

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ciencias y Humanidades



Validación de un kit comercial para la genotipificación de ADN
mitocondrial para la realización de un estudio poblacional que incluye
los grupos etnolingüísticos achí, ixil, kaqchikel, ladino, queqchí y quiché
en Guatemala a partir de muestras referenciales

Trabajo de graduación en modalidad de tesis presentado por

Alicia María Mazariegos Medina

para optar al grado académico de Licenciada en Bioquímica y Microbiología

Guatemala,

2019

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ciencias y Humanidades



Validación de un kit comercial para la genotipificación de ADN mitocondrial para la realización de un estudio poblacional que incluye los grupos etnolingüísticos achí, ixil, kaqchikel, ladino, queqchí y quiché en Guatemala a partir de muestras referenciales

Trabajo de graduación en modalidad de tesis presentado por

Alicia María Mazariegos Medina

para optar al grado académico de Licenciada en Bioquímica y Microbiología

Guatemala,

2019

Vo. Bo. :

(f) 

Wendy Carolina Argueta Marroquín

Tribunal Examinador:

(f) 

Wendy Carolina Argueta Marroquín

(f) 

Edwin David Castañeda

(f) 

Krisztina Fulop Rios Gonzalez

Fecha de aprobación: Guatemala, 2 de diciembre de 2019

CONTENIDO

Lista de cuadros	iii
Lista de figuras	iv
Resumen	v
I. Introducción.....	1
II. Objetivo	2
A. Objetivo general.....	2
B. Objetivos específicos.....	2
III. Justificación	3
IV. Marco teórico.....	5
A. Ácido Desoxirribonucleico.....	5
B. Genotipificación de ADNmt.....	6
C. ADN ancestral	8
D. Toma de muestra.....	9
E. Secuenciación	10
F. Validación de método	10
V. Metodología.....	13
A. Sitio de estudio	13
B. Sujetos de estudio	13
C. Diseño, enfoque y tipo de investigación.....	13
D. Tipo y tamaño de muestra.....	13
E. Criterios de inclusión y exclusión.....	14
F. Variables	15
G. Instrumentos de medición.....	15
H. Metodología.....	16
VI. Resultados.....	19
A. Exactitud	19
B. Precisión	19
C. Sensibilidad.....	19
D. Estabilidad	22

E.	Repetibilidad.....	22
F.	Reproducibilidad.....	23
G.	Contaminación.....	23
H.	Estudio poblacional	25
VII.	Análisis de resultados.....	27
VIII.	Conclusiones	31
IX.	Recomendaciones	32
X.	Bibliografía.....	34
XI.	Anexos	38
A.	Anexo.....	38
B.	Anexo.....	52
C.	Anexo.....	58

LISTA DE CUADROS

Cuadro	Página
1. Contenido de ADN en muestras biológicas.	5
2. Criterios de inclusión y exclusión.	14
3. Variables.	15
4. Haplotipos generados en laboratorio externo para la muestra M2 y M3.	19
5. Haplotipo generado en la FAFG para la muestra M2 y M3.	19
6. Haplotipos para muestras de hisopado bucal tomadas en diferentes años (extracción por Chelex).	20
7. Haplotipo para muestras de diferente sustrato de un mismo donante (extracción por QIAamp).	20
8. Haplotipos para extractos óseos con diferente cuantificación y perfil genético reportado.	21
9. Haplotipo para mezcla de muestras.	24
10. Parámetros forenses e índices de diversidad de secuencias de las regiones HVI y II de ADNmt para grupos etnolingüísticos guatemaltecos.	26
11. Haplotipos para diferentes concentraciones de una muestra ósea.	58
12. Haplotipos de muestras procesadas por analistas diferentes.	58

LISTA DE FIGURAS

Figura	Página
1. Material genético en los gametocitos y cigoto.....	6
2. Muestras útiles para comparación genética de ADNmt.....	7
3. Mapa de la República de Guatemala con las áreas de muestreo.....	8
4. Procedimiento para procesamiento de muestras con ADNmt.....	16
5. Gel de agarosa 1% correspondiente a la región HVI para diferentes diluciones de una muestra ósea.....	22
6. Picos característicos de un electroferograma en caso de mezclas.....	25

RESUMEN

La Genética Forense es una disciplina que surgió con el fin de resolver casos criminales utilizando información genética dejada en escenas del crimen. Sin embargo, también ha ofrecido grandes ventajas en la identificación humana a través de comparaciones directas y filiaciones a través de familiares. Estas características aportan a la identificación de desaparecidos del Conflicto Armado Interno (CAI), ocurrido entre los años 1960 a 1996 en Guatemala. La búsqueda y recuperación de estas víctimas no ocurrió sino hasta décadas después de que ocurrieran los hechos, lo cual ha implicado trabajar con restos esqueletizados para lograr su identificación a través del ADN. Es frecuente obtener perfiles parciales o fallidos cuando se analizan los marcadores STRs de ADN nuclear provenientes de restos antiguos. Debido a esta limitación, la utilización de ADN mitocondrial (ADNmt) para complementar los estudios, ya que se presenta en mayor número de copias en las células y se conserva por más tiempo. La técnica de ADNmt mejorará los procesos de la Fundación de Antropología Forense de Guatemala, cuya misión es la identificación de los desaparecidos del CAI. El objetivo del estudio es validar los kits BigDye™ Direct Cycle Sequencing y BigDye™ XTerminator Purification en muestras humanas para la obtención del haplotipo de ADNmt. Se concluyó que el rango ideal de concentración de ADN para la secuenciación y que es posible la obtención del haplotipo de ADNmt a partir de hisopados bucales, frotos bucales fijados en tarjeta FTA, sangre fijada en FTA, raíz de cabello, fragmentos óseos y piezas dentales.

I. INTRODUCCIÓN

La Genética Forense ha contribuido a la identificación de víctimas a través de la comparación de los perfiles genéticos de marcadores STR autosómicos (por las siglas en inglés de repeticiones cortas en tándem) obtenidos a partir del ADN nuclear (ADNn) (Butler, 2010). Esta técnica se ve limitada ante muestras con baja cuantificación de ADN, degradación y en caso de filiaciones cuando únicamente se cuenta con familiares lejanos (Hares *et al.*, 2014; King *et al.*, 2014). Para superar estas limitaciones al trabajar con ADN ancestral, se ha recurrido a la genotipificación del ADN mitocondrial (ADNmt) (Bandelt, van Oven, & Salas, 2012; Hares *et al.*, 2014).

El ADNmt presenta un número de copias más alto que el ADNn en las células humanas (Hares *et al.*, 2014). La mitocondria contiene entre 2 a 10 copias de ADNmt, y pueden haber hasta 1000 mitocondrias por célula somática (Budowle, Allard, Wilson, & Chakraborty, 2003). Por esta razón el análisis de ADNmt es útil en el área forense cuando no se logra obtener un perfil genético de STR autosómicos satisfactorio (Hares *et al.*, 2014).

La baja fidelidad de la polimerasa de ADNmt y la aparente ausencia de mecanismos de reparación dan lugar a una tasa de mutación más alta que la del ADNn (Budowle *et al.*, 2003). El ADNmt posee regiones polimórficas conocidas como Hipervariable I, II y III (HVI, HVII y HVIII, respectivamente) (Bandelt *et al.*, 2012). Las variaciones de secuencia entre individuos se encuentran especialmente en las regiones HVI (16024-16365) y HVII (73-340) (Budowle *et al.*, 2003; Hares *et al.*, 2014). El reporte de haplotipos de ADNmt para casos forenses o estudios poblacionales requiere identificar solamente las diferencias con respecto a la versión revisada (rCRS, NC001807) de la primera secuencia de ADNmt humano (Hares *et al.*, 2014). Según este requisito, se tiene por objetivo validar un procedimiento operacional estándar que permita la correcta secuenciación y genotipificación de las regiones HVI y HVII del ADNmt de muestras forenses y reportar los haplotipos obtenidos según las recomendaciones de SWGDAM (*Scientific Working Group on DNA Analysis Methods*) (Scientific Working Group on DNA Analysis Methods, 2016).

Es posible la genotipificación de ADNmt utilizando los kits BigDye™ Direct Cycle Sequencing (Applied Biosystems™, CA, EE. UU.) y Big Dye™ XTerminator Purification (Applied Biosystems™, CA, EE. UU.) en muestras referenciales y esqueléticas con fines de identificación humana para las aplicaciones del Laboratorio de Genética Forense de la Fundación de Antropología Forense de Guatemala. El método cuenta con la información necesaria para amplificar, secuenciar y analizar los resultados para reportar haplotipos de ADNmt.

II. OBJETIVO

A. OBJETIVO GENERAL

Validar los kits BigDye™ Direct Cycle Sequencing (Applied Biosystems™, CA, EE. UU.) y BigDye™ XTerminator Purification (Applied Biosystems™, CA, EE. UU.) para la secuenciación y obtención de haplotipos del ADNmt en muestras humanas (hisopado bucal, frote bucal fijado en tarjeta FTA, sangre fijada en FTA, raíz de cabello, fragmentos óseos y piezas dentales), para el estudio poblacional de los grupos etnolingüísticos achí, ixil, kaqchikel, ladino, queqchí y quiché en Guatemala a partir de 102 muestras referenciales.

B. OBJETIVOS ESPECÍFICOS

Evaluar la exactitud y precisión de la técnica propuesta para la genotipificación de ADN mitocondrial (ADNmt), utilizando los kits BigDye™ Direct Cycle Sequencing (Applied Biosystems™, CA, EE. UU.) y BigDye™ XTerminator Purification (Applied Biosystems™, CA, EE. UU.), realizando una comparación intralaboratorio al reanalizar una muestra previamente caracterizada por un laboratorio externo al de la Fundación de Antropología Forense de Guatemala.

Determinar el rango de concentración óptimo de ADN que asegure la genotipificación del ADNmt a través del estudio de sensibilidad.

Realizar el estudio de estabilidad de la técnica para la obtención de un haplotipo de ADNmt al procesar muestras de hisopado bucal colectadas desde el año 2009 hasta el año 2018.

Evaluar la repetitividad de la técnica, procesando diversos sustratos (hisopado bucal fijado en tarjeta FTA, sangre fijada en tarjeta FTA y raíz de cabello) colectados del mismo donante y comparando el haplotipo obtenido de cada uno de ellos.

Comprobar la reproducibilidad de la técnica de genotipificación al obtener el haplotipo de las mismas muestras por dos operadores diferentes en tiempos diferentes.

Determinar cómo se vería una muestra con más de un donante de ADNmt (mezcla).

Genotipificar el ADNmt de los grupos etnolingüísticos achí, ixil, kaqchikel, ladino, queqchí y quiché (n=17 por grupo) para el estudio poblacional guatemalteco.

III. JUSTIFICACIÓN

Las investigaciones en la Fundación de Antropología Forense de Guatemala (FAFG) consisten en la identificación de víctimas del Conflicto Armado Interno (CAI). La FAFG contribuye al fortalecimiento del sistema de justicia y los derechos humanos ya que provee de evidencia que puede ser utilizada ante un proceso legal o solicitada por el Ministerio Público.

La Genética Forense contribuye a la identificación de víctimas a través de la comparación de los perfiles genéticos de marcadores STRs autosómicos obtenidos a partir del ADN nuclear (ADNn) de un indicio indubitado (se conoce al donante del ADN) y un indicio dubitado (se desconoce la identidad del donante del ADN) En el contexto del trabajo de la FAFG, las muestras indubitadas provienen de familiares que continúan en la búsqueda de sus desaparecidos y que donan una muestra genética, comúnmente llamadas muestras referenciales. Los indicios dubitados son los restos óseos de las víctimas que son recuperados de fosas clandestinas, cementerios, etc. que en su mayoría datan de la década de los años 1980. Los diferentes contextos y tiempos de deposición de los restos esqueletizados provocan la degradación del ADNn lo que dificulta la obtención de un perfil genético de STRs autosómicos completo que pueda ser comparado con los obtenidos de los grupos familiares. Un segundo reto en la investigación son las filiaciones complejas, aquellos casos en los que no se cuenta con muestras de familiares cercanos (por ejemplo, del núcleo familiar) sino de familiares en segundo o tercer grado (por ejemplo, tíos abuelos). En ambos casos, contar con una herramienta genética que pueda complementar la información generada al analizar STRs autosómicos aumentaría las probabilidades de generar una coincidencia genética y por lo tanto la identificación de una osamenta.

