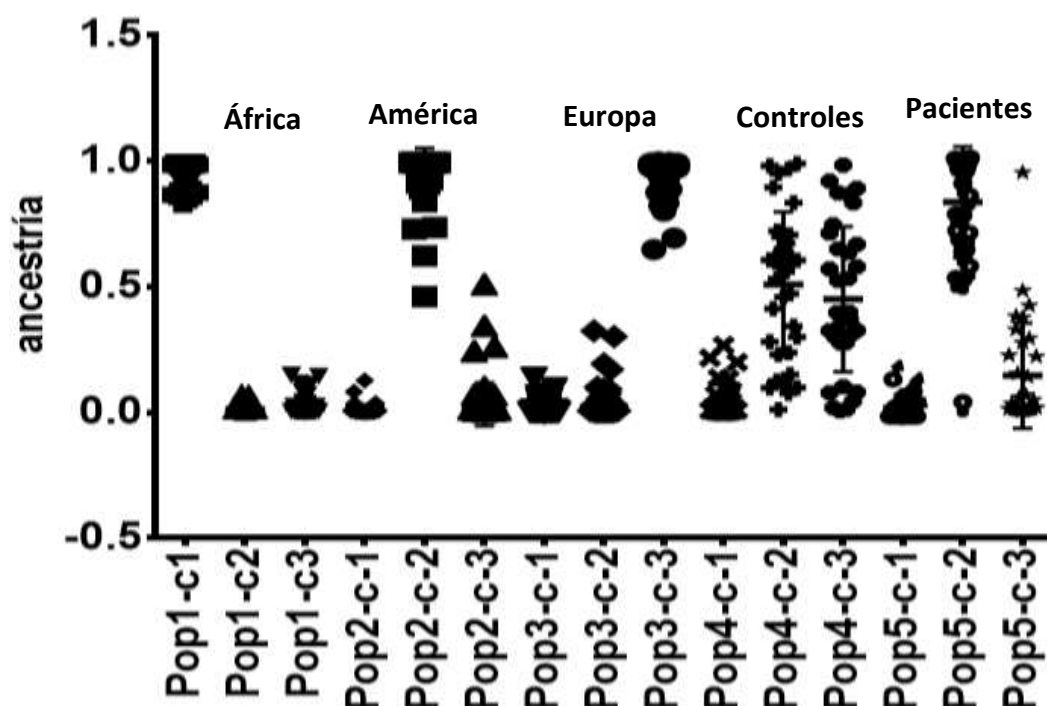


Figura 17. Valores de la matriz Q para cada uno de los grupos evaluados (controles y pacientes) en el contexto de las poblaciones de referencia (África, América y Europa). La identidad de cada una de las poblaciones (Pop 1-pop 5) corresponde a lo indicado en la Tabla 3.

5.3 Evaluación de parámetros genéticos

Los estadísticos F de Wright (F_{IS} y F_{ST}) permitieron analizar la estructura genética de las poblaciones (Wright, 1931). Los valores del estadístico F_{IS} se evaluaron mediante dos algoritmos distintos en el contexto de tres pruebas estadísticas en las que se consideró como hipótesis nula (H_0) que existía equilibrio de HW. Las hipótesis alternativas (H_1) estuvieron en función del valor de probabilidad P, el déficit y exceso de heterocigotidad. Los valores de F_{IS} obtenidos para las tres aproximaciones fueron promediados para cada uno de los algoritmos (W&C y R&H). Los valores positivos indicaron que el número de heterocigotos observados fue menor que el esperado, mientras que los valores negativos mostraron que hay un exceso de heterocigotos, en relación al valor esperado.

Aunque se evaluaron 46 marcadores genéticos con la aplicación GENEPOP en el caso de los controles no se obtuvo información para 2 de esos marcadores, mientras que en el caso de los pacientes no se obtuvo información para 6 de los 46 marcadores evaluados (Anexo 3). Esto pudo deberse a la poca diversidad genética de los mismos (Raymond & Rousset, 2016)

Mediante el estadístico F_{IS} se estableció la deficiencia y exceso de heterocigotos en la población global con el fin de determinar el equilibrio de Hardy-Weinberg. Los valores obtenidos fueron comparados con cero, dando como resultado que mientras más alejado de este valor se encontraban más alejada del equilibrio estaba la población para dicho locus, en caso contrario se aceptó la hipótesis nula (la población está en equilibrio de Hardy-Weinberg).

El valor mínimo de F_{IS} determinado para los controles y pacientes estuvo alrededor de -0,3. Para el grupo conformado por los pacientes este valor fue aproximadamente de -0,33. Los máximos en ambos grupos estuvieron

alrededor de 0,68. Los valores medios para los controles fueron 0,012-0,013 (Figura 18A) y para los pacientes 0,029-0,030 (Figura 18B). El análisis de los datos implicó que solo en el caso de los pacientes los datos se ajustaron a una distribución normal (Anexo 4).

Para los controles, los marcadores en los que se detectó mayor deficiencia heterocigótica fueron 593 ($F_{IS} = 0,67$), 94 ($F_{IS} = 0,34$), 159 ($F_{IS} = 0,3$), 1603 ($F_{IS} = 0,2$), 2264 ($F_{IS} = 0,26$), para los pacientes estos marcadores fueron 798 ($F_{IS} = 0,7$), 51 ($F_{IS} = 0,4$), 2313 ($F_{IS} = 0,35$), 406 ($F_{IS} = 0,3$). Los marcadores más asociados a un exceso de heterocigotos fueron: en los controles, 2275 ($F_{IS} = -0,3$), 128 ($F_{IS} = -0,26$) y 2719 ($F_{IS} = -0,21$) mientras que en pacientes, 159 ($F_{IS} = -0,33$), 584 ($F_{IS} = -0,3$), 1734 ($F_{IS} = -0,28$), 250 ($F_{IS} = -0,26$).

Debido a los bajos valores obtenidos para el estadístico F_{IS} se determinó que la población se encontraba en el equilibrio de Hardy-Weinberg, lo que cumple con lo establecido en el artículo base y en estudios posteriores sobre ancestría (Cabrera et al., 2016; Halder et al., 2008; Ossa et al., 2015; Pereira et al., 2012).