Cabe mencionar que, la genotipificación del ADNmt no es individualizante ya que es heredado sin eventos de recombinación entre madres e hijos. En comparación, la genotipificación del ADNn sí logra la individualización ya que cada perfil genético de STRs nucleares es único y diferente para cada individuo, con la excepción de gemelos idénticos. Sin embargo, el compartir el mismo haplotipo entre diferentes generaciones dentro del mismo linaje materno permite la filiación aún con familiares distantes dentro de este linaje. De tal forma, que la genotipificación del ADNmt en muestras comprometidas (baja concentración de ADN, degradación, etc.) ofrece varias ventajas, como: mejor conservación en herencia vertical, mayor número de copias del genoma mitocondrial por célula y sufrir menor degradación por ser un genoma circular y de menor tamaño. La integración de esta técnica forense permitiría mejorar los procedimientos de identificación humana a través de la Genética Forense en el trabajo de la FAFG.

Adicionalmente, para poder incluir el peso estadístico de coincidencias genéticas utilizando el haplotipo de ADNmt, es necesario llevar a cabo un estudio poblacional para conocer la distribución de los haplotipos presentes en la población guatemalteca. Debido a que Guatemala cuenta con varios grupos etnolingüísticos, es necesario considerarlos y establecer los haplogrupos presentes, estableciendo la frecuencia de cada uno de ellos. Debido a la alta conservación de del ADNmt se busca maximizar el tamaño de muestra ($n=17$) por grupo etnolingüístico, con la limitación de un máximo de 102 muestras. Se seleccionaron los grupos achí, ixil, kaqchikel, ladino, queqchí y quiché para el estudio poblacional debido a la alta disponibilidad de muestras (almacenadas desde el 2009 hasta el 2018) y su utilidad para fines de identificación ya que provienen de familiares de víctimas de desaparición.

IV. MARCO TEÓRICO

A. ÁCIDO DESOXIRRIBONUCLEICO

Los ácidos nucleicos, incluido el ADN, se constituye por nucleótidos los cuales constan de una base (variante), una azúcar y un fosfato (que constituyen la cadena estructural). Los nucleótidos, también conocidos como bases nitrogenadas o solamente bases, son la adenina (A), tiamina (T), citosina (C) y guanina (G) (Butler & Butler, 2010). El azúcar que contiene el ADN es desoxirribosa lo que explica el nombre de “Desoxirribonucleico” (Krebs, Goldstein, & Kilpatrick, 2014). El genoma humano está constituido por ADN contenido en el núcleo (ADNn) y en las mitocondrias. El ADN mitocondrial (ADNmt) contribuye con un 0.0006% al genoma humano total (Mark Stoneking, 1994). En el Cuadro 1 se muestran algunos datos sobre la cantidad de ADN contenido en muestra biológica.

Cuadro 1. Contenido de ADN en muestras biológicas.

Tipo de muestra	Cantidad de ADN
Sangre líquida	20,000 – 40,000 ng/mL
Mancha de sangre	250 -500 ng/cm ³
Semen líquido	150,000 – 300,000 ng/mL
Hisopado vaginal poscoital	10 – 3,000 ng/hisopo
Cabello con raíz arrancado	1 -750 ng/raíz
Cabello con raíz rasurado	1 – 10 ng/raíz
Saliva líquida	1,000 – 10,000 ng/mL
Hisopado bucal	100 – 1,500 ng/hisopo
Orina	1 – 20 ng/mL
Hueso	3 – 10 ng/mg
Tejido	50 – 500 ng/mg

Las muestras de evidencia son significativamente afectadas por factores en el ambiente (Lee & Ladd, 2001).

La contribución genética en la descendencia proviene de 22 cromosomas autosómicos, un cromosoma X y el ADNmt en el óvulo de la madre; 22 cromosomas autosómicos y un cromosoma X o Y en el espermatozoide (Figura 1) (Butler & Butler, 2010). Debido a que el ADNmt no sufre recombinación es posible trazar la descendencia materna entre generaciones bastante separadas (Budowle *et al.*, 2003). El espermatozoide posee mitocondrias en la región del cuello y la cola, pero son destruidas durante o poco después de la fertilización (Budowle *et al.*, 2003). Las mitocondrias espermáticas desaparecen en etapas tempranas de la embriogénesis (Amorim, Fernandes, & Taveira, 2019), por destrucción selectiva, inactivación o dilución. Recientemente se describió inherencia biparental de ADNmt entre 17 miembros de tres familiares con múltiples generaciones, los resultados fueron confirmados por dos laboratorios externos (Luo *et al.*, 2018).

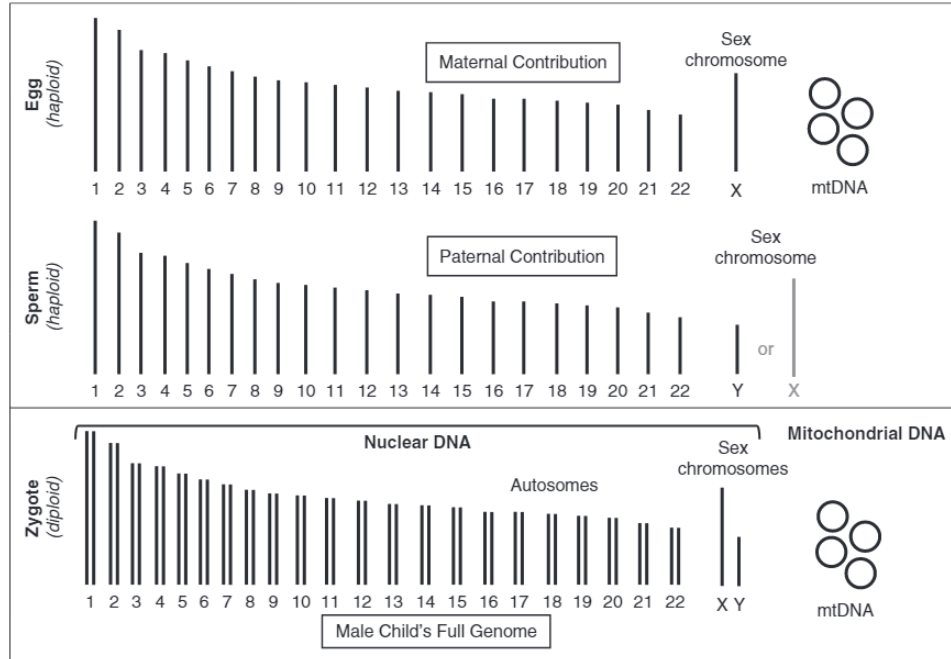


Figura 1. Material genético en los gametocitos y cigoto.

Fuente: (Butler & Butler, 2010).

El óvulo contribuye con 22 cromosomas autosómicos, un cromosoma sexual y ADN mitocondrial. El espermatozoide contribuye con 22 cromosomas autosómicos y un cromosoma sexual. LA formación del cigoto posee 22 pares de cromosomas autosómicos, dos cromosomas sexuales y ADN mitocondrial.

B. GENOTIPIFICACIÓN DE ADNmt

El análisis de ADNmt se utiliza en casos forenses cuando los marcadores nucleares no pueden ser aplicados. El uso de haplogrupos, según las regiones variables, sugieren un control adicional del código regional, lo cual incrementa el poder de discriminación del ADNmt (Bandelt *et al.*, 2012). Debido a la inherencia materna del ADNmt solamente algunas muestra de comparación son útiles según se muestra en la siguiente figura.

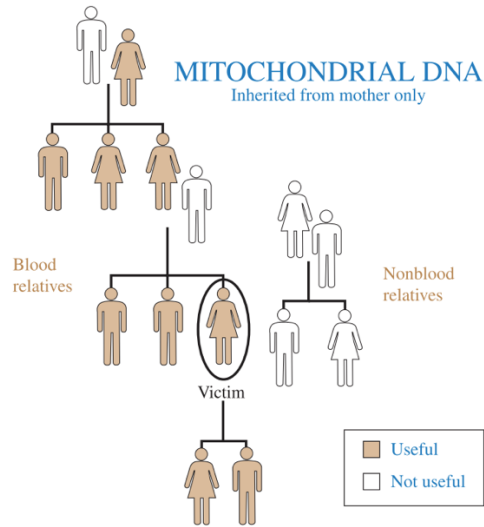


Figura 2. Muestras útiles para comparación genética de ADNmt.

Fuente: (Nij, 2005)

Las muestras útiles para la comparación genética de ADNmt se muestran en color café y está relacionado con el linaje materno heredado a toda su descendencia, pero solamente transmitido por mujeres. Lo que permite la trazabilidad del ADNmt entre la familia, incluso, entre parientes lejanos o comparaciones complejas.

El ADNmt presenta un número de copias más alto que el ADNn en células humanas (Hares *et al.*, 2014). La mitocondria contiene entre 2 a 10 copias de ADNmt, y pueden haber hasta 1000 mitocondrias por célula somática (Budowle *et al.*, 2003). Por esta razón el análisis de ADNmt es útil en el área forense cuando no se logra obtener un perfil genético de STRs autosómicos satisfactorio (Hares *et al.*, 2014).

La baja fidelidad de la polimerasa de ADNmt y la aparente ausencia de mecanismos de reparación dan lugar una tasa de mutación más alta que la del ADNn (Budowle *et al.*, 2003). Se ha reportado la estimación de la tasa de mutación en las dos regiones hipervariables en 0.32×10^{-6} / sitio / año (IC 95% $0.065 \times 10^{-6} - 0.97 \times 10^{-6}$) lo cual es mayor en comparación a 0.5×10^{-9} / sitio / año en genoma nuclear (Amorim *et al.*, 2019). El ADNmt posee regiones polimórficas conocidas como Hipervariable I, II y III (HVI, HVII y HVIII respectivamente) (Bandelt *et al.*, 2012). Las variaciones de secuencia entre individuos se encuentran especialmente en las regiones HVI (16024-16365) y HVII (73-340) (Budowle *et al.*, 2003; Hares *et al.*, 2014).

Para el reporte de haplotipos de ADNmt para casos forenses o estudios poblacionales requiere identificar solamente las diferencias con respecto a la versión revisada (rCRS, NC001807) de la primera secuencia de ADNmt humano (Hares *et al.*, 2014).

Estudios por Söchtig *et al.* (2015) evalúan la variabilidad genética de los grupos maya y ladinos de Guatemala (Figura 3) a través de marcadores uniparentales y ancestrales. En este estudio llegan a concluir que las características filogenéticas del ADNmt son compatibles con la endogamia moderada local, el aislamiento de los mayas combinado con episodios de intercambio de genes dentro los grupos étnicos.



Figura 3. Mapa de la República de Guatemala con las áreas de muestreo.

Fuente: (Söchtig *et al.*, 2015)

El mapa recalca las áreas de muestreo que abarcan Alta y Baja Verapaz, El quiché y la ciudad Capital. Las muestras provenían de los siguientes grupos etnolingüísticos: qeqchí, achí, poqomchí, quiché y kaqchikel.

C. ADN ANCESTRAL

Debido a las ventajas del ADNmt mencionadas anteriormente ante el ADNn se han logrado realizar estudios antropológicos tales como los hecho en México sobre la demografía histórica de la población indígena (Rubi-Castellanos *et al.*, 2015) o la identificación de los restos del rey Ricardo III (King *et al.*, 2014). Aunque existen algunos desafíos al momento de trabajar con ADN ancestral, debido a las modificaciones que ocurren después de la muerte como: el rompimiento de

una hebra de ADN, modificación de una base por oxidación e hidrólisis y pérdida de bases (Cipollaro, Galderisi, & Di Bernardo, 2005). La degradación del ADN está relacionada a las condiciones ambientales, sobre todo la temperatura de deposición y no tanto sobre su edad arqueológica o tiempo de deposición (Allentoft *et al.*, 2012).

También existen desafíos externos a la integridad del ADN, tales como la presencia de inhibidores que intervienen en los procesos de amplificación y secuenciación de las muestras, generando resultados pobres y de baja calidad (Alaeddini, 2012). Algunos de estos inhibidores son los ácidos húmicos y fúlvicos, residuos de parafirina o sus productos de degradación (Montiel, Malgosa, & Subirà, 1997; Pääbo, 1989).

Debido a la especificidad de los kits la contaminación por parte de bacterias, virus, animales y hongos no compiten en las reacciones para generar perfiles genéticos, pero sí compite con la de otro donador. Como último desafío se debe resaltar la contaminación de ADN humano exógeno. Esta contaminación pudo haber ocurrido durante la muerte, en los rituales de sepultura, durante la exhumación y su análisis antropológico (Anastasiou & Mitchell, 2013).

D. TOMA DE MUESTRA

El hisopado bucal es un método no invasivo de recolección de tejido que consiste en transferir células epiteliales por frotación a un hisopo. La toma de muestra con esta técnica tiene las siguientes ventajas: representa un menor riesgo biológico tanto para el donante como para quien toma la muestra, no existe problema de derrame de material biológico y se puede almacenar a temperatura ambiente (Jehuda, 1993). En la FAFG se utiliza este método de recolección de muestras y adicionalmente se fijan en tarjetas FTA.

Las tarjetas FTA (Flinders Technology Associates) constan de un papel filtro recubierto con algunos agentes químicos que permiten la fijación y preservación de muestras de ADN (Patent No. 5,496,562, 1996; Green, Tillmar, Pettersson, & Montelius, 2019). Existen métodos de extracción de ADN para las tarjetas FTA, pero han salido al mercado kits que utilizan pequeños fragmentos de la tarjeta directamente en la reacción de PCR, lo cual es una ventaja costo efectiva en comparación a otras técnicas de toma de muestra (Mundorff *et al.*, 2018). Se reportó que es posible obtener perfiles de STR completos a partir de muestras almacenadas por 16 años en tarjetas FTA, aunque detectaron que la cantidad y degradación del ADN decrece con relación al tiempo de almacenamiento (Rahikainen, Palo, de Leeuw, Budowle, & Sajantila, 2016).

E. SECUENCIACIÓN

En 1975 Frederick Sanger desarrolló un método para determinar la secuencia de nucleótidos de ADN, y el método fue nombrado Secuenciación Sanger. El procedimiento consiste en preparar cuatro reacciones en las que se incluían trifosfatos de desoxinucleicos (dNTP) normales A, G, C y T, pero cada uno contiene un nivel bajo de uno de los cuatro trifosfatos de didesoxinucleótidos (ddNTP). Al incorporarse uno de los ddNTP en la síntesis de la cadena de ADN termina. Los productos de cada reacción se separan por electroforesis en gel; cuya lectura se realiza de abajo hacia arriba según aparezcan las bandas en cada carril (Sanger, Nicklen, & Coulson, 1977).

Algunos de los eventos estructurales del ADN son las sustituciones, deleciones o inserciones de bases (Arsham & Shaffer, 2017). Estos eventos son detectados al comparar la secuencia de las muestras con una Secuencia de Referencia, que se genera a partir de un consenso (Scientific Working Group on DNA Analysis Methods, 2019). Las sustituciones consisten en los cambios de bases; las deleciones, en la pérdida de las bases; las inserciones, en la adquisición de bases inexistentes en la referencia (Arsham & Shaffer, 2017; Butler & Butler, 2010).

F. VALIDACIÓN DE MÉTODO

La FAFG se rige bajo la norma ISO 17025, la cual exige la validación de todos sus métodos. Estas validaciones implican realizar pruebas con los equipos, reactivos, operadores y muestras que permitan garantizar resultados confiables (ISO, 2017). Por lo que se utilizará la guía para la validación de métodos de análisis de ADN que contempla los siguientes estudios: La exactitud que se evalúa consta en analizar la habilidad de la técnica (reactivos e instrumentos) para proveer resultados (secuencias del ADNmt) cercanas a la secuencia conocida (previamente obtenida por un laboratorio Forense acreditado y externo al de la FAFG). El estudio de sensibilidad busca establecer los rangos dinámicos e ideales de concentraciones de ADN para la obtención de la secuencia completa del ADNmt necesarios para reportar el haplotipo de forma correcta a partir de muestras referenciales y esqueléticas. La repetitividad consta de la evaluación de la exactitud y precisión de los resultados obtenidos por un mismo operador y/o equipo, para el cual se espera obtener los mismos resultados para una muestra procesada varias veces. El estudio de reproducibilidad busca evaluar la exactitud y precisión para obtener los mismos resultados para dos operadores y/o equipos diferentes. La evaluación de mezclas como parte de la validación permite reconocer resultados de mezclas o muestras contaminadas. El programa de análisis también debe ser validado para completar el proceso, por lo que se requiere de compararlo con los resultados de otro programa

previamente validado o con resultados provistos por otro laboratorio validado (Scientific Working Group on DNA Analysis Methods, 2016).

El programa de análisis para este estudio es Sequencher® (GeneCodes), que permite ensamblar y alinear secuencias, detectar polimorfismos de un solo nucleótido (SNP por sus siglas en inglés). Entre sus funciones está la generación de haplotipos de ADNmt bajo estándares forenses, pero posee otras herramientas de análisis de secuenciación de Sanger o de secuenciación en paralelo o nueva generación (Gene Codes Corporation, 2018).

El estudio poblacional consiste en la parte final de la validación, en la que se caracteriza a un grupo de interés, que en el caso de ADNmt consiste en la identificación de los haplogrupos de la población (conjunto de eventos estructurales propios del grupo) (Budowle *et al.*, 2003).

1. Genotipificación de ADNmt

Se realizaron 5 recomendaciones por parte de la Sociedad internacional para la Genética Forense (Hares *et al.*, 2014):

- Se debe trabajar bajo buenas prácticas de laboratorio y protocolos específicos para trabajar ADNmt.
- En todo el proceso deben considerarse los controles negativos, positivos y blancos requeridos.
- Se deben reportar solamente las secuencias consenso basado en la información redundante de la secuencia al utilizar secuencias directas e indirectas.
- Debe evitarse la transcripción manual de datos y reportar haplotipo consenso por parte de dos operadores.
- Los laboratorios que tipifiquen ADNmt para casos forenses deben participar regularmente en pruebas de aptitud.

2. Lineamientos para la interpretación del análisis de ADNmt

Son necesarios el uso y evaluación de controles, el obtener los resultados esperados garantiza cumplir con estándares de calidad. Deben cubrirse la región hipervariable I (HVI, en las posiciones 16024 a 16365) y la región hipervariable II (HVII, en las posiciones 73 a 340). El llamado de bases o designación de bases deben ser según el sistema de nomenclatura IUPAC (International Union of Pure and Applied Chemistry) incluso para las ambigüedades. (Scientific Working Group on DNA Analysis Methods, 2019). La presentación de haplotipos debe cumplir con las reglas SWGDAM (existen otros dos reglamentos):

Regla 1: Mantener los patrones conocidos de polimorfismos.

Regla 2: Usar todas las diferencias requeridas, al menos que viole los patrones conocidos de polimorfismos.

Regla 3: Los estiramientos en C en HVI deben interpretarse con un 16189C cuando el T anclado en la posición 16189 no está presente. Las variaciones de longitud en el tracto A corto que precede a 16184 deben notarse como transversiones. Los estiramientos en C en HVII deben interpretarse con un 310C cuando la T anclada en la posición 310 no está presente. Los estiramientos en C deben interpretarse con un 311T cuando la T anclada en la posición 310 es seguida por una segunda T.

La regla 4 será omitida debido a que aplica para la región HVIII.

Regla 5: Se prefieren sustituciones ante inserciones y deleciones.

Regla 6: Se prefieren transiciones a transversiones.

Regla 7: Colocar inserciones y deleciones continuamente cuando sea posible.

Regla 8: Colocar inserciones y deleciones en el extremo 3' de la hebra ligera.

La regla 9 será omitida debido a que no aplica para las regiones hipervariables I y II.

V. METODOLOGÍA

A. SITIO DE ESTUDIO

No se puede delimitar un sitio específico debido a la heterogeneidad de las muestras, sin embargo, todas provienen del territorio guatemalteco.

B. SUJETOS DE ESTUDIO

Individuos en búsqueda de sus familiares desaparecidos.

Individuos de los cuales se logró recuperar la osamenta o parte de ella.

Colaboradores de la Fundación de Antropología Forense de Guatemala.

Individuos pertenecientes a las comunidades achí, ixil, kaqchikel, ladino, queqchí y quiché en Guatemala.

C. DISEÑO, ENFOQUE Y TIPO DE INVESTIGACIÓN

El diseño experimental de tipo verdadero es el apropiado para este estudio ya que se busca evaluar variables para determinar el rango de concentración de ADN en el cual es posible obtener un haplotipo, siempre acompañado de controles positivos y negativos que permitan validar y proporcionar confiabilidad en las pruebas. También consta de un diseño no experimental de tipo transversal en la que se caracterizará las diferencias entre los haplogrupos de algunas comunidades guatemaltecas; y de tipo longitudinales en la que se evaluará la estabilidad de muestras tomadas entre los años 2009 al 2018 y muestras del mismo individuo tomadas en distintos años.

Este estudio consta de un enfoque mixto en el cual la fase cualitativa abarca la validación de un método para la generación de haplotipos de ADNmt, en lo que se evaluará sensibilidad, precisión, exactitud, repetitividad, reproducibilidad, estabilidad y mezclas. En la fase cualitativa se busca caracterizar a las comunidades achí, ixil, kaqchikel, ladino, queqchí y quiché a través de la evaluación de frecuencias en cabios presentes en las diferentes posiciones de las regiones HVI y HVII.

D. TIPO Y TAMAÑO DE MUESTRA

El estudio requiere de extractos de muestras óseas, muestras referenciales que incluyen: hisopados bucales, muestras de sangre y bucales fijadas en FTA y cabellos con raíz.

El estudio poblacional requiere de 17 muestras por comunidad, para un total de 102 muestras.

Tamaño de muestra y tipo de muestreo es con fines especiales para los estudios de sensibilidad, precisión, exactitud, repetitividad, reproducibilidad, estabilidad y mezclas; y un muestreo por conveniencia para el de población.

E. CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN

Cuadro 2. Criterios de inclusión y exclusión.

Estudio	Criterios	
	Inclusión	Exclusión
Sensibilidad	1 muestra con un extracto de concentración mínima de 1.50 ng/μL. Muestras bucales o de sangre fijadas en tarjetas FTA. Muestras con haplotipo de ADNmt conocido (muestra de referencia, evaluado por un laboratorio acreditado externo a la FAFG). Extractos de muestras óseas cercanas a las siguientes concentraciones 0.298, 0.130, 0.050, 0.025, 0.013 y 0.002 ng/μL con perfil de STRs autosómicos. Un extracto de muestra ósea con cuantificación mayor a 4.0 ng/μL.	Muestras no cuantificadas (exceptuando en caso de perforación de FTA).
Estabilidad	Muestras de hisopados bucales tomadas en los años 2009, 2011, 2014, 2017 y 2018,	Muestras que se desconoce el año de toma de muestra.
Precisión	Diferentes diluciones de ADN para una misma muestra.	Muestras con cuantificación inferior a 1.50 ng/μL. Muestra de donante diferente.
Exactitud	2 muestras de haplotipo de ADNmt conocido. Haplotipo que abarca la región HVI y HVII. Muestras bucales o sangre fijadas en tarjeta FTA.	Muestras tomadas en el laboratorio. Muestras sin respaldo de haplotipos generados en otros laboratorios.
Repetibilidad ¹	Muestra bucal y de sangre fijada en tarjeta FTA y cabello con raíz del mismo donante.	Muestras evaluadas por diferentes analistas.
Reproducibilidad ²	Muestras previamente evaluadas por otro analista.	Muestras procesadas por el mismo analista.
Mezcla	Se conoce la conformación de la muestra (quienes son los donadores). Se generó o conoce los haplotipos de los donadores.	Los donadores pertenecen al mismo linaje materno.
Poblacional	Pertenecer a uno de los siguientes grupos etnolingüísticos: Achí, Ixil, Kaqchikel, Ladino, Queqchí y Quiché. Disponibilidad de extractos de ADN.	Muestras de individuos pertenecientes al mismo linaje materno (corroborado con los genogramas).