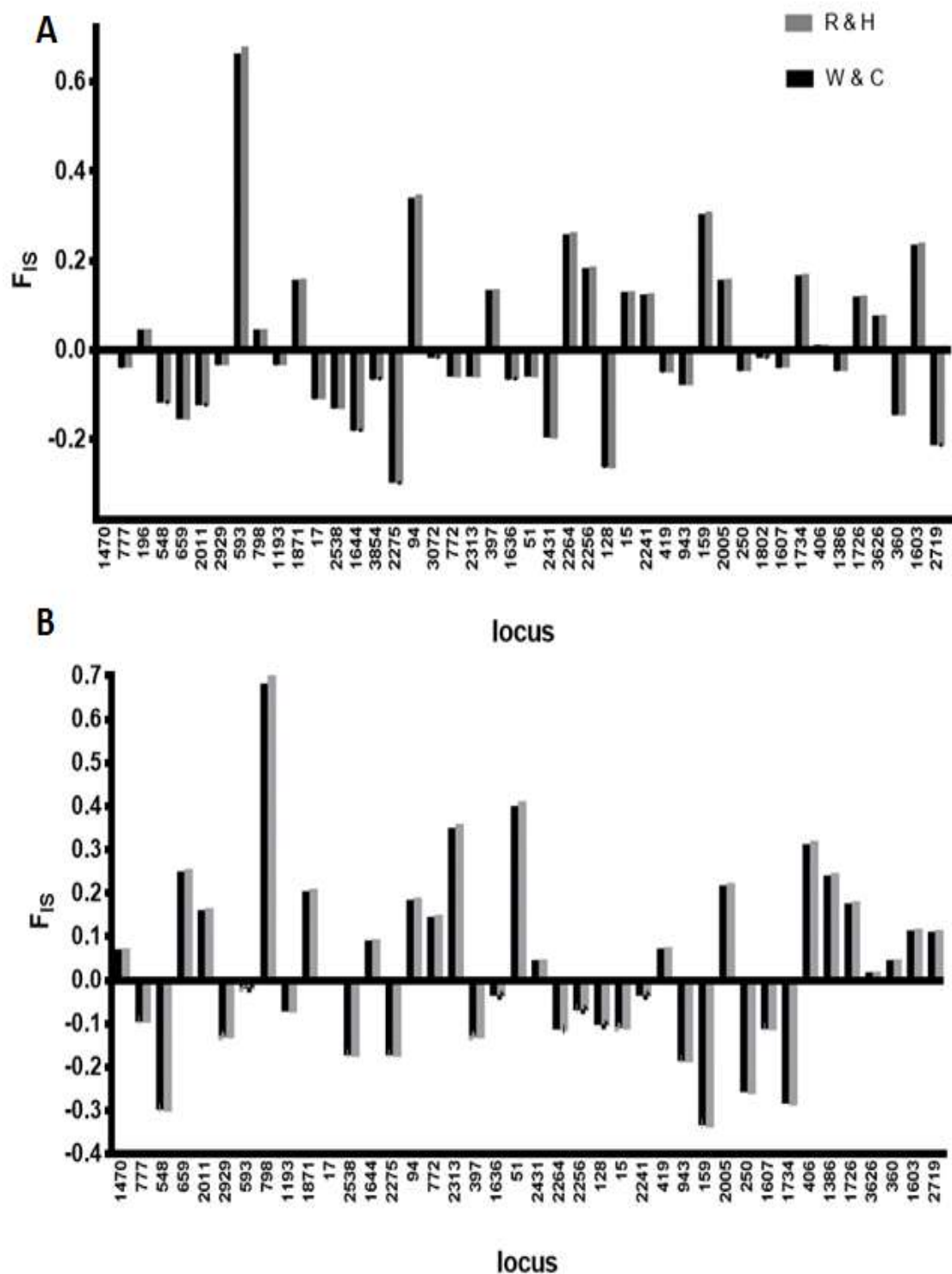
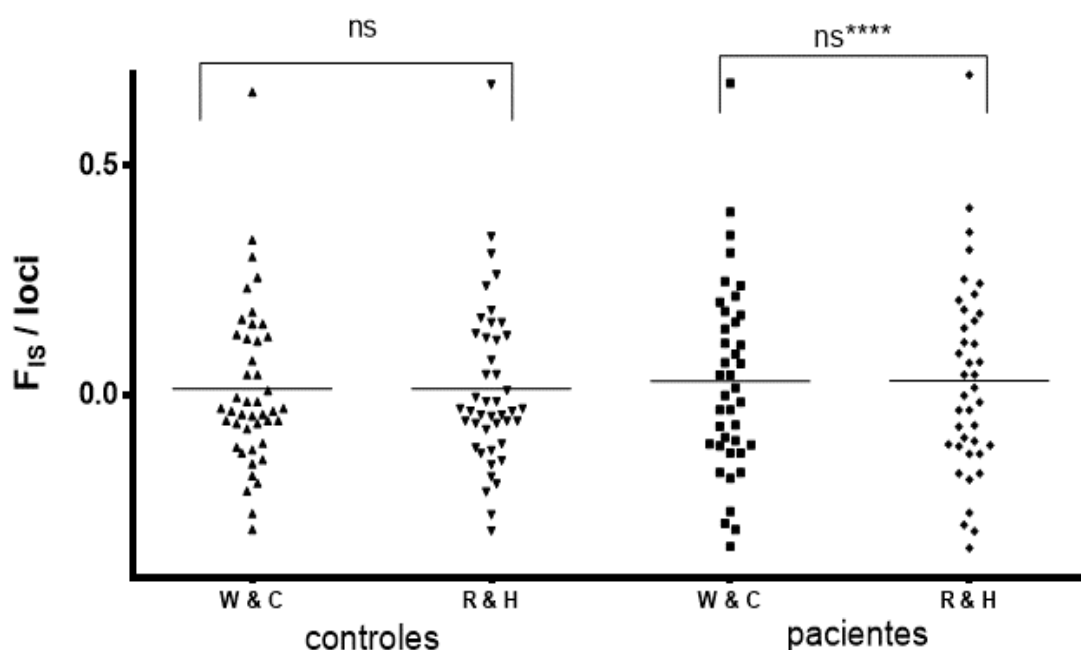


Figura 18. Prueba de Hardy-Weinberg. Valores promedio del estadístico F de Wright para cada uno de las tres hipótesis alternativas (H_1). Los valores de F_{IS} fueron determinados por los métodos de W&C (barras en negro) y R&H (barras en gris).

A) Controles conformados por individuos aparentemente sanos.

B) Pacientes afectados con la enfermedad de Huntington.

El resultado del análisis de los valores de F_{IS} mediante una prueba de *Student* (Figura 19) indicó que no hay diferencias significativas ($p=0,05$) entre los valores evaluados, utilizando las dos pruebas estadísticas descritas previamente en metodología. Los resultados para los controles (Anexo 5) fueron obtenidos por un método no paramétrico mientras que los pacientes fueron analizados en empleando un método paramétrico (Anexo 6).



Método de determinación de F_{IS}

Figura 19. Prueba de Student para los valores de F_{IS} determinados por dos algoritmos distintos en tres análisis con diferentes hipótesis alternativas para las poblaciones individuales conformadas por los controles y los pacientes.

Los F_{IS} determinados por los métodos W&C y R&H presentaron una correlación lineal. En ambos grupos (controles y pacientes) la pendiente de la correlación es 1,019 y el intercepto 0,0006081 (Figura 20A y B, respectivamente). Los resultados obtenidos mostraron que, independientemente del método que se utilice para su cálculo la población conformada por los pacientes, cumple con el equilibrio de Hardy-Weinberg ya que la media de los valores de este estadístico es muy cercana a cero.

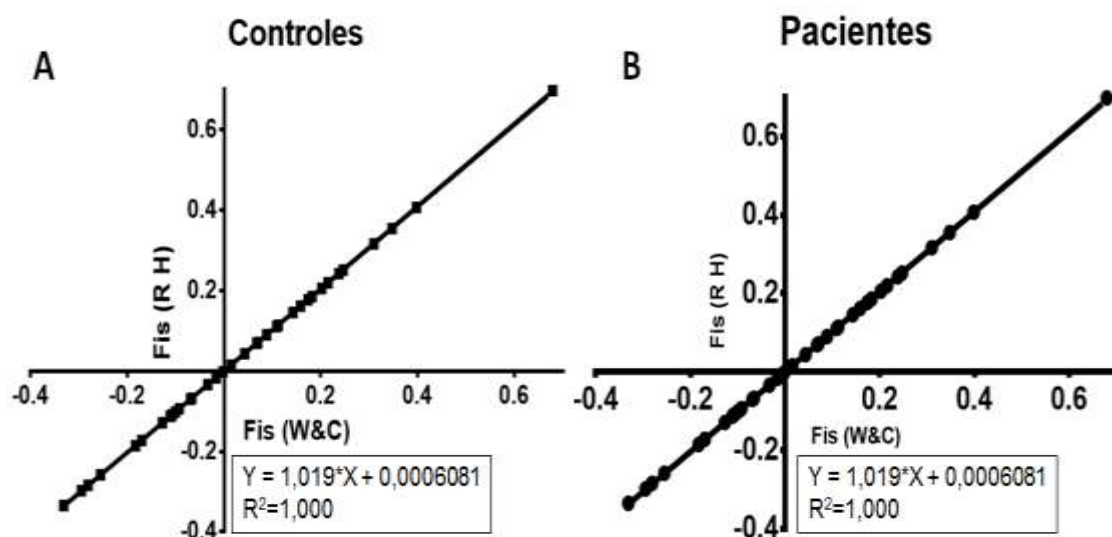


Figura 20. Correlación entre los F_{IS} determinados por los métodos W&R y R&H.

A) Controles, B) pacientes.

El estadístico F_{ST} , permitió evidenciar la diferenciación que existe entre subpoblaciones. Sus valores se encuentran dentro del rango de cero a uno, donde cero significa que no hay estructura en la población y uno indica que las poblaciones son independientes o desde el punto de vista genético se encuentran muy separadas (Powell y Wright, 1979; Wright, 1931). Mediante este estadístico se calcularon las distancias genéticas además de inferir los niveles de relación entre las distintas poblaciones a partir de frecuencias alélicas tal como lo explica Rondón y Barreto (2013).

El análisis de la distancia entre la población de los pacientes afectados con la enfermedad de Huntington, los controles y las tres poblaciones de referencia evidenció que la población de pacientes estudiada se encuentra más relacionada con las referencias de América, así como a la población empleada como control (Figura 21). Esto puede ser la consecuencia de su origen geográfico, de su cercanía genética y de que provienen de una misma población muestral (Hincapié et al., 2009; Long, 1991).

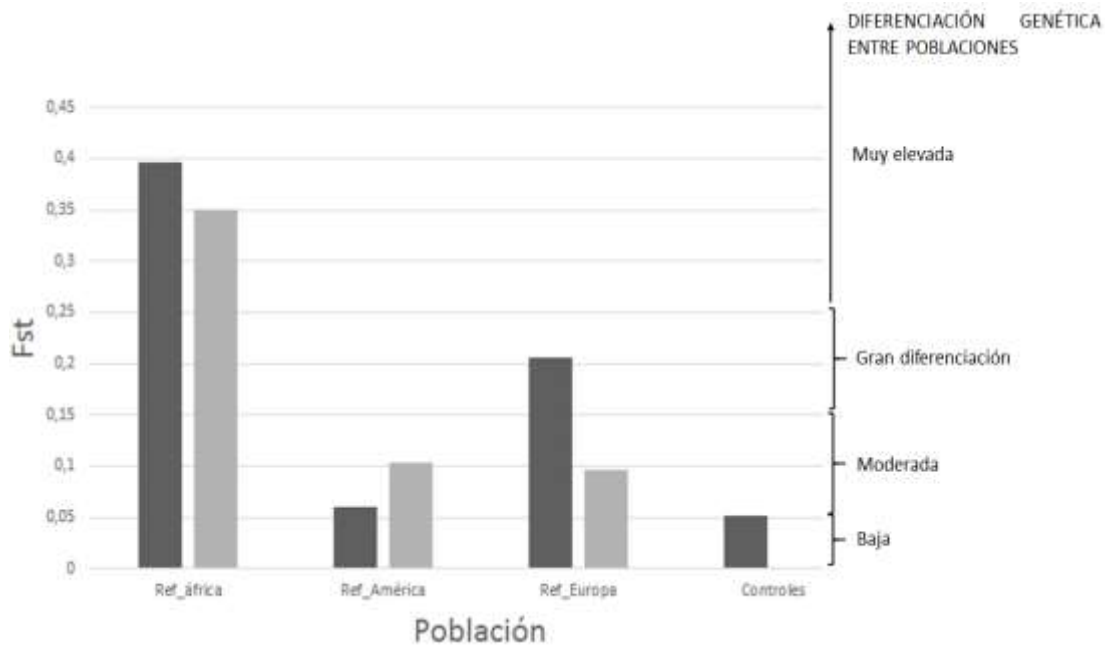


Figura 21. Evaluación de la diferenciación entre poblaciones mediante el estadístico F_{ST} de Wright.