¹ Evaluación de la precisión y exactitud en la obtención de resultados por un mismo operador.

² Evaluación de la precisión y exactitud en la obtención de un resultado por varios operadores.

F. VARIABLES

Cuadro 3. Variables.

Variable	Descripción general	Naturaleza	Nivel de medición	Codificación/unidad de medición
Haplotipo	Serie de posiciones y bases que difieren entre la secuencia de ADNmt de la muestra con la rCRS ³ .	Cualitativo	Nominal	Sí (se generó un haplotipo) No (no se pudo generar un haplotipo)
Tipo de muestra	Tejido y/o sustrato que contiene el material genético de interés.	Cualitativo	Nominal	Óseo (evidencia) Hisopado bucal (referencial) Bucal fijado en tarjeta FTA (referencial) Sangre fijada en tarjeta FTA (referencial) Raíz de cabello (referencial)
Año	Año en el que la muestra fue tomada.	Cuantitativo	Discreta	Año de toma de muestra.
Población	Muestras de referencia provenientes de una comunidad guatemalteca	Cualitativo	Nominal	Achí Ixil Kaqchikel Ladino Queqchí Quiché
Concentración	Cantidad de ADN suspendido en determinado volumen.	Cuantitativo	Continua	ng/ μ L
Analista	Investigador que procesa la muestra dese PCR hasta el reporte del Haplotipo.	Cualitativo	Nominal	Analista 1 Analista 2

G. INSTRUMENTOS DE MEDICIÓN

No se utilizó ningún instrumento para la colección de datos.

³ Revised Cambridge Reference Sequence.

H. METODOLOGÍA

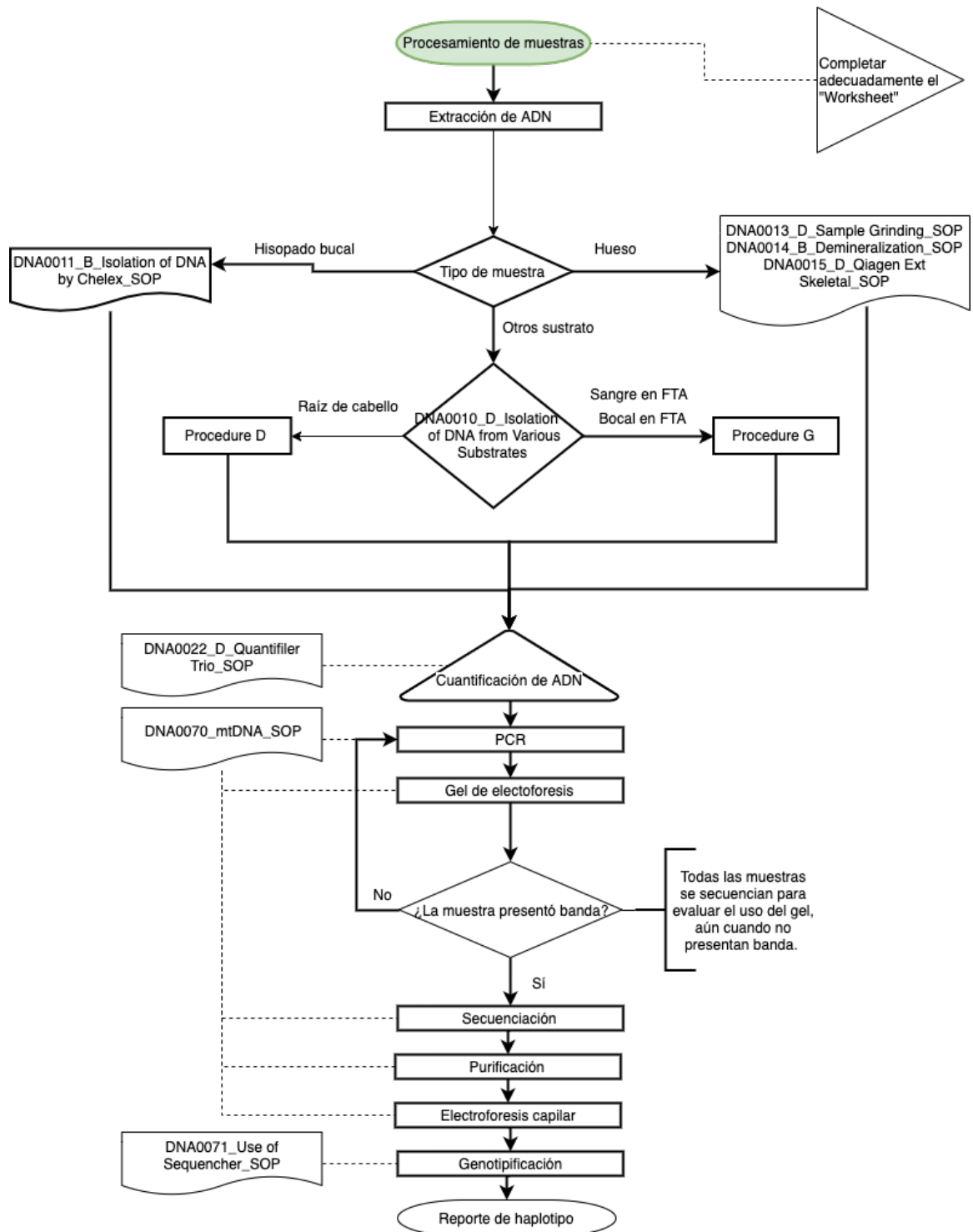


Figura 4. Procedimiento para procesamiento de muestras con ADNmt.

1. Extracción de ADN

Se utilizó la extracción de ADN por Chelex para hisopados bucales. El proceso emplea la resina Chelex 100® (Laboratorios Bio-rad, CA, EE. UU.) al 5%, altamente purificado aplicando el procedimiento operacional estándar (POE) DNA0011⁴.

Se utilizó la extracción de ADN de muestras de raíz de cabello, material manchado con sangre o saliva aplicando el procedimiento D del POE DNA0010⁵. Este procedimiento, hace uso del kit QIAamp® DNA Micro (QIAGEN, CA, EE. UU.) diseñado para purificación de ADN total para varios sustratos y muestras de volumen o tamaño reducido. Se llevó a cabo la purificación de ADN de muestras de células bucales o sangre fijadas en FTA siguiendo el procedimiento G del POE DNA0010 que utiliza “reactivo de purificación de FTA” (GE, NJ, EE. UU.) y solución amortiguadora TE.

2. Cuantificación de ADN

Se utilizó el kit Quantifiler™ Trio DNA Quantification (Applied Biosystems™CA, EE. UU.) con el sistema 7500 Fast & 7500 Real-Time PCR System (Applied Biosystems™, CA, EE. UU.) para la cuantificación de ADN total humano presente en extractos según POE DNA0022⁶.

3. Reacción en cadena de la Polimerasa

Se amplificó las regiones HV1 y HV2 para todos los sustratos de ADN utilizando el kit BigDye™ Direct Cycle Sequencing (Applied Biosystems™, CA, EE. UU.) y los cebadores F15978, R16391, F30 y R408 según el POE DNA0070 (Anexo A).

4. Gel de electroforesis

Los productos de PCR se evaluaron a través de electroforesis en gel de agarosa empleando el sistema E-Gel® iBase™ (Invitrogen™, CA, EE. UU.) que emplea luz ultravioleta para la visualización de bandas según el POE DNA0070.

5. Secuenciación

Los productos de PCR se procesaron con el kit BigDye™ Direct Cycle Sequencing (Applied Biosystems™, CA, EE. UU.) y BigDye™ XTerminator Purification (Applied Biosystems™, CA, EE. UU.), según el POE DNA0070 para la obtención de la secuencia de ADN mitocondrial de las

⁴ DNA0011_B_Isolation of DNA by Chelex_SOP, método validado en la FAFG.

⁵ DNA0010_D_Isolation of DNA from Various Substrates, método validado en la FAFG.

⁶ DNA0022_D_Quant Trio_SOP, método validado en la FAFG.

regiones HVI y HVII a través de la técnica de Secuenciación de Sanger con el Analizador Genético 3500 (Applied Biosystems™, CA, EE. UU.).

6. Determinación de haplotipos

Los resultados de la secuenciación se analizaron con el programa Sequencher® (GeneCodes, MI, EE.UU.) según el POE DNA0071 (Anexo B).

7. Estudio poblacional

El número de haplotipos fue calculado por el método de conteo directo. La capacidad de discriminación fue estimada al dividir el número de haplotipos diferentes por el total de individuos (Rakha *et al.*, 2017). Se estimó la diversidad de Haplotipos (DH), $DH = n(1 - \sum f^2)/(n - 1)$, donde f es la frecuencia de la muestra de cada haplotipo y n es el número de individuos (Chen *et al.*, 2015). La probabilidad de coincidencia aleatoria (PCA) es $PCA = \sum f^2$, donde f representa la frecuencia de cada haplotipo de ADNmt (M. Stoneking, Hedgecock, Higuchi, Vigilant, & Erlich, 1991). Se estimó el rendimiento como $(R) = N/H$, donde N corresponde al número total de individuos evaluados y H la cantidad de individuos que se logró genotipificar. Los resultados de los cálculos se multiplicaron por 100 para poder reportar los resultados como porcentajes.

VI. RESULTADOS

A. EXACTITUD

Hubo un 100% de coincidencia entre el haplotipo esperado (Cuadro 4) y obtenido utilizando la técnica propuesta en este proyecto de investigación (Cuadro 5) para las muestras M2 y M3.

Cuadro 4. Haplotipos generados en laboratorio externo para la muestra M2 y M3.

ÍTEM	HVI			HVII								
	N	n	%	(16024-16365)		N	n	%	(73-340)			
M1	53	52	98	T16224C	T16311C	50	48	96	A73G	C114T	A263G	-315.1C
M2	53	53	100	Sin diferencias		51	50	98	A73G	C150T	A263G	-315.1
M3	53	51	96	G16213A	A16293G	52	50	96	A263G -315.1C			

N = número de laboratorios que han realizado el análisis. n = número de laboratorios con consenso. Los haplotipos (gonotipos de ADNmt) en genética forense se reportan con la nomenclatura UPAC y los lineamientos de SWGDAM. Los haplotipos muestran las diferencias entre la referencia y la muestra, la posición se indica con números, seguidos por la base (A, T, C o G) en caso de que la diferencia sea una sustitución. La posición seguido de “-“ para las deleciones. La posición seguido de “. #” (el # es un número correlativo de inserciones) acompañados por la base insertada.

Cuadro 5. Haplotipo generado en la FAFG para la muestra M2 y M3.

Muestra	HVI (16024-16365)	HVII (73-340)
M2	Sin diferencias	73G 150T 263G 315.1C
M3	16,213A 16,293G	263G 315.1C

No se muestran resultados para la muestra M1 debido a que el extracto de ADN se perdió por evaporación.

B. PRECISIÓN

Se obtuvo el mismo haplotipo ($f = 100\%$, $N = 16$) para la muestra M2 en diferentes procesamientos al trabajar con concentraciones de 1.780, 1.000, 0.500, 0.250, 0.125, 0.060, 0.030, 0.015 ng/ μ L (en duplicado). Estos resultados también se utilizaron para evaluar la sensibilidad del método (rango ideal).

C. SENSIBILIDAD

Se evaluó la sensibilidad para muestras extraídas por el método Chelex (hisopado bucal) y QIAamp (hisopado bucal fijado en tarjeta FTA, sangre fijada en tarjeta FTA y raíz de cabello), los métodos de rutina para muestras referenciales en las instalaciones de la FAFG. En los Cuadros 6 y 7 se muestran los haplotipos obtenidos para las muestras evaluadas.

Cuadro 6. Haplotipos para muestras de hisopado bucal tomadas en diferentes años (extracción por Chelex).

Muestra	Toma de muestra (año)	HV1 (16024-16365)	HV2 (73-340)
1	2009	Sin amplificación*	Sin amplificación
2	2009	16,093C 16,129A 16,223T 16,298C 16,325C 16,327T	73G 214G 249- 263G 290- 291- 309.1C 315.1C
3	2011	16,111T 16,223T 16,290T 16,319A	73G 146C 153G 235G 263G 309.1C 315.1C
4	2011	Sin amplificación	Sin amplificación
5	2014	16,111T 16,223T 16,290T 16,299G 16,319A 16,362C	73G 146C 153G 235G 263G 315.1C
6	2014	16,111T 16,223T 16,290T 16,319A 16,362C	73G 146C 152C 153G 200G 235G 263G 309.1C 315.1C
7	2017	16,111T 16,223T 16,290T 16,319A 16,362C	73G 146C 152C 153G 214G 235G 263G 309.1C 315.1C
8	2017	16,111T 16,290T 16,319A 16,362C	73G 146C 153G 235G 263G 309.1C 315.1C
9	2012 2018	16,223T 16,290T 16,319A 16,362C	73G 146C 153G 182T 184G 235G 263G 309.1C 315.1C

La muestra 9 corresponde a un mismo donante, se evaluó precisión y exactitud al obtener el haplotipo en diferentes años de toma de muestra. *Sin amplificación significa que no se logró obtener banda al evaluar los productos de PCR en Gel.

Cuadro 7. Haplotipo para muestras de diferente sustrato de un mismo donante (extracción por QIAamp).

Muestra	N	n	HV1 (16024-16365)	HV2 (73-340)	Frecuencia (%)
A	3	2*	16,233T 16,290T 16,319A 16,362C	73G 146C 153G 182T 184A 235G 263G 309.1C 315.1C	100
B	3	3	16,104T 16,111T 16,223T 16,290T 16,311C 16,319A 16,362C	73G 146C 153G 235G 263G 309.1C 309.2C 315.1C	100
C	3	3	16,223T 16,290T 16,319A 16,362C	73G 146C 153G 235G 263G 315.1C	100

*No se logró obtener haplotipo para la muestra de sangre, debido a que la cuantificación falló.

Se usaron muestras óseas que presentaran una concentración cercana a la deseada a evaluar (entre 0.002 a 0.300 ng/μL) y con previa evaluación de STR (datos de la FAFG) para detectar

indicios de correlación de esos dos factores en la obtención de haplotipos reportables, los resultados se muestran en el Cuadro 8.

Cuadro 8. Haplotipos para extractos óseos con diferente cuantificación y perfil genético reportado.

Muestra	Perfil STR	Concentración (ng/μL)	HVI (16024-16365)	HVII (73-340)
1h	18 loci	0.298	No se logró amplificar*	73G 263G 309.1C 309.2C 315.1C
2h	Completo	0.130	No se logró amplificar	73G 263G 309.1C 309.2C 315.1C
3h	Completo	0.050	No se logró amplificar	73G 146C 153G 235G 263G 309.1C 315.1C
4h	Completo	0.025	16,111T 16,223T 16,290T 16,319A 16,360T 16,362C	No se logró amplificar
5h	Falló	0.013	No se logró amplificar	No se logró amplificar
6h	Falló	0.002	No se logró amplificar	73G 204C 263G 315.1C

*"No se logró amplificar" significa que los fragmentos no cubrieron por completo la región, no se obtuvo secuencia o faltó cubrir la región con alguno de los dos fragmento (sentido o antisentido).

Se evaluaron diferentes diluciones para una muestra esquelética con una alta concentración de ADN. Las concentraciones fueron 4.000, 1.000, 0.500, 0.100, 0.050, 0.025 y 0.005 ng/μL. No se logró amplificar con concentraciones de 0.100, 0.050, 0.025 ni 0.005 ng/μL como se evidencia en la Figura 5. La frecuencia del haplotipo es del 29 % (N = 7, Cuadro 11, Anexo C) para las concentraciones de 1.000 y 0.500 ng/μL.

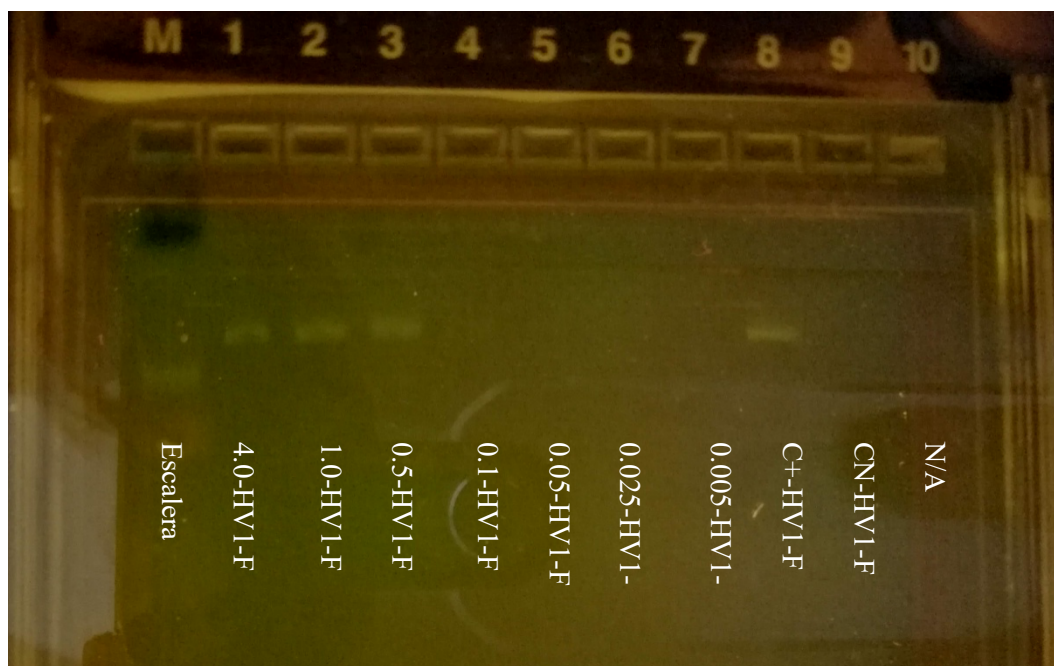


Figura 5. Gel de agarosa 1% correspondiente a la región HVI para diferentes diluciones de una muestra ósea.

El primer carril corresponde a la escalera, los carriles numerados con 1, 2, 3, 4, 5, 6 y 7 corresponden a la muestra evaluada a la concentración de 4.000, 1.000, 0.500, 0.100, 0.050, 0.025 y 0.005 ng/ μ L, respectivamente. Los carriles 8 y 9 corresponden al control positivo y negativo, respectivamente. Todos los productos de PCR son de fragmento líder de la región HVI.

Adicionalmente se evaluó la sensibilidad para diferentes diámetros de perforaciones en FTA con hisopado bucal. Las perforaciones de 0.5 mm de diámetro con lavados produjeron haplotipos reportables. Los diámetros de 1.2 y 2.0 mm con lavado solamente produjeron haplotipos reportables para el 50 % de los procesamientos (se trabajó en duplicado). Las perforaciones sin lavados no proporcionaron haplotipo reportable para ninguno de los tres diámetros antes descritos.

D. ESTABILIDAD

Se evaluó la estabilidad de muestras referenciales tomadas en diferentes años (dos muestras por año), y determinar si el periodo de almacenaje es un factor por considerar para la genotipificación. Se obtuvo haplotipos reportables (Cuadro 6) para todos los años de toma de muestra (2009, 2011, 2012, 2014, 2017, 2018), exceptuando una de las muestras tomadas en el 2009 y otra del 2011.

E. REPETIBILIDAD

La repetibilidad evaluó extractos de sangre fijada en tarjeta FTA, bucales fijados en tarjeta FTA y cabello con raíz de un mismo donante. Este mismo muestreo se hizo para 3 donantes diferentes

(A, B y C), para los cuales se logró obtener haplotipos reportables (Cuadro 7) con frecuencia del 100%. La cuantificación falló para la muestra de sangre del donante A y tampoco se logró obtener el haplotipo.

F. REPRODUCIBILIDAD

En la reproducibilidad se evaluó si cualquier analista entrenado podría generar haplotipos reportables, por lo que el Analista 2 reprocesó el control positivo (CEPH) y dos muestras previamente procesadas por el Analista 1. Los resultados entre analistas tuvieron una frecuencia de 100% de coincidencia (Cuadro 11, Anexo C).

G. CONTAMINACIÓN

Se preparó una mezcla de sangre de los donantes A y B. Se observó que en las posiciones en las que diferían entre los donantes (16104, 16111, 16311, 182, 309.1, 309.2 y 315.1) no se lograron resolver las ambigüedades identificadas por Sequencher (Cuadro 9).

Cuadro 9. Haplotipo para mezcla de muestras.

Referencia	Muestra	Muestra	Mezcla de a y	
	a	b	b	
16,10 4	C	C	T	N*
16,11 1	C	C	T	N
16,22 3	C	T	T	T
16,29 0	C	T	T	T
16,31 1	T	T	C	N
16,31 9	G	A	A	A
16,36 2	T	C	C	C
73	A	G	G	G
146	T	C	C	C
153	A	G	G	G
182	C	T	C	N
184	G	A	G	N
235	A	G	G	G
263	A	G	G	G
309.1	:	C	C	C
309.2	:	:	C	:
315.1	:	C	C	C

*Sequencher utiliza “N” según la nomenclatura UPAC para señalar ambigüedades. Estas posiciones no lograron resolverse por el analista.

En la Figura 6 se muestra el pico característico para mezcla de muestras que podría ser útil para la identificación de contaminación. Estos picos están caracterizados por estar sobrepuesto y que alcanzan una altura similar.

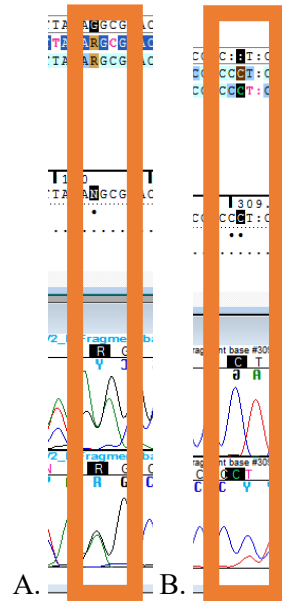


Figura 6. Picos característicos de un electroferograma en caso de mezclas.

A. Lecturas incongruentes debido a que las muestras difieren por cambio de base. En el electroferograma inferior se observa picos sobrepuestos y la disminución de ambas señales. B. Lecturas incongruentes debido que las muestras difieren por inserciones.

H. ESTUDIO POBLACIONAL

La menor capacidad de discriminación poblacional es corresponde al grupo etnolingüístico achí con 82.35 %, seguido por queqchí con 87.50 %, quiché con 92.86 %, mestizo, ixil y kaqchikel con 100 % (Cuadro 10). La probabilidad de que dos muestras aleatorias de la población posean genotipos de ADNmt idéntico corresponden a la Probabilidad de coincidencia aleatoria; considerando todos estos grupos la Probabilidad de coincidencia aleatoria corresponde a un 1.3 %. Con una diversidad de haplotipos total del 99.8%; siendo el más bajo con 97.8 % para el grupo achí y el más alto con 100 % para mestizos, ixiles y kaqchikeles. Se obtuvo un rendimiento (R) del 90 % para muestras referenciales.