En barras negras (pacientes) y grises (controles) se muestran los valores del parámetro estadístico en relación con cada una de las poblaciones de referencia.

Los valores de F_{ST} obtenidos indicaron que según los rangos establecidos por Wright (1979), la diferenciación de las frecuencias alélicas de la población de pacientes con Huntington es muy elevada con relación a la población africana ($F_{ST}=0,39$), con respecto a la población europea presenta gran diferenciación ($F_{ST}=0,21$) mientras que con relación a la población americana y los controles presenta una diferenciación moderada ($F_{ST}=0,07$) y baja ($F_{ST}=0,05$) respectivamente.

Estos resultados están en contradicción con lo establecido por otros grupos sobre el origen de la enfermedad, razón por la que se debe tener en cuenta la composición y tamaño de la muestra poblacional evaluada además de considerar que esta es autóctona de una región geográfica determinada (Bryc et al., 2010; González et al., 2007). Un estudio realizado utilizando ADNmt estableció que la población ecuatoriana tiene una contribución genética amerindia elevada, independientemente de que la presencia de la población

europea generó la extinción de una gran cantidad de grupos étnicos amerindios (Baeta et al., 2012,p. 299).

En relación al desequilibrio de unión (*linkage disequilibrium*) de cada par de loci no se obtuvieron resultados estadísticamente significativos. Por tal motivo, se puede afirmar que los loci evaluados están relacionados en modo aleatorio, lo que concuerda con lo descrito por Pereira et al., (2012) en su investigación sobre ancestría. En genética poblacional se debe asegurar que no haya transferencia ligada, con el fin de que cada locus analizado pase a la descendencia de manera independiente y no se generen problemas en este tipo de estudios poblacionales que trabajan con un locus a la vez como lo indica Griffiths et al.,(2002).

El análisis de la variación génica intraindividuos (1-Qintra) e intrapoblaciones (1-Qinter) indica que la diversidad genética tuvo valores entre 0,1 y 0,5 (Figura 22). Adicionalmente, se realizó el análisis de la diversidad solo en contexto de las muestras poblacionales de los pacientes afectados con la enfermedad de Huntington y los controles de modo independiente. Los parámetros evaluados fueron el 1-Qintra (Figura 23A) y 1-Qinter (Figura 23B) y no se encontraron diferencias estadísticamente significativas ($p=0,05$) entre estos parámetros en relación a los pacientes y controles.

La diversidad genética intraindividuos es la misma que interindividuos de cada población. Genéticamente en el contexto de diversidad todos los loci son muy parecidos. Por lo general, los grupos étnicos conformados por mestizos presentan una alta diversidad como consecuencia de los diferentes orígenes ancestrales y geográficos a los que pertenecen y como consecuencia se tiene una alta presencia de heterogeneidad (Baeta et al., 2012, p.300-302; Kay et al., 2015; Tang et al., 2006).

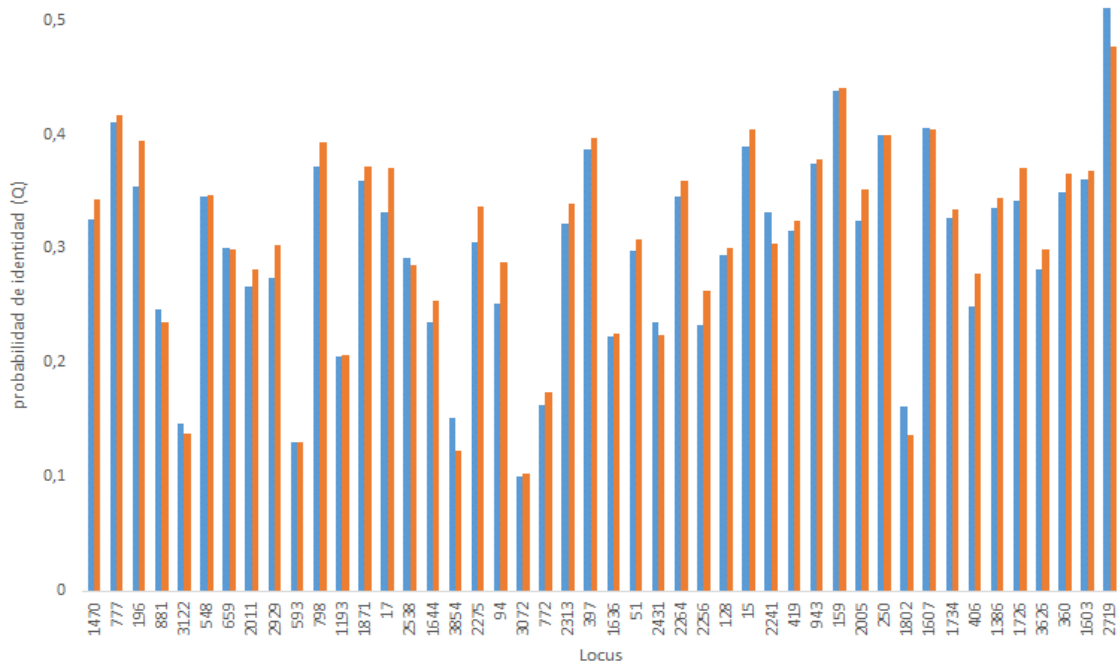


Figura 22. Valores de 1-Qintra (diversidad génica intraindividuos, mostrado en azul) y 1-Qinter (diversidad génica interindividuos-intrapoblaciones, mostrado en color anaranjado) para las cinco poblaciones analizadas, determinados mediante un ANOVA.

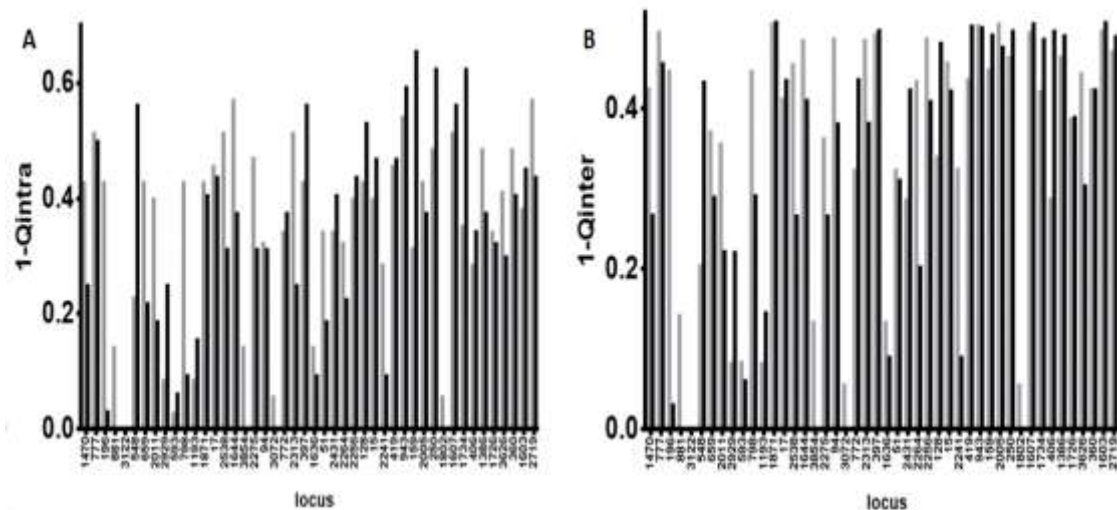


Figura 23. Diferenciación génica de los individuos de las muestras poblacionales conformadas por los pacientes afectados con la enfermedad de Huntington (barras en negro) y los controles conformados por individuos aparentemente sanos (barras mostradas en gris).

A) 1-Qintra (diversidad génica intraindividuos)

B) 1-Qinter (diversidad génica interindividuos-intrapoblación)

6. CONCLUSIONES Y RECOMENDACIONES

6.1 Conclusiones

La población conformada por los pacientes afectados con la enfermedad de Huntington, se divide en tres subpoblaciones ($K=3$) que corresponden a los orígenes ancestrales de la población ecuatoriana (africano, amerindio y europeo).

Desde el punto de vista ancestral el grupo conformado por los pacientes está enmarcado fundamentalmente en los orígenes nativos-americanos (83,7 %) seguido en menor proporción por el origen europeo (14,5 %).

La diferenciación genética (medida mediante el estadístico F_{ST}) entre los pacientes fue mayor con la población de referencia de África ($F_{ST}=0,39$), la diferenciación con la población de Europa también fue elevada ($F_{ST}=0,21$) y baja-moderada en relación a las poblaciones de origen nativo-americano ($F_{ST}=0,07$) y los controles ($F_{ST}=0,05$) esto se puede deber al origen ancestral de la muestra, a su complejidad y tamaño.

A pesar de que investigaciones realizadas indican que el origen de la enfermedad es africano este componente ancestral tuvo poca contribución en la estructura genética de los pacientes con Huntington estudiados.

6.2 Recomendaciones

Evaluar una muestra poblacional conformada por mayor cantidad de pacientes que vinculen individuos afrodescendientes con el fin de ampliar el panorama de la estructura genética poblacional.

Incorporar al estudio individuos de los distintos grupos étnicos y sub-étnicos del Ecuador que se originen ancestralmente de grupos indígenas propios del país, con el fin de obtener resultados representativos.

Evaluar marcadores genéticos que estén directamente relacionados con el desarrollo de la enfermedad de Huntington para evidenciar INDELS relacionados con la misma.

REFERENCIAS

- Aguirre, Z. (2012). Diversidad Étnica-Cultural del Ecuador. In *Estudios Universitarios* (Vol. 2002010010). Loja-Ecuador. Recuperado el 16 de Octubre del 2016 de <https://dspace.unl.edu.ec/jspui/bitstream/123456789/445/1/DIVERSIDAD%20ETNICA.pdf>.
- Andrew, S., Theilmann, J., Almqvist, E., Norremolle, A., Lucotte, G., Anvret, M., ... Hayden, M. R. (1993). DNA analysis of distinct populations suggests multiple origins for the mutation causing Huntington disease. *Clinical Genetics*, 43(6). <https://doi.org/10.1111/j.1399-0004.1993.tb03820.x>
- Aparicio, M., & Lucas, J. (2002). Molecular basis of Huntington's disease and possible pathogenic mechanisms. *Revista de Neurologia*, 35(3). <http://www.ncbi.nlm.nih.gov/pubmed/12235581>
- Applied Biosystems. (2005). GeneScan™ 600 LIZ © Size Standard v2.0. USA.
- Aquino, J., Jannuzzi, J., Carvalho, E., & Gusmão, L. (2015). Assessing the suitability of different sets of InDels in ancestry estimation. *Forensic Science International: Genetics Supplement Series*, 5, e34–e36. <https://doi.org/http://dx.doi.org/10.1016/j.fsigss.2015.09.014>
- Aronin, N., Chase, K., Young, C., Sapp, E., Schwarz, C., Matta, N., ... Beal, M. F. (1995). CAG expansion affects the expression of mutant Huntingtin in the Huntington's disease brain. *Neuron*, 15(5). [https://doi.org/10.1016/0896-6273\(95\)90106-X](https://doi.org/10.1016/0896-6273(95)90106-X)
- Ayala, E. (2008). *Resumen de historia del Ecuador. Biblioteca General Cultural* (Vol. 8). Quito-Ecuador. Recuperado el 10 de octubre de 2016 de <http://repositorio.uasb.edu.ec/bitstream/10644/836/1/AYALAE-CON0001-RESUMEN.pdf>

- Baeta, M., Núñez, C., Sosa, C., Bolea, M., Casalod, Y., González-Andrade, F., ... Martínez-Jarreta, B. (2012). Mitochondrial diversity in Amerindian Kichwa and Mestizo populations from Ecuador. *International Journal of Legal Medicine*, 126(2). <https://doi.org/10.1007/s00414-011-0656-4>
- Bardien, S., Abrahams, F., Soodyall, H., Van der Merwe, L., Greenberg, J., Brink, T., & Carr, J. (2007). A South African mixed ancestry family with Huntington disease-like 2: Clinical and genetic features. *Movement Disorders*, 22(14). <https://doi.org/10.1002/mds.21672>
- Bates, G. (2003). Huntingtin aggregation and toxicity in Huntington's disease. *The Lancet*, 361(9369). [https://doi.org/10.1016/S0140-6736\(03\)13304-1](https://doi.org/10.1016/S0140-6736(03)13304-1)
- Bates, G., Tabrizi, S., & Jones, L. (2014). Clinical Aspects of Huntington's Disease. In *Huntington's Disease* (Oxford Uni). UK. [https://books.google.com.ec/books?hl=es&lr=&id=i4kAwAAQBAJ&oi=fnd&pg=PP1&dq=Bates+G,+Harper+P,+Jones+L.+\(2002\).+Huntington+Disease.+Oxford:+Oxford+University.&ots=1d-fxM_MIJ&sig=6ktmNy-YJG3Lx_ss-3bKrrTyINA#v=onepage&q=Bates G%2C Harper P%2C Jones L. \(2002\).+Huntington+Disease.+Oxford:+Oxford+University.](https://books.google.com.ec/books?hl=es&lr=&id=i4kAwAAQBAJ&oi=fnd&pg=PP1&dq=Bates+G,+Harper+P,+Jones+L.+(2002).+Huntington+Disease.+Oxford:+Oxford+University.&ots=1d-fxM_MIJ&sig=6ktmNy-YJG3Lx_ss-3bKrrTyINA#v=onepage&q=Bates%20Harper%20Jones%20L.%20(2002).+Huntington+Disease.+Oxford:+Oxford+University.)
- Boll, T., Heaton, R., & Reitan, R. (1974). Neuropsychological and emotional correlates of Huntington's Chorea. *The Journal of Nervous and Mental Disease*, 158(1).
- Bryc, K., Velez, C., Karafet, T., Moreno, A., Reynolds, A., Auton, A., ... Ostrer, H. (2010). Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proceedings of the National Academy of Sciences*, 107(Supplement_2). <https://doi.org/10.1073/pnas.0914618107>
- Cabrera, A., Paz-Y-Miño, C., Burgos, G., López-Cortés, A., Herrera, C., Gaviria, A., & Tejera, E. (2016). A study of the molecular variants associated with lactase persistence in different Ecuadorian ethnic groups. *American Journal of Human Biology*, 0(March). <https://doi.org/10.1002/ajhb.22865>
- Campbell, M., & Tishkoff, S. (2008). African genetic diversity: implications for

human demographic history, modern human origins, and complex disease mapping. *Annual Review of Genomics and Human Genetics*, 9. <https://doi.org/10.1146/annurev.genom.9.081307.164258.AFRICAN>

Cárdenas, J., Heinz, T., Pardo, J., Álvarez, V., Taboada, P., Sánchez, P., ... Salas, A. (2015). The multiethnic ancestry of Bolivians as revealed by the analysis of Y-chromosome markers. *Forensic Science International: Genetics*, 14. <https://doi.org/10.1016/j.fsigen.2014.10.023>

Cattaneo, E., Zuccato, C., & Tartari, M. (2005). Normal huntingtin function: an alternative approach to Huntington's disease. *Nature Reviews Neuroscience*, 6(12). <https://doi.org/10.1038/nrn1806>

Cazeneuve, C., & Durr, A. (2014). Genetic of Huntingtons Disease. In G. Bates, S. Tabrizi, & L. Jones (Eds.), *Huntington's Disease* (4th ed.). New York: Oxford Monographs on medical genetics. Recuperado el 06 de Septiembre del 2016 de [https://books.google.com.ec/books?id=_ifAgAAQBAJ&pg=PT168&lpg=PT168&dq=Huntingtons+ancestry&source=bl&ots=0A6tHaMEHb&sig=vHTJl_JPAR17KPRIB3pULJtLnXw&hl=es&sa=X&ved=0ahUKEwj8l9KbnrjQAhUF6yYKHXBzBE4Q6AEITzAG#v=onepage&q=Huntingtons ancestry&f=false](https://books.google.com.ec/books?id=_ifAgAAQBAJ&pg=PT168&lpg=PT168&dq=Huntingtons+ancestry&source=bl&ots=0A6tHaMEHb&sig=vHTJl_JPAR17KPRIB3pULJtLnXw&hl=es&sa=X&ved=0ahUKEwj8l9KbnrjQAhUF6yYKHXBzBE4Q6AEITzAG#v=onepage&q=Huntingtons%20ancestry&f=false)

Crespillo, M., Bañón, R., & Valverde, J. (2011). Aprendizaje y reflexiones de la identificación de cadáveres mediante marcadores genéticos monoparentales (ADN mitocondrial, cromosoma Y). A propósito de un caso. *Revista Española de Medicina Legal*, 37(1). [https://doi.org/10.1016/S0377-4732\(11\)70056-6](https://doi.org/10.1016/S0377-4732(11)70056-6)

Cuenca, P., & Morales, F. (1999). Mutaciones inestables: Causa de algunas enfermedades neurológicas hereditarias. *Acta Médica Costarricense*.

DeGiorgio, M., Jakobsson, M., & Rosenberg, N. a. (2009). Explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *Proceedings of the National Academy of Sciences*, 106(38).