Cuadro 10. Parámetros forenses e índices de diversidad de secuencias de las regiones HVI y II de ADNmt para grupos etnolingüísticos guatemaltecos.

Población	Mestizo	Ixil	Achí	Kaqchikel	Queqchí	Quiché	General
No. de muestras	17	17	17	15	16	14	97
No. de haplotipos	17	17	14	15	14	13	88
No. de sitios polimórficos	73	52	41	47	49	32	128
CD (%)	100	100	82.35	100	87.50	92.86	90.72
PCA (%)	5.9	5.9	8.0	6.7	7.8	8.2	1.3
DH (%)	100	100	97.8	100	98.3	98.9	99.8

CD es la Capacidad de discriminación; PCA, probabilidad de coincidencia aleatoria; DH, Diversidad de haplotipos.

VII. ANÁLISIS DE RESULTADOS

La coincidencia en un 100% del haplotipo esperado y el obtenido para las muestras M2 y M3 reflejan que el protocolo propuesto y las condiciones del laboratorio son apropiados para obtener secuencias exactas. De igual manera la coincidencia de los haplotipos generados para diferentes procesamientos de una muestra evidencia la precisión del método.

Los resultados del estudio de sensibilidad evidenciaron que el rango ideal para las concentraciones de ADN es entre 0.015 a 1.780 ng/ μ L, el cual está por debajo del 4.0 ng/ μ L, recomendado por los fabricantes del kit BigDye™ Direct Cycle Sequencing (Applied Biosystems, 2011). Los resultados de los Cuadros 6 y 7 evidenciaron que los métodos de extracción Chelex y QIAamp no afecta la sensibilidad del método ya que se lograron genotipificar las muestras de ambos métodos de extracción.

Es posible que el valor recomendado no está dentro del rango ideal debido a que la cuantificación es específica para ADNn humano y no para ADNmt (Thermo Fisher Scientific, 2018). Debido a que el ADN en exceso puede actuar como inhibidor de la PCR (Alaeddini, 2012), la reacción requiere bajas concentraciones de ADNn.

Los resultados de sensibilidad para las tarjetas FTA evidenciaron que el uso de las perforaciones de 0.5, 1.2 y 2.0 mm de diámetro directo en la reacción de PCR no permite la obtención de haplotipos reportables. Aun qué las perforaciones de 1.2 y 2.0 mm con tratamientos de lavado permitió obtener resultados reportables, el rendimiento fue bajo (50%). Debido a que se desconoce la cantidad de ADN fijado en las perforaciones no se puede atribuir con certeza como fuente de error la inhibición por exceso de ADN. La viabilidad de las perforaciones con lavados puede contribuirse a que este proceso extra elimina inhibidores de la PCR, sin embargo, el tratamiento no es suficiente como lo evidencia el rendimiento.

Los extractos de ADN de muestras óseas revelaron que el rango ideal es de 0.5 a 1.0 ng/ μ L. El rango está por debajo del valor recomendado por el fabricante del kit. Este rango bajo es útil debido a que las muestras óseas tienden a obtener extractos de ADN con baja cantidad de ADN. Es posible que el rango se vea disminuido en comparación al de las muestras referenciales debido a la baja cantidad y la degradación del ADNn, así como de los inhibidores de PCR presentes. El ácidos húmicos y fúlvicos, residuos de parafirina, otros productos de degradación y principalmente el calcio son posibles inhibidores (Alaeddini, 2012; Montiel *et al.*, 1997; Pääbo, 1989). Los resultados de las muestras del Cuadro 6 evidencian que los perfiles de STR completos no garantizan una a genotipificación de ADNmt reportable, cuando las muestras presentan concentraciones inferiores

al rango ideal para muestras óseas. Esto respalda que los resultados para muestras óseas son variables debido a su naturaleza, las diferentes condiciones de deposición y degradación. La congruencia entre los resultados en los geles de agarosa y los resultados de secuenciación, evidencian que este paso es crítico para decidir si se continúa con el procesamiento o volver a reprocesar. En este paso en el caso de obtener una banda se debe continuar con el procedimiento; en el caso de ausencia de banda, se debe de regresar al proceso de extracción. Esta decisión evitará la pérdida de recursos y tiempo con extractos y productos de PCR que seguramente no se lograrían secuenciar.

Los resultados de los hisopados bucales tomados desde el 2009 al 2018, mostraron resultados satisfactorios. Las muestras previas al 2009 pueden procesarse, pero considerando la posibilidad de que no se pueda genotipificar con éxito o que requieran más de un procesamiento. Uno de los donantes poseía una muestra tomada en el 2012 y otra en el 2018 (Cuadro 9, Anexo C), esto permitió garantizar que la exactitud del método no está afectada por el tiempo ni método de almacenaje. En este precedente, es posible que los haplotipos fallidos se relacionen con la degradación del material genético o la selección de la porción del hisopo tomado en esta ocasión. No fue posible determinar extender el estudio de estabilidad más años atrás por falta de disponibilidad de muestras tomadas años previos al 2009.

Los resultados positivos del estudio de repetibilidad y reproducibilidad evidencian que es posible obtener el mismo haplotipo para la misma muestra por el personal apropiadamente capacitado, utilizando los equipos, siguiendo los protocolos de operación estándar y bajo las condiciones del laboratorio. Lo cual puede contribuirse en parte a que el laboratorio está acreditado bajo la norma ISO 17025 que estandariza y mantiene controlados los equipos y condiciones del laboratorio. Esta evidencia es suficiente para aceptar que un procesamiento sea completado por varios analistas.

El estudio de contaminación proveyó de las características necesarias para identificar la presencia de más de un haplotipo en una muestra, con el fin de conocer y poder identificar eventos de contaminación en una muestra. La falta de resolución de ambigüedades en el llamado de bases por parte de los analistas son indicios de contaminación. En los electroferogramas cuando para una posición se presentan resultados incongruentes o se presentan picos traslapados y de altura media se deben considerar como indicios de contaminación.

El estudio poblacional evidenció que del total de muestras (N=102) solamente se logró genotipificar a 97, con un total de 88 haplotipos diferentes. A pesar de las genotipificaciones fallidas, se cumple con el mínimo de 50 muestras para la validación según los lineamientos de

SWGDAM (2016). La capacidad de discriminación (CD) general es del 90.72 % debido a que 6 individuos comparten haplotipos, pero para los grupos la CD mínima es del 82.35 % y máxima del 100%. Los haplotipos compartidos entre los grupos concuerda con la cercanía geográfica que poseen entre sí (Pepkin, 2019). La probabilidad de coincidencia aleatoria más alta corresponde al grupo quiché con 8.2 % y mínima para la general del 1.3 %. Se evidenció que el grupo achí es el menos diverso (97.8%) en comparación a los otros, lo cual se respalda con un haplotipo compartido por tres individuos.

El procesar 102 muestras para el estudio poblacional permite conocer la diversidad genética para el ADNmt de los guatemaltecos. Los datos son útiles para la formación de base de datos poblacional y así determinar la frecuencia de los haplotipos, lo cual es un indicador de único o común que puede ser la información genética de una muestra. Los haplotipos únicos son preferentes ya que tienen menor probabilidad de ser atribuidos a un individuo al azar. La combinación de resultados estadísticos de los análisis de STR autosómicos junto con los de ADNmt aportarían a alcanzar el umbral (según las políticas de la FAFG) de identificación humana.

La probabilidad de coincidencia aleatoria es inversamente proporcional a la capacidad de discriminación. Un valor bajo para la probabilidad de coincidencia aleatoria es un indicador de lo único o poco frecuente que es un haplotipo dentro de un grupo lo cual refuerza la identificación de un individuo (la coincidencia con un individuo al azar con la misma información genética es mínima). Las poblaciones con menor probabilidad de coincidencia aleatoria fueron la mestiza y la ixil, según el parámetro de capacidad de discriminación. Lo cual significa que los grupos restantes presentan mayor probabilidad de reportar una coincidencia por eventos de azar.

Se identificaron coincidencias de haplotipos, lo cual puede indicar que el grupo etnolingüístico posee baja diversidad genética. Los factores geográficos y lingüísticos pueden limitar a los individuos a relacionarse entre sí. Lo anterior podría explicar la redundancia entre haplotipos en un mismo grupo a pesar de que se seleccionaron individuos que aparentemente no están relacionados (lo cual fue verificado con genogramas indicados por los donadores). Los genogramas presentan las siguientes limitantes: pueden contener información solamente del núcleo familiar, por lo que no se podría detectar el linaje materno por parte de abuelas u otras descendientes; debido a que es información provista por los donantes no se puede asegurar que sea correcta.

Las muestras de la validación y las del estudio poblacional no presentaron un patrón en cuanto a los fallos de genotipificación. Debido a que los controles obtuvieron los resultados esperados se considera como paso crítico la amplificación de PCR. Es posible que existan problemas de hibridación entre la muestra y los cebadores, causado por polimorfismos de un solo nucleótido. En

contraste con el diseño de este estudio, cuando se realiza la amplificación de la región control de ADNmt se utilizan más de cuatro cebadores, tratando que los fragmentos en un mismo sentido se traslapen y garantizar una cobertura completa en ambos sentidos (Eduardoff *et al.*, 2017).

VIII. CONCLUSIONES

El método utilizando los kits BigDye™ Direct Cycle Sequencing y BigDye™ XTerminator Purification es exacto y preciso, respaldado por la comparación con laboratorios externos para las muestras M2 y M3.

El rango ideal de concentración de ADNn es de 0.015 a 1.780 ng/μL y 0.500 a 1.000 ng/μL para muestras referenciales y óseas, respectivamente, para la genotipificación de ADNmt utilizando los kits BigDye™ Direct Cycle Sequencing y BigDye™ XTerminator Purification. Las muestras óseas, incluso dentro del rango ideal, puede no obtener haplotipos reportables debido a eventos de degradación.

La estabilidad del método está limitada para hisopados bucales tomados a partir del año 2014. Las muestras tomadas previas al 2014 pueden requerir de más de un procesamiento para garantizar un haplotipo reportable, debido a que las muestras pierden su estabilidad ante tantos años de almacenamiento.

La repetitividad de la técnica está garantizada al presentar precisión y exactitud al procesar diferentes muestras (hisopado bucal, sangre fijada en tarjeta FTA y raíz de cabello) de un mismo donador y obtener el mismo haplotipo de ADNmt en todos los procesamientos.

La reproducibilidad de la técnica da lugar a que cualquier analista con la capacitación y los protocolos DNA0070 y DNA0071 logrará genotipificar el ADNmt con exactitud y precisión. El procesamiento puede ser completado por más de un analista.

Las muestras contaminadas (mezcla) se caracterizan por ambigüedades no resueltas debido a la falta de resolución del llamado de bases por electroferogramas inconsistentes. Las mezclas pueden respaldarse al compararse contra un banco de haplotipos.

El método de genotipificación de ADNmt con los kits BigDye™ Direct Cycle Sequencing y BigDye™ XTerminator Purification fue exitosamente validado según las recomendaciones de SWGDAM. Está limitado a muestras de hisopados bucales, a extractos de tarjetas FTA fijadas con saliva o sangre, cabello con raíz, extractos de ADN piezas dentales y partes óseas.

El estudio poblacional que incluye a los grupos etnolingüísticos mestizos, ixiles, achíes, kaqchikeles, queqchíes y quichés, representa la primera publicación de haplotipos ADNmt de estas poblaciones realizada en Guatemala.

IX. RECOMENDACIONES

1. Se recomienda utilizar el kit Absolute Human Mitochondrial DNA Copy Number Quantification qPCR Assay (ScienCell Research Laboratories, 2019) para la cuantificación de ADNmt de las extracciones. Los resultados permitirán obtener el rango de sensibilidad verdadero del método, en contraste al establecido en este estudio corresponde a ADNn. En cuanto a las muestras óseas y de cabello de raíz, los resultados de una cuantificación de ADNmt serían un parámetro apropiada para decidir si se requiere concentrar las muestras previo a su procesamiento para la genotipificación de ADNmt.
2. El uso de tarjetas FTA con y sin lavados no se recomienda para la genotipificación de ADNmt. En el caso de que las muestras estén fijadas en FTA se recomienda realizar una extracción de ADN de la tarjeta. Otra opción es tratar las perforaciones de 1.2 o 2.0 mm con lavados y procesar las muestras en duplicado.
3. Al trabajar con muestras óseas se recomienda concentrar los extractos si se encuentran por debajo del rango ideal (0.500 a 1.000 ng/ μ L), además de verificar siempre en gel de agarosa la amplificación de los fragmentos.
4. Se recomienda validar otros cebadores que amplifique 200 o menos pares de bases en diferentes posiciones de los fragmentos, en caso de que no se logre secuenciar alguno de los fragmentos para muestras referenciales o para muestras óseas degradadas.
5. Los hisopados bucales del 2011 y anteriores se recomienda procesar en duplicado y verificar los productos de PCR en gel de agarosa para obtener un parámetro de decisión sobre proseguir o no con el procesamiento.
6. Se recomienda genotipificar el ADNmt de los trabajadores del laboratorio y del personal que se considere necesario. Adicionalmente se recomienda comparar todas las muestras que presenten indicios de contaminación contra todos los haplotipos reportados para descartar o identificar la fuente de contaminación. Se requiere de un estudio para validar muestras con heteroplasmía, utilizando muestras provistas por otros laboratorios con el método de genotipificación de ADNmt validados.
7. Se recomienda aumentar el número de individuos considerados para el estudio poblacional, ya que otros estudios muestran un mínimo de 16 individuos por grupo (Aziz, Nawaz, Afridi, & Khan, 2019), aunque se han realizado estudios con 99 individuos por grupo poblacional

(Rakha *et al.*, 2017). Una alternativa es aumentar el número de grupos y evaluar la mejor alternativa para trabajar la información poblacional, como grupos etnolingüísticas o como un solo grupo.

X. BIBLIOGRAFÍA

- Alaeddini, R. (2012). Forensic implications of PCR inhibition—A review. *Forensic Science International: Genetics*, 6(3), 297–305. <https://doi.org/10.1016/j.fsigen.2011.08.006>
- Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., ... Bunce, M. (2012). The half-life of DNA in bone: Measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences*, 279(1748), 4724–4733. <https://doi.org/10.1098/rspb.2012.1745>
- Amorim, A., Fernandes, T., & Taveira, N. (2019). Mitochondrial DNA in human identification: a review. *PeerJ*, 7, e7314. <https://doi.org/10.7717/peerj.7314>
- Anastasiou, E., & Mitchell, P. D. (2013). Evolutionary anthropology and genes: Investigating the genetics of human evolution from excavated skeletal remains. *Gene*, 528(1), 27–32. <https://doi.org/10.1016/J.GENE.2013.06.011>
- Applied Biosystems. (2011). *BigDye® Direct Cycle Sequencing Kit PROTOCOL*.
- Arsham, M. S., & Shaffer, L. G. (2017). ISCN: the universal language of cytogenetics. In *The AGT Cytogenetics Laboratory Manual* (pp. 359–428). <https://doi.org/10.1002/9781119061199.ch8>
- Aziz, S., Nawaz, M., Afridi, S. G., & Khan, A. (2019). Genetic structure of Kho population from north-western Pakistan based on mtDNA control region sequences. *Genetica*, 0(0), 0. <https://doi.org/10.1007/s10709-019-00060-8>
- Bandelt, H.-J., van Oven, M., & Salas, A. (2012). Haplogrouping mitochondrial DNA sequences in Legal Medicine/Forensic Genetics. *International Journal of Legal Medicine*, 126(6), 901–916. <https://doi.org/10.1007/s00414-012-0762-y>
- Budowle, B., Allard, M. W., Wilson, M. R., & Chakraborty, R. (2003). FORENSICS AND MITOCHONDRIAL DNA: Applications, Debates, and Foundations. *Annual Review of Genomics and Human Genetics*, 4(1), 119–141. <https://doi.org/10.1146/annur>

ev.genom.4.070802.110352

Burgoyne, L. A. (1996). *Patent No. 5,496,562*. United States.

Butler, J. M. (John M., & Butler, J. M. (John M. (2010). *Fundamentals of forensic DNA typing*. Academic Press/Elsevier.

Chen, F., Yin, C. Y., Qian, X. Q., Fan, H. T., Deng, Y. J., Zhang, Y. D., ... Xu, P. (2015). Single nucleotide polymorphisms of mitochondrial DNA HVSI and HVS-II in Chinese Bai ethnic group. *Electrophoresis*, 36(6), 930–936. <https://doi.org/10.1002/elps.201400493>

Cipollaro, M., Galderisi, U., & Di Bernardo, G. Ancient DNA as a multidisciplinary experience. , 202 *Journal of Cellular Physiology* § (2005).

Eduardoff, M., Xavier, C., Strobl, C., Casas-Vargas, A., Parson, W., Eduardoff, M., ... Parson, W. (2017). Optimized mtDNA Control Region Primer Extension Capture Analysis for Forensically Relevant Samples and Highly Compromised mtDNA of Different Age and Origin. *Genes*, 8(10), 237. <https://doi.org/10.3390/genes8100237>

Gene Codes Corporation. (2018). *Sequencher 5.4.6* (pp. 1–8). pp. 1–8.

Green, H., Tillmar, A., Pettersson, G., & Montelius, K. (2019). The use of FTA cards to acquire DNA profiles from postmortem cases. *International Journal of Legal Medicine*. <https://doi.org/10.1007/s00414-019-02015-2>

Hares, D. R., Prinz, M., Parson, W., Parsons, T. J., Schneider, P. M., Salas, A., ... Pokorak, E. (2014). DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing. *Forensic Science International: Genetics*, 13, 134–142. <https://doi.org/10.1016/j.fsigen.2014.07.010>

ISO. (2017). ISO/IEC 17025. Retrieved September 16, 2019, from <https://www.isotools.org/2017/12/12/iso-iec-17025-nueva-version-2017/>

Jehuda, Y. (1993). Advances in Forensic Haemogenetics. *Most*, 5(3), 94720–94720. <https://doi.org/10.1007/978-3-642-78782-9>

King, T. E., Fortes, G. G., Balaesque, P., Thomas, M. G., Balding, D., Delsler, P. M., ... Schürer, K. (2014). Identification of the remains of King Richard III. *Nature Communications*, 5(1), 5631. <https://doi.org/10.1038/ncomms6631>

Krebs, J. E., Goldstein, E. S., & Kilpatrick, stephen T. (2014). *Lewin's GENES XI*. <https://doi.org/10.1111/j.1365-3148.2005.00580.x>

- Lee, H., & Ladd, C. (2001). Preservation and collection of biological evidence. *Croatian Medical Journal*, 42(3), 225–228.
- Luo, S., Valencia, C. A., Zhang, J., Lee, N. C., Slone, J., Gui, B., ... Huang, T. (2018). Biparental inheritance of mitochondrial DNA in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 115(51), 13039–13044. <https://doi.org/10.1073/pnas.1810946115>
- Montiel, R., Malgosa, A., & Subirà, M. (1997). Overcoming PCR inhibitors in ancient DNA extracts from teeth. In *Ancient Biomolecules*.
- Mundorff, A. Z., Amory, S., Huel, R., Bilić, A., Scott, A. L., & Parsons, T. J. (2018). An economical and efficient method for postmortem DNA sampling in mass fatalities. *Forensic Science International: Genetics*, 36, 167–175. <https://doi.org/10.1016/j.fsigen.2018.07.009>
- Nij. (2005). *Identifying Victims Using DNA: A Guide for Families DNA INITIATIVE*. Retrieved from <http://www.genome.gov>.
- Pääbo, S. (1989). Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the United States of America*, 86(6), 1939–1943. <https://doi.org/10.1073/pnas.86.6.1939>
- Pepkin, A. (2019). Revitalización de las Lenguas Indígenas de Guatemala (University of Mary Washington). Retrieved from https://scholar.umw.edu/student_research/269
- Rahikainen, A. L., Palo, J. U., de Leeuw, W., Budowle, B., & Sajantila, A. (2016). DNA quality and quantity from up to 16 years old post-mortem blood stored on FTA cards. *Forensic Science International*, 261, 148–153. <https://doi.org/10.1016/j.forsciint.2016.02.014>
- Rakha, A., Fatima, Peng, M.-S., Adan, A., Bi, R., Yasmin, M., & Yao, Y.-G. (2017). mtDNA sequence diversity of Hazara ethnic group from Pakistan. *Forensic Science International: Genetics*, 30, e1–e5. <https://doi.org/10.1016/J.FSIGEN.2017.07.004>
- Rubi-Castellanos, R., Tirado, S., Coble, M. D., González-Martín, A., Regalado-Liu, L., Rangel-Villalobos, H., ... Sandoval, K. (2015). Demographic History of Indigenous Populations in Mesoamerica Based on mtDNA Sequence Data. *PLOS ONE*, 10(8), e0131791. <https://doi.org/10.1371/journal.pone.0131791>
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), 5463–5467. <https://doi.org/10.1073/pnas.74.12.5463>

- ScienCell Research Laboratories. (2019). *Absolute Human Mitochondrial DNA Copy Number Quantification qPCR Assay Kit (AHMQ)*.
- Scientific Working Group on DNA Analysis Methods. (2016). *Scientific Working Group on DNA Analysis Methods Validation Guidelines for DNA Analysis Methods SWGDAM Validation Guidelines for DNA Analysis Methods*. (December 2016), 1–15. Retrieved from https://docs.wixstatic.com/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf
- Scientific Working Group on DNA Analysis Methods. (2019). *Interpretation Guidelines for Mitochondrial DNA Analysis by Forensic DNA Testing Laboratories*. Retrieved from https://docs.wixstatic.com/ugd/4344b0_f61de6abf3b94c52b28139bff600ae98.pdf
- Söchtig, J., Álvarez-Iglesias, V., Mosquera-Miguel, A., Gelabert-Besada, M., Gómez-Carballa, A., & Salas, A. (2015). Genomic insights on the ethno-history of the Maya and the “Ladinos” from Guatemala. *BMC Genomics*, *16*(1). <https://doi.org/10.1186/s12864-015-1339-1>
- Stoneking, M., Hedgecock, D., Higuchi, R. G., Vigilant, L., & Erlich, H. A. (1991). Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *American Journal of Human Genetics*, *48*(2), 370–382.
- Stoneking, Mark. (1994). Mitochondrial DNA and human evolution. *Journal of Bioenergetics and Biomembranes*, *26*(3), 251–259. <https://doi.org/10.1007/BF00763097>
- Thermo Fisher Scientific. (2018). *Quantifiler HP and Trio DNA Quantification Kits User Guide (Pub. No. 4485354 Rev. H*

XI. ANEXOS

A. ANEXO

Fundación de Antropología Forense de Guatemala CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0070	mtDNA Sequencing	Revision: Original	Effective date:05/03/2019	Page 1 of 15



Mitochondrial DNA Sequencing using BigDye® Direct Cycle Sequencing Kit and BigDye® XTerminator™ Purification Kit Standard Operating Procedure

1.0 PRINCIPLE

This protocol describes all steps required to sequence human mtDNA for identification purposes. This SOP includes PCR amplification, cycle sequencing and post-PCR clean-up procedures using the BigDye® Direct Cycle Sequencing Kit and BigDye® XTerminator™ Purification Kit.