<https://doi.org/10.1073/pnas.0903341106>

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

Enoch, M. (2006). Using ancestry-informative markers to define populations and detect population stratification. *Journal of Psychopharmacology*, 20(4 suppl). <https://doi.org/10.1177/1359786806066041>

Eurostemcell. (2013). Eurostemcell. Recuperado el 10 de Octubre de 2016 <http://www.eurostemcell.org/factsheet/huntington's-disease-how-could-stem-cells-help>

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

Falush, D., Wirth, T., Linz, B., Pritchard, J., Stephens, M., Kidd, M., ... Suerbaum, S. (2003). Traces of human migrations in *Helicobacter pylori* populations. *Science (New York, N.Y.)*, 299(5612). <https://doi.org/10.1126/science.1080857>

Faria, A., Ratnanather, J., Tward, D., Lee, D., Van Den Noort, F., Wu, D., ... Miller, M. (2016). Linking white matter and deep gray matter alterations in premanifest Huntington disease. *NeuroImage: Clinical*, 11. <https://doi.org/10.1016/j.nicl.2016.02.014>

Fernandes, J. (2001). La enfermedad de huntington: Una visión biomolecular. *Revista de Neurologia*, 32(8).

Ferreira, C., García, K., Macías, L., Pérez, A., & Tomsich, C. (2013). *Mujeres y hombres del Ecuador en Cifras III serie información estratégica*. Inec. Quito-Ecuador: Editorial Ecuador.

- Ferriol, M., Picó, B., & Nuez, F. (2003). Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. *TAG Theoretical and Applied Genetics*, *107*(2). <https://doi.org/10.1007/s00122-003-1242-z>
- Fumagalli, D., Desmedt, C., Ignatiadis, M., Loi, S., Piccart, M., & Sotiriou, C. (2011). Gene Profiling Assay and Application: The Predictive Role in Primary Therapy. *JNCI Monographs*, *2011*(43). <https://doi.org/10.1093/jncimonographs/lgr040>
- Garavito, G., Martinez, B., Builes, J., Aguirre, D., Mendoza, L., Afanador, C., ... Marrugo, J. (2015). Indels markers set and ancestry estimates in a population sample from Atlantic Department of Colombia. *Forensic Science International: Genetics Supplement Series*, *5*, e177–e178. <https://doi.org/10.1016/j.fsigss.2015.09.071>
- Griffiths, J., Miller, R., Suzuki, T., Lewontin, M., & Gelbart, L. (2002). Patrones de Herencia. In *Genética* (7a edición). McGraw-Hill Interamericana. Recuperado el 08 de Octubre del 2016 http://baunne.unne.edu.ar/material_griffiths/Genetica_7a_Ed_-_Cap_02.pdf
- González, F., Sánchez, D., González, J., Gascón, S., & Martínez, B. (2007). Sex-Specific Genetic Admixture of Mestizos, Amerindian Kichwas, and Afro-Ecuadorans from Ecuador. *Human Biology*, *79*(1). <https://doi.org/10.1353/hub.2007.0024>
- Guerra, R., & Yu, Z. (2006). Single Nucleotide Polymorphisms and Their Applications. In *Computational and Statistical Approaches to Genomics*. Boston, MA: Springer US. https://doi.org/10.1007/0-387-26288-1_16
- Halder, I., Shriver, M., Thomas, M., Fernandez, J., & Frudakis, T. (2008). A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. *Human Mutation*, *29*(5). <https://doi.org/10.1002/humu.20695>

- Hardy, G. (1908). Mendelian proportions in a mixed population. *Science*, 28(706). <https://doi.org/10.1126/science.28.706.49>
- Hayden, M., & Fremer, B. (2000). The metabolic & molecular bases of inherited disease. In *Huntington's disease*.
- HDSA. (2016). Huntington's Disease Society of America. <http://hdsa.org>
- Hincapié, M., Gil, A., Pico, A., Gusmão, L., Rondón, F., Vargas, C., & Castillo, A. (2009). Análisis de la estructura genética en una muestra poblacional de Bucaramanga , departamento de Santander * Colombia Médica. *Colombia Medica*, 40(4).
- Hughes, R., & Lo, D. (2010). Huntington's disease: clinical features and routes to therapy. In *Neurobiology of Huntington's Disease: Applications to Drug Discovery* (CRC PRESS).
- Huntington, G. (1872). On Chorea. *The Medical and Surgical Reporter*.
- Jakobsson, M., & Rosenberg, N. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14). <https://doi.org/10.1093/bioinformatics/btm233>
- Jarve, M. (2012). *Different genetic perspectives on human history in Europe and the Caucasus: the stories told by uniparental and autosomal markers*. Tartu University Press. Recuperado el 12 de Octubre del 2016 http://dspace.ut.ee/bitstream/handle/10062/25366/jarve_mari.pdf?sequence=1&isAllowed=y
- Jorde, L. (2001). Population genomics: a bridge from evolutionary history to genetic medicine. *Human Molecular Genetics*, 10(20). <https://doi.org/10.1093/hmg/10.20.2199>
- Kay, C., Collins, J., Skotte, N., Southwell, A., Warby, S., Caron, N., ... Hayden, M. (2015). Huntingtin Haplotypes Provide Prioritized Target Panels for

- Allele-specific Silencing in Huntington Disease Patients of European Ancestry. *Molecular Therapy*, 23(11). <https://doi.org/10.1038/mt.2015.128>
- Kirkwood, S., Su, J., Conneally, P., & Foroud, T. (2001). Progression of Symptoms in the Early and Middle Stages of Huntington Disease. *Archives of Neurology*, 58(2). <https://doi.org/10.1001/archneur.58.2.273>
- Krause, A., Mitchell, C., Essop, F., Tager, S., Temlett, J., Stevanin, G., ... Margolis, R. (2015). Junctophilin 3 (JPH3) expansion mutations causing Huntington disease like 2 (HDL2) are common in South African patients with African ancestry and a Huntington disease phenotype. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 168(7). <https://doi.org/10.1002/ajmg.b.32332>
- Kwok, P., & Chen, X. (2003). Detection of single nucleotide polymorphisms. *Curr Issues Mol Biol*, 5(2). Recuperado el 10 de Octubre del 2016 de http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12793528
- Landles, C., & Bates, G. (2004). Huntingtin and the molecular pathogenesis of Huntington's disease. *EMBO Reports*, 5(10). <https://doi.org/10.1038/sj.embor.7400250>
- Lazar, A., Panin, F., Goodman, A., Lazic, S., Lazar, Z., Mason, S., ... Barker, R. (2015). Sleep deficits but no metabolic deficits in premanifest Huntington's disease. *Annals of Neurology*, 78(4). <https://doi.org/10.1002/ana.24495>
- Li, J., Absher, D., Tang, H., Southwick, A., Casto, A., Ramachandran, S., ... Myers, R. (2008). Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. *Science*, 319(5866). <https://doi.org/10.1126/science.1153717>
- Liou, S. (2010). Huntington's Outreach project for education. Recuperado el 11 de Noviembre del 2016 de http://web.stanford.edu/group/hopes/cgi-bin/hopes_test/population-genetics-and-hd/

- Long, J. (1991). The genetic structure of admixed populations. *Genetics*, 127(2). Recuperado el 18 de Noviembre de <http://www.ncbi.nlm.nih.gov/pubmed/2004712>
- López, F. (Universidad C. de M. (2015). Aspectos emergentes de la biomedicina actual en torno al riesgo: un estudio antropológico sobre el consejo genético y la detección de la Corea de Huntington.
- Luthi-Carter, R. (2002). Polyglutamine and transcription: gene expression changes shared by DRPLA and Huntington's disease mouse models reveal context-independent effects. *Human Molecular Genetics*, 11(17). <https://doi.org/10.1093/hmg/11.17.1927>
- Mahjoubi, F., Montazeri, M., Zare-kahrizi, S., Nafisi, S., & Engineering, G. (2013). Employing Real Time PCR for the Diagnosis of Huntington Disease. *Zahedan Journal of Research in Medical Sciences*, 15(7).
- Margolis, R. (2003). Diagnosis of Huntington Disease. *Clinical Chemistry*, 49(10). <https://doi.org/10.1373/49.10.1726>
- Martínez, A., & Rában, A. (2002). Anatomía patológica de la enfermedad de Huntigton. *Revista Española de Patología*, 35(4).
- Martínez, L. (2011). *Estructura genetica de la población de Guatemala*. Universidad de Granada, Granada-Guatemala. Recuperado el 16 de Octubre de 2016, de <https://hera.ugr.es/tesisugr/20353716.pdf>
- Martínez, M., Sánchez, M., & Martínez, G. (2010). Resumen de la evolución de las técnicas de citogenética y genética molecular para la identificación de las alteraciones genéticas del desarrollo embrionario. *SEMERGEN - Medicina de Familia*, 36(9). <https://doi.org/10.1016/j.semerng.2010.09.002>
- Merriwether, A. (2003). The evolution and genetics of Latin American populations. *American Journal of Human Biology*, 15(4). <https://doi.org/10.1002/ajhb.10174>

- Miller, R., Phillips, M., Jo, I., Donaldson, M., Studebaker, J., Addleman, N., ... Kwok, P. (2005). High-density single-nucleotide polymorphism maps of the human genome. *Genomics*, 86(2). <https://doi.org/10.1016/j.ygeno.2005.04.012>
- Mirkin, S. (2007). Expandable DNA repeats and human disease. *Nature*, 447(7147). <https://doi.org/10.1038/nature05977>
- Moily, N., Kota, L., Anjanappa, R., Venugopal, S., Vaidyanathan, R., Pal, P., ... Kandasamy, M. (2014). Trinucleotide repeats and haplotypes at the Huntingtin locus in an Indian sample overlaps with European haplogroup A. *PLoS Currents*, 6. <https://doi.org/10.1371/currents.hd.a3ad1a381ab1eed117675145318c9a80>
- Molón, L., Yáñez, R., & López-Ibor, M. (2010). Juvenile Huntington ' s disease : a case report and literature review. *Salud Mental*, 38(5).
- Mueller, U., & Wolfenbarger, L. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution*, 14(10). [https://doi.org/10.1016/S0169-5347\(99\)01659-6](https://doi.org/10.1016/S0169-5347(99)01659-6)
- Mulligan, C., Hunley, K., Cole, S., & Long, J. (2004). Population genetics, history, and health patterns in native americans. *Annual Review of Genomics and Human Genetics*, 5(1). <https://doi.org/10.1146/annurev.genom.5.061903.175920>
- Ossa, H., Aquino, J., Sierra, S., Ramírez, A., Carvalho, E., & Gusmão, L. (2015). Analysis of admixture in Native American populations from Colombia. *Forensic Science International: Genetics Supplement Series*, 5, e332–e334. <https://doi.org/10.1016/j.fsigss.2015.09.132>
- Panagopoulos, I., Lassen, C., Kristoffersson, U., & Aman, P. (1999). A novel PCR-based approach for the detection of the Huntington disease associated trinucleotide repeat expansion. *Hum Mutat*, 13(3).

- Paz-y-Miño, C., & Burgos, G. (2015). La Historia del Ecuador contada por los genes. In *Historia de las ciencias en el Ecuador*. Quito-Ecuador: de ResearchGate. Recuperado el 17 de Julio del 2016 de https://www.researchgate.net/profile/Cesar_Paz-y-Miño/publication/287201628_La_Historia_del_Ecuador_contada_por_los_genes/links/5672cda808aee7a427435ab0.pdf?origin=publication_list
- Paz-y-Miño, C., Pavón-Realpe, V., Jaramillo, G., López, A., & Domínguez, J. (2014). Estado de la mutación del gen IT-15 (HTT) en familias ecuatorianas con enfermedad de Huntington. *Arch Neurocién*, 19(2). Recuperado el 20 de Julio del 2016 <http://www.medigraphic.com/pdfs/arcneu/ane-2014/ane142b.pdf>
- Penacino, G. (2004). Analisis de ADN: 12 Marcadores del Cromosoma Masculino en la Investigacion Criminal. In *Informedica*. Buenos Aires-Argentina.
- Peñuelasa, I., Miguela, C. de, & Gullón, A. (1997). Expansión de triplete repetidos en enfermedades neurodegenerativas hereditarias. *Med Clin (Barc)*, 108(14). Recuperado el 27 de Agosto del 2016 de http://cekm.unlugar.com/expasion_tripletes_repetidos_enf_ndg.pdf
- Pereira, R., Phillips, C., Pinto, N., Santos, C., Santos, S., Amorim, A., ... Gusmão, L. (2012). Straightforward Inference of Ancestry and Admixture Proportions through Ancestry-Informative Insertion Deletion Multiplexing. *PLoS ONE*, 7(1). <https://doi.org/10.1371/journal.pone.0029684>
- Phillips, C., Salas, A., Sánchez, J., Fondevila, M., Gómez-Tato, A., Álvarez-Dios, J., ... Carracedo, Á. (2007). Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. *Forensic Science International: Genetics*, 1(3–4). <https://doi.org/10.1016/j.fsigen.2007.06.008>
- Plagnol, V., & Wall, J. (2006). Possible ancestral structure in human populations. *PLoS Genetics*, 2(7). <https://doi.org/10.1371/journal.pgen.0020105>

- Powell, J., & Wright, S. (1979). Evolution and the Genetics of Populations, Vol. 4: Variability within and among Natural Populations. misc, JSTOR.
- Pritchard, J., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2). <https://doi.org/10.1111/j.1471-8286.2007.01758.x>
- Pritchard, J., Wen, X., & Falush, D. (2010). Documentation for structure software: Version 2.3. *Software: Practice and Experience*, 6(3). <https://doi.org/10.1002/spe.4380060305>
- Ramachandran, S., Deshpande, O., Roseman, C., Rosenberg, N., Feldman, M., & Cavalli-Sforza, L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44). <https://doi.org/10.1073/pnas.0507611102>
- Ramos, M., Trujillo-Tiebas, M., & Milá, M. (2012). Recomendaciones de buena práctica para el diagnóstico genético de la enfermedad de Huntington. *Medicina Clínica*, 138(13). <https://doi.org/10.1016/j.medcli.2011.03.001>
- Raymond, M., Rousset, F. (1995). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *J. Heredity*, 86.
- Raymond, M., & Rousset, F. (2016). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity Collections*, 107(6).
- Raymund, R. (2010). Huntington's disease: a clinical review. *Orphanet Journal of Rare Diseases*, 5(1), 40. <https://doi.org/10.1186/1750-1172-5-40>
- Rojas, W., Parra, M., Campo, O., Caro, M., Lopera, J., Arias, W., ... Bedoya, G. (2010). Genetic make up and structure of Colombian populations by means of uniparental and biparental DNA markers. *American Journal of Physical Anthropology*, 143(1). <https://doi.org/10.1002/ajpa.21270>

- Rondón, F., & Barreto, G. (2013). Estructura genética , ancestralidad y su relación con los estudios en salud humana. *Médicas Uis*, 2(1).
- Rosales, M., & Barros, P. (2008). Diagnóstico molecular de la enfermedad de Huntington. *Gaceta Médica de México*, 144(3).
- Rosenberg, N. (2002). Genetic Structure of Human Populations. *Science*, 298(5602). <https://doi.org/10.1126/science.1078311>
- Rosenberg, N., Burke, T., Elo, K., Feldman, M., Freidlin, P., Groenen, M., ... Weigend, S. (2001). Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics*, 159(2).
- Rourke, D., & Raff, J. (2010). The Human Genetic History of the Americas: The Final Frontier. *Current Biology*, 20(4). <https://doi.org/10.1016/j.cub.2009.11.051>
- Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8(1). <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Salamanca, F. (2003). Avances en el tratamiento de las enfermedades neurodegenerativas. *Gaceta Medica de Mexico*, 22(Supl 1).
- Salzano, F., & Cátira, M. (2003). *The evolution and genetics of Latin American populations* (Vol. 15). Cambridge University Press. <https://doi.org/10.1002/ajhb.10174>
- Sánchez, D., González, F., & Martínez, B. (2003). Population Genetics of 12 STR Loci in a Sample of Mestizos from Ecuador (South America). *Journal of Forensic Sciences*, 48(2), 2002218. <https://doi.org/10.1520/JFS2002218>
- Santos, N., Ribeiro, E., Ribeiro, A., Pereira, R., Gusmão, L., Amorim, A., Guerreiro, J., ... Santos, S. (2010). Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Human Mutation*, 31(2).

<https://doi.org/10.1002/humu.21159>

- Serna, Y. (2014). ¿ Qué es la Corea de Huntington (EH)? *Nosocomio*.
- Sobrino, O., De Cós Milas, A., Moreno, C., Trigo, A., Lapastora de Mingo, P., & Jiménez, M. (2007). Manifestaciones neuropsiquiátricas de la enfermedad de huntington. *Interpsiquis. 8 Congreso Virtual de Psiquiatría.*, (1).
- Song, N., Zhang, X., Han, Z., & Gao, T. (2016). Genetic differentiation between cultured and wild populations of *Paralichthys olivaceus* based on AFLP markers. *Biochemical Systematics and Ecology*, 68. <https://doi.org/10.1016/j.bse.2016.07.025>
- Stefflova, K., Dulik, M., Barnholtz-Sloan, J., Pai, A., Walker, A., & Rebbeck, T. (2011). Dissecting the Within-Africa Ancestry of Populations of African Descent in the Americas. *PLoS ONE*, 6(1). <https://doi.org/10.1371/journal.pone.0014495>
- Stout, J., Glikmann-Johnston, Y., & Andrews, S. (2016). Cognitive assessment strategies in Huntington's disease research. *Journal of Neuroscience Methods*, 265. <https://doi.org/10.1016/j.jneumeth.2015.12.007>
- Tang, H., Coram, M., Wang, P., Zhu, X., & Risch, N. (2006). Reconstructing Genetic Ancestry Blocks in Admixed Individuals. *The American Journal of Human Genetics*, 79(1). <https://doi.org/10.1086/504302>
- Tasset, I., Sánchez, F., & Túnez, I. (2009). Bases moleculares de la enfermedad de Huntington: Papel del estrés oxidativo. *Revista de Neurología*, 49(8).
- Townhill, J., Hughes, A., Thomas, B., Busse, M. E., Price, K., Dunnett, S., ... Rosser, A. E. (2016). Using Actiwatch to monitor circadian rhythm disturbance in Huntington' disease: A cautionary note. *Journal of Neuroscience Methods*, 265. <https://doi.org/10.1016/j.jneumeth.2016.01.009>

- Urbano, L., Portilla, E., Muñoz, W., Sierra, C., Bolaños, H., Arboleda, Y., ... Builes, J. (2015). Ancestral genetic composition in a population of South Western Colombian using autosomal AIM-INDELS. *Forensic Science International: Genetics Supplement Series*, 5. <https://doi.org/10.1016/j.fsigss.2015.09.076>
- Väli, Ü., Brandström, M., Johansson, M., & Ellegren, H. (2008). Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC Genetics*, 9(1). <https://doi.org/10.1186/1471-2156-9-8>
- Vásquez, M., Cuenca, P., Campos, D., Gutiérrez, B., Fernández, H., & Morales, F. (2011). Abordaje integral de pacientes costarricenses afectados con la enfermedad de Huntington y sus familiares. *Acta Méd. Costarric*, 53(3). Recuperado el 26 de Septiembre del 2016 de <http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&src=google&base=LILACS&lang=p&nextAction=lnk&exprSearch=648415&indexSearch=ID>
- Vásquez, M., & Montero, F. (2008). Diagnóstico molecular de la enfermedad de Huntington en Costa Rica. *Acta Medica Costarricense*, 50(1). Recuperado el 24 de Septiembre del 2016 de http://www.actamedica.medicos.sa.cr/index.php/Acta_Medica/article/view/350%5Cnpapers2://publication/uuid/A5EA26A1-4BB3-48FF-93DB-19E40B19F045
- Vázquez, Y., Laguna, L., Laffita, J., González, Y., Almaguer, L., Rodríguez, R., ... Velázquez, L. (2013). Genetic features of Huntington disease in Cuban population: Implications for phenotype, epidemiology and predictive testing. *Journal of the Neurological Sciences*, 335(1–2). <https://doi.org/10.1016/j.jns.2013.08.037>
- Vekemans, X., Beauwens, T., Lemaire, M., & Roldan-Ruiz, I. (2002). Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11.

- Wagner, R. (1999). Anecdotal, historical and critical commentaries on genetics. Rudolph Virchow and the genetic basis of somatic ecology. *Genetics*, 151(3).
- Warby, S., Visscher, H., Collins, J., Doty, C., Carter, C., Butland, S., ... Hayden, M. (2011). HTT haplotypes contribute to differences in Huntington disease prevalence between Europe and East Asia. *European Journal of Human Genetics*, 19(5). <https://doi.org/10.1038/ejhg.2010.229>
- Weinberg, W. (1908). *Über den nachweis der vererbung beim menschen*. BOOK, publisher not identified.
- Weinberg, W. (1912). Zur vererbung des zwergwuchses. *Arch Rass Ges Biol*, 9, 710–718. JOUR.
- Weir, B., & Cockerham, C.. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, 38(6). <https://doi.org/10.2307/2408641>
- Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16(2).
- Wright, S. (1949). The genetical structure of populations. *Annals of Eugenics*, 15(1). <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Yhnell, E., Lelos, M., Dunnett, S., & Brooks, S. (2016). Cognitive training modifies disease symptoms in a mouse model of Huntington's disease. *Experimental Neurology*, 282. <https://doi.org/10.1016/j.expneurol.2016.05.008>

ANEXOS

Anexo 1. Ejemplo de la tabla de frecuencias alélicas de los pacientes obtenida a partir del programa GeneMapper®.

Alust./Mar	1470	777	196	881	3122	548	659	2011	2929	593	790	1193	1871	17	2538	1644	3854	2275	94	3072	772	2313	397	1636	51	2431	2264	2256	128	15	2241	419		
PCVCH	2	1	2	?	1	1	2	1	1	2	1	1	1	1	2	2	2	2	1	1	1	1	1	1	1	1	2	1	1	1	2	1		
	2	1	2	?	1	2	2	1	1	2	2	2	1	2	2	2	2	2	2	1	2	1	2	1	1	1	1	2	1	1	2	2	2	
OMMCH	1	1	2	?	1	1	2	1	1	2	1	2	2	1	2	1	2	1	2	1	1	1	1	1	2	2	2	1	1	1	2	1		
	2	2	2	?	1	2	2	2	1	2	1	2	2	1	2	1	2	2	2	1	2	1	2	1	2	2	2	2	2	2	2	2		
PCCH2	2	1	2	?	1	2	2	1	1	2	2	2	2	1	2	1	2	2	1	1	2	2	1	1	1	1	2	1	1	1	2	2		
	2	1	2	?	1	2	2	1	1	2	2	2	2	1	2	2	2	2	2	1	2	2	2	1	1	1	1	2	1	1	1	2	2	
OJCH	1	1	2	?	1	2	1	1	1	2	2	2	1	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	1	
	2	2	2	?	1	2	2	2	2	2	2	2	2	1	2	2	2	2	2	1	1	2	2	1	2	2	2	2	2	2	2	2	1	
MLCH	2	1	2	?	1	1	1	1	1	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	2	1	2	1	2	1	
	2	1	2	?	1	2	2	1	2	2	2	2	2	1	2	2	2	2	2	1	1	1	1	2	1	1	2	2	2	2	1	2	2	
APCCH	2	1	2	?	1	1	2	1	1	2	2	2	1	1	1	2	2	1	1	1	1	1	1	1	1	2	2	1	1	2	2	2		
	2	2	2	?	1	2	2	1	1	2	2	2	1	2	2	2	2	2	2	1	2	1	2	1	1	2	2	1	2	2	2	2	2	
SMCHCH	2	2	2	1	1	1	2	1	1	2	2	2	2	2	2	2	2	2	2	1	2	1	2	1	1	1	2	2	1	2	2	2	1	
	2	2	2	1	1	2	2	1	1	2	2	2	2	2	2	2	2	2	1	1	2	1	2	1	1	2	2	2	2	2	2	2	2	
MGDHTT	2	1	2	1	1	1	2	1	1	2	2	2	1	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2	1	1	1	2	1	
	2	2	2	?	1	2	2	1	1	2	2	2	1	1	2	2	2	2	2	1	1	2	1	1	1	1	2	2	1	2	2	2	1	
SICH	2	1	2	?	1	2	1	1	1	2	2	2	1	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	
	2	1	2	?	1	2	2	1	1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	2	1	1	2	2	1	2	1	2	2	
VLCH2	1	2	2	?	1	2	2	2	1	2	2	1	2	1	2	1	2	2	2	1	2	2	2	1	2	2	2	2	2	2	2	2	1	
	1	2	2	?	1	2	2	2	2	2	2	2	2	2	2	1	2	2	2	1	2	2	2	1	2	2	2	2	2	2	2	2	1	
AACH	2	1	2	1	1	1	2	1	1	2	2	2	1	1	2	1	2	2	1	1	1	1	2	1	1	1	1	2	1	1	1	2	1	
	2	2	2	1	1	2	2	1	2	2	2	2	1	2	2	2	2	2	1	2	1	2	1	1	2	1	2	1	2	1	2	2	2	
XCHT	1	1	2	?	1	2	2	1	1	2	2	2	2	1	2	2	2	2	2	1	1	1	1	1	1	1	1	2	1	2	1	2	2	
	2	2	2	?	1	2	2	1	1	2	2	2	2	1	2	2	2	2	2	1	1	1	2	1	1	2	2	2	2	2	1	2	2	
AMCOHTT	2	1	2	1	1	1	2	1	1	2	1	2	2	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	
	2	2	2	1	1	2	2	2	1	2	2	2	2	2	2	2	2	2	2	1	1	1	2	2	2	2	2	2	2	1	1	2	2	2
RCHT	2	1	2	?	1	2	2	1	1	2	1	2	1	1	1	1	2	2	1	1	1	2	1	1	1	1	1	1	1	1	2	1	2	1
	2	1	2	?	1	2	2	1	1	2	2	2	1	2	2	2	2	2	2	1	2	2	2	1	1	1	2	2	1	2	2	2	2	
BSCH	2	1	2	1	1	1	1	1	1	2	2	2	1	1	1	2	2	2	1	2	1	2	1	2	1	1	1	2	1	1	1	2	2	
	2	2	2	1	1	1	1	1	2	2	2	2	2	1	2	2	2	2	1	1	2	2	2	2	1	1	2	2	1	2	1	2	2	
PDAHTT	2	1	2	1	1	1	2	1	1	2	2	2	1	1	2	2	2	2	1	1	1	1	1	1	1	1	2	2	1	1	1	2	1	
	2	2	2	1	1	2	2	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	2	2	1	1	2	2	1	2	1	2	2	
CGDHTT	2	1	2	1	1	1	2	1	1	2	2	2	2	1	2	2	2	1	1	1	1	1	1	1	1	1	2	2	1	1	1	2	1	
	2	2	2	1	1	2	2	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	
DCCH	2	1	2	1	1	2	2	1	1	1	2	1	2	1	2	2	2	2	2	1	1	1	1	1	1	1	1	2	2	1	1	2	1	
	2	1	2	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2	2	1	1	1	2	1	2	2	2	2	1	2	2	2	2	
JCCH	1	1	2	?	1	1	2	1	1	2	2	1	2	2	1	2	2	2	1	1	1	1	1	1	1	1	2	2	1	1	1	2	1	
	2	2	2	?	1	2	2	1	1	2	2	2	2	2	2	2	2	2	2	1	1	2	2	1	1	1	2	2	1	2	1	2	1	
JPCH	2	1	2	?	1	2	1	1	1	2	2	2	1	1	2	1	2	2	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	
	2	1	2	?	1	2	2	1	1	2	2	2	2	1	2	2	2	2	1	1	2	1	2	1	1	2	2	2	2	1	2	2	2	
JECGHT	2	1	2	?	1	1	2	1	1	2	2	1	2	1	2	2	2	2	2	1	1	1	1	1	1	1	2	2	1	1	1	2	1	
	2	1	2	?	1	2	2	1	1	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	1	1	2	2	2	2	2	2	1	
LCCH	1	1	1	?	1	1	2	1	1	2	1	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1
	2	1	2	?	1	2	2	2	1	2	1	2	2	2	2	2	2	2	2	1	1	2	1	1	1	1	2	2	2	2	2	2	2	
MCHT	2	1	2	?	1	1	2	1	1	2	2	2	1	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2	1	1	1	2	1	
	2	2	2	?	1	2	2	1	1	2	2	2	1	1	2	2	2	2	1	1	2	1	1	1	1	2	2	1	2	2	2	2	1	

Anexo 2. Ejemplo de la matriz Q obtenida a partir del programa STRUCTURE. Proporciones de cada individuo para los tres clústeres inferidos.

```

Inferred ancestry of individuals:
Label (%Miss) Pop: Inferred clusters
1 1031 (0) 1 : 0.134 0.049 0.817
2 1028 (0) 1 : 0.059 0.048 0.894
3 1030 (0) 1 : 0.060 0.099 0.842
4 1034 (0) 1 : 0.048 0.081 0.871
5 1035 (0) 1 : 0.063 0.101 0.837
6 452 (0) 1 : 0.001 0.002 0.997
7 453 (0) 1 : 0.001 0.001 0.998
8 454 (0) 1 : 0.001 0.001 0.999
9 455 (0) 1 : 0.001 0.004 0.995
10 457 (0) 1 : 0.003 0.004 0.993
11 458 (0) 1 : 0.001 0.001 0.998
12 459 (0) 1 : 0.001 0.001 0.999
13 460 (0) 1 : 0.002 0.003 0.995
14 461 (0) 1 : 0.001 0.001 0.999
15 464 (0) 1 : 0.002 0.003 0.996
16 465 (0) 1 : 0.001 0.001 0.998
17 466 (2) 1 : 0.001 0.001 0.998
18 469 (0) 1 : 0.002 0.001 0.997
19 470 (0) 1 : 0.001 0.001 0.999
20 472 (0) 1 : 0.001 0.001 0.998
21 473 (0) 1 : 0.001 0.001 0.998
22 475 (0) 1 : 0.002 0.028 0.970
23 479 (0) 1 : 0.001 0.001 0.998
24 985 (0) 1 : 0.001 0.005 0.994
25 986 (0) 1 : 0.001 0.001 0.998
26 1086 (0) 1 : 0.001 0.002 0.997
27 1090 (0) 1 : 0.001 0.001 0.998
28 1094 (0) 1 : 0.001 0.001 0.998
29 449 (0) 1 : 0.002 0.002 0.997
30 450 (0) 1 : 0.001 0.001 0.998
31 456 (0) 1 : 0.001 0.001 0.998
32 462 (0) 1 : 0.001 0.001 0.998
33 463 (0) 1 : 0.014 0.015 0.971
    
```

Anexo 3. Resultados del análisis de los Fis obtenidos por los tres algoritmos evaluados mediante los métodos W&C y R&H.

	Controles		Pacientes	
	44,000	44,000	40,000	40,000
Número de valores				
Mínimo	-0,294	-0,297	-0,331	-0,335
25% Percentil	-0,098	-0,099	-0,111	-0,112
Mediana	-0,033	-0,033	0,006	0,007
75% Percentile	0,125	0,127	0,170	0,173
Maximo	0,660	0,676	0,679	0,698
Media	0,012	0,013	0,029	0,030
desviación estándar de la media	0,176	0,179	0,208	0,212
Error estándar de la media	0,027	0,027	0,033	0,034
máximo con 95 % CI media	-0,041	-0,042	-0,038	-0,038
mínimo con 95 % CI media	0,066	0,067	0,096	0,098
Suma	0,541	0,569	1,161	1,206

Anexo 4. Pruebas de normalidad de los datos correspondientes a los FIS de los grupos comprendidos por los controles y los pacientes.

	controles		pacientes	
	44	44	40	40
Number of values				
Minimum	-0,294	-0,2971	-0,3312	-0,3349
25% Percentile	-0,09817	-0,09927	-0,1106	-0,1124
Median	-0,03297	-0,03345	0,00645	0,00665
75% Percentile	0,1253	0,1274	0,1698	0,173
Maximum	0,66	0,676	0,6793	0,6976
Mean	0,01228	0,01293	0,02903	0,03015
Std. Deviation	0,1758	0,179	0,2083	0,2123
Std. Error of Mean	0,02651	0,02699	0,03293	0,03366
Lower 95% CI of mean	-0,04118	-0,0415	-0,03758	-0,03773
Upper 95% CI of mean	0,06575	0,06736	0,09564	0,09804
D'Agostino & Pearson omnibus normality test				
K2	17,01	17,53	5,719	6,096
P value	0,0002	0,0002	0,0573	0,0475
Passed normality test (alpha=0.05)?	No	No	Yes	No
P value summary	***	***	ns	*
Shapiro-Wilk normality test				
W	0,921	0,9188	0,9664	0,9648
P value	0,0052	0,0043	0,2744	0,2435
Passed normality test (alpha=0.05)?	No	No	Yes	Yes
P value summary	**	**	ns	ns
KS normality test				
KS distance	0,155	0,1557	0,07824	0,07657
P value	0,0097	0,0091	0,2	0,2
Passed normality test (alpha=0.05)?	No	No	Yes	Yes
P value summary	**	**	ns	ns
Sum	0,5405	0,569	1,181	1,206

Anexo 5. Prueba de *Student* para el grupo conformado por los controles.

Table Analyzed	Controles Fis
Column B	R & H
vs.	vs,
Column A	W & C
Wilcoxon matched-pairs signed rank test	
P value	0,8016
Exact or approximate P value?	Exact
P value summary	ns
Significantly different? (P < 0.05)	No
One- or two-tailed P value?	Two-tailed
Sum of positive, negative ranks	517,0 , -473,0
Sum of signed ranks (W)	44
Median of differences	
Median	-0,0002167
How effective was the pairing?	
rs (Spearman)	1
P value (one tailed)	
P value summary	
Significant correlation? (P > 0.05)	

Anexo 6. Prueba de *Student* para el grupo conformado por los pacientes

Table Analyzed	pacientes Fis
Column B	R & H
vs.	vs,
Column A	W & C
Paired t test	
P value	0,094
P value summary	ns
Significantly different? (P < 0.05)	No
One- or two-tailed P value?	Two-tailed
t, df	t=1,717 df=39
Number of pairs	40
How big is the difference?	
Mean of differences	0,001124
SD of differences	0,004142
SEM of differences	0,0006548
95% confidence interval	-0,0002003 to 0,002449
R square	0,07026
How effective was the pairing?	
Correlation coefficient (r)	1
P value (one tailed)	< 0,0001
P value summary	****
Significant correlation? (P > 0.05)	No



D. M. Quito, 16 de diciembre de 2016

Estefanía Nathaly Eras
Estudiante investigador

De mis consideraciones:

Por medio de la presente el Comité de Ética y Bioética para la Investigación de la Universidad de Las Américas le informa que el protocolo y los instrumentos de investigación del proyecto de titulación **"Determinación de los niveles preliminares de ancestría en pacientes de la población ecuatoriana con mutación en el gen IT-15"**, con fecha de hoy 16 de diciembre de 2016, han sido aprobados por el Comité.

Esta aprobación, tiene una duración de un año, después del cual se debe pedir extensión si fuera necesaria.

En toda correspondencia con el Comité de Ética y Bioética para la investigación, favor referirse al siguiente código de aprobación: 2016-1206.

El Comité estará dispuesto a lo largo de la implementación del estudio a responder tanto a los participantes como a la investigadora en relación a cualquier inquietud que pudiere surgir. Es importante remarcar que cualquier novedad debe ser comunicada al Comité; todo cambio en el protocolo de investigación implica que debe pedir una nueva aprobación al CEBE. Cualquier evento adverso debe ser informado al Comité dentro de un máximo de 24 horas luego de ocurrido.

El Comité de Ética y Bioética para la Investigación de la Universidad de Las Américas ha otorgado esta probación con base en la información entregada por los solicitantes, quienes al presentarla asumen la veracidad, corrección y autoría de los documentos entregados. Los solicitantes de la aprobación son los responsables de aplicar y respetar la información, procedimientos y condiciones expresados en estos documentos aprobados por el Comité; también son responsables de respetar la legislación vigente aplicable y los estándares nacionales e internacionales en la materia.

Atentamente,

Diego R. Chauvin P.
Presidente del CEBE-UDLA