2.0 BACKGROUND

Analysis of mitochondrial DNA in forensic applications has enabled the identification of missing persons through comparison with additional maternal relatives. Most forensic applications are based on sequencing both hypervariable regions of the mtDNA (HVI and HVII). Sequencing of these regions has been commonly done using Sanger-type sequencing (STS), this method is based on the use of dideoxynucleotides (ddNTPs) as inhibitors of the DNA polymerase activity during chain elongation due to the absence of a 3'-hydroxyl group [9.1]. Currently fluorescently labeled terminating nucleotides are used allowing the detection of the different sequencing fragments by capillary electrophoresis (CE) [9.2]. At the end, an mtDNA haplotype is reported for each sample processed using an analysis software. In a forensic context, haplotype specific polymorphisms are the basis for identification and comparison between evidence and reference samples purposes.

The BigDye® Direct Cycle Sequencing kit provides the required reagents for the sequencing reaction of DNA. This kit is based on fluorescence-based cycle sequencing reaction to be detected using a Genetic Analyzer through CE [manual del kit]. By using universal sequencing primers, this kit is versatile to different sequencing purposes, including obtaining the sequence of HVI and HVII of mtDNA (see Appendix A for specific primer sequence). This kit combines PCR purification and cycle sequencing steps in the same reaction plate, enabling a streamlined and time-saving protocol.

The BigDye® XTerminator™ Purification Kit enables the elimination of unincorporated dye-labeled components that can interfere with electrokinetic injection, electrophoretic separation and data analysis. Components of this kit are compatible with BigDye® Direct Cycle Sequencing kit allowing the elimination of salt ions, unincorporated dye terminators and dNTPs, generating a high signal strength.

FAFG's validation studies demonstrated that PCR amplification and a good quality sequence is obtained with a genomic DNA concentration ranging from 0.015 - 1.78 ng/μL (DNA input of 0.015 - 1.78 ng). Nevertheless, the kit recommends a DNA concentration of 4.0 ng/μL;

DNA0070_mtDNA_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0070	mtDNA Sequencing	Revision: Original	Effective date:05/03/2019	Page 2 of 15



internal validation showed to generate good quality results with this DNA concentration as well. The reagents are designed and optimized to be used with the following Applied Biosystems instruments: Veriti® 96-Well Thermal Cycler and 3500 Genetic Analyzer.

3.0 SCOPE

This SOP applies for the PCR amplification and cycle sequence of Hypervariable Regions I and II from the mtDNA in DNA extracts from reference (buccal and blood samples extracted by Chelex and QIAGEN methods) and skeletal remains samples for identification purposes.

4.0 POLICY

This SOP was created to comply with the COGUANOR NTG/ISO/IEC 17025 accreditation, in order to sequence mtDNA correctly.

5.0 REAGENTS, SUPPLIES & EQUIPMENT

5.1. Kits:

- BigDye® Direct Cycle Sequencing Kit includes
 - BigDye® Direct PCR Master Mix
 - BigDye® Direct Sequencing Master Mix
 - BigDye® Direct M13 Fwd Primer
 - BigDye® Direct M13 Rev Primer
 - Control DNA CEPH 1347-02, 20-µL, 50-ng/µL

NOTE: All components, except BigDye® Direct PCR Master Mix should be stored at -15 to -25 °C. BigDye® Direct PCR Master Mix should be stored at -20 °C until first use, then 4°C. For optimal performance, freeze and thaw the BigDye® Direct Sequencing Master Mix no more than 5 times.

- BigDye® X Terminator™ Purification Kit includes
 - SAM™ Solution
 - X Terminator® Solution

NOTE: The X Terminator® Solution should be stored at 4°C. Avoid freezing the reagents. If the work time last more than 30 minutes, keep them on ice with the caps closed between pipetting steps. The SAM™ Solution should be stored at 4°C. Storage at room temperature may decrease the likelihood of precipitation.

5.2. Reagents

- Specific primers for the amplification of HVI and HVII (see Appendix A)

DNA0070_mtDNA_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0070	mtDNA Sequencing	Revision: Original	Effective date:05/03/2019	Page 3 of 15



- E-Gel® EX (1% Agarose Gels)
- E-Gel® 1 Kb Plus DNA Ladder
- Deionized water
- TE (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0)
- Bleach, 10% commercial

5.3. Equipment

- Micro or mini centrifuge
- Plate centrifuge
- Vortex
- Veriti® 96-Well Thermal Cycler
- E-Gel®iBase™
- 3500 Genetic Analyzers
- Freezer (at least -17 °C)
- Refrigerator (2 – 6 °C)
- Pipettes (as needed)

5.4. Supplies

- Calculator
- Sterile microcentrifuge tubes, 1.5-mL, or 2.0-mL, or other appropriate clean, unused tube (for Master Mix preparation)
- 96-well reaction plate for PCR
 - 96-well base
 - 96-well full plate cover (or reasonable substitute)
- Gloves, disposable
- Marker pen, permanent
- Microcentrifuge tube rack
- Tips, aerosol-resistant
- Waste container(s)
- Laboratory Coat
- Safety Glasses

6.0 SAFETY

- 6.1. Treat all biological samples as potentially infectious. Protective gloves, lab coats must be worn at all times when working of biological samples; the use of any other protective equipment, such as glasses, is optional. Frequent changing of gloves and/or re-bleaching of gloves throughout the procedure can reduce the occurrence of random and cross-contamination.
- 6.2. All work surfaces, utensils, and instruments used must be cleaned as necessary with 10% bleach.
- 6.3. All appropriate MSDS must be read PRIOR to beginning this procedure.
- 6.4. The head block of the thermal cycler can become hot during use. Be careful not to touch the heat block when working with the thermal cycler.

DNA0070_mtDNA_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0070	mtDNA Sequencing	Revision: Original	Effective date:05/03/2019	Page 4 of 15



7.0 QUALITY ASSURANCE

- 7.1. Always change pipette tips between liquid transfers. Use only aerosol-barrier pipet tips.
- 7.2. All samples must be appropriately labeled at all times.
- 7.3. At least one (1) PCR positive control (sample of known DNA profile and that is non-probative) must be run with each amplification reaction. It is recommended that two (2) positive controls be run per reaction should one of the controls fail. FAFG's requires that at least one (1) positive control amplify correctly per PCR reaction.
- 7.4. At least one (1) PCR negative control (sterile, deionized water or TE) must be run with each amplification reaction.
- 7.5. Refer to a supervisor or technical lead for questions regarding this procedure.
- 7.6. Ensure that all appropriate worksheets are completed throughout the procedure. This includes sample processing worksheet, reagent and equipment logs.

8.0 PROCEDURE

8.1. PCR Amplification using BigDye® Direct Cycle Sequencing Kit:

NOTE: Before using this kit make sure your PCR primers meet the design requirement: all specific primers must include the M13 universal primer sequences (see Appendix A) at the 5' end.

- 8.1.1. Determine the total number of samples, including controls.
- 8.1.2. Thaw and vortex the following reagent for 5 sec:
 - 8.1.2.1. Samples (genomic DNA),
 - 8.1.2.2. M13-tailed forward and reverse PCR primer set for each hypervariable region,
 - 8.1.2.3. BigDye® Direct PCR Master Mix
 - 8.1.2.4. Control DNA CEPH 1347-02; 20- μ L; 50-ng/ μ L. Dilute to 4-ng/ μ L before use.

NOTE: All components, except BigDye® Direct PCR Master Mix, should be stored at -15 to -25 °C. BigDye® Direct PCR Master Mix should be stored at -20 °C until first use, then 4°C. For optimal performance, freeze and thaw the BigDye® Direct Sequencing Master Mix no more than 5 times.

NOTE: Control DNA CEPH must be diluted to 4.0 ng/ μ L.

NOTE: M13-tailed PCR primers must be prepared at a concentration of 0.8 μ M.

- 8.1.3. Spin the tubes briefly in a microcentrifuge to remove any liquid from the caps.
- 8.1.4. Select a clean, unused tube for the master mix.
- 8.1.5. Prepare 2 reactions (one "reverse" and one "forward") for each sample. For EACH forward or reverse reaction, prepare your master mix by adding the following components for each sample/control:

DNA0070_mtDNA_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

Component	Volume if Gel WONT be run	Volume if Gel WILL be run
BigDye® Direct PCR Master Mix	5.00 µL	7.50 µL
Deionized water	2.50 µL	3.60 µL
M13-tailed Reverse primer (0.8 µM)	0.75 µL	1.20 µL
M13-tailed Forward primer (0.8 µM)	0.75 µL	1.20 µL
Total volume:	9.00 µL	13.50 µL

NOTE: The FAFG worksheets automatically include an extra sample to account for handling variation. The analyst may change this value as needed.

NOTE: Before preparing the master mix for amplification, consider that you must run a gel if a DNA extract was not quantified and/or for quantified samples with a DNA concentration below the optimal range (0.015 ng/µL).

- 8.1.6. Vortex the master mix at medium speed for 5 sec and quick-spin to remove liquid from the cap.
- 8.1.7. Dispense 9.00 µL (or 13.50 µL if a gel will be run) of master mix per well for reaction.
- 8.1.8. Dispense 1.00 µL (or 1.50 µL if a gel will be run) of mtDNA (0.015 – 1.78 ng/µL) to the appropriate well tube.

NOTE: The final volume in each PCR tube is 10 µL if a gel won't be run.

NOTE: The final volume in each PCR tube is 15 µL if a gel will be run.

- 8.1.9. Prepare a negative control by dispense 1.00 µL (or 1.50 µL if a gel will be run) deionized water or TE buffer instead of genomic DNA.
- 8.1.10. Prepare a positive control by dispense 1.00 µL (or 1.50 µL if a gel will be run) of Control DNA CEPH 1347-02 instead of genomic DNA. Recommended concentration 4.0 ng/µL (dilute stock concentration of 50.0 ng/µL accordingly).
- 8.1.11. Seal the plate with adhesive film or caps and then spin the briefly.
- 8.1.12. Place the tray in the thermal cycler.
- 8.1.13. Close the heated cover.
- 8.1.14. When prompted, select the program "Big dye direct PCR" and change the reaction volume to 10 µL or 15 µL, accordingly.
- 8.1.15. Start the thermal cycler
- 8.1.16. The PCR amplification conditions are as following:



Initial incubation step	Denature	Anneal	Extend	Final Extension	Final Step
Hold	35 cycles			Hold	Hold
95 °C 10 min	96 °C 3 sec	62 °C 15 sec	68 °C 30 sec	72 °C 2 min	4°C ∞

NOTE: This is a stopping point, you might store the amplified DNA at 4°C overnight or at -15 °C to -25 °C for long-term storage.

8.2. Gel Electrophoresis (Optional):

NOTE: The minimum quantity of PCR product to use for sequencing according to manufacturer’s recommendation is 20-ng. This can be checked by running on an agarose gel, if necessary. For optimized amplification this step can be omitted.

For FAFG, a gel is required if a DNA extract was not quantified and/or for quantified samples with a DNA concentration below the optimal range (0.015 ng/μL).

- 8.2.1. Prepare the appropriate number of 0.2 mL tubes (depending on the number of samples to run) with 15.0 μl of water.
- 8.2.2. Remove the plate seal and take 5.0 μl of each amplification reaction (remember that you have two reactions – one “reverse” and one “forward” – for each sample).
- 8.2.3. Add this 5.0 μl of sample to each 15.0 μl of deionized water/TE buffer/loading buffer.
- 8.2.4. Plug the E-Gel® iBase™ device into an electrical outlet using the adapter plug on the base.
- 8.2.5. Remove the gel from the package DO NOT remove the comb until you start loading the samples.
- 8.2.6. Insert the gel into the base by sliding it into the two electrode connections on the iBase™. Make sure that the two electrodes on the right side of the cassette are in contact with the two contacts on the iBase™. The Invitrogen logo should be located at the bottom of the base. A steady, red light illuminates on the base if the gel is correctly inserted.
- 8.2.7. Take out the comb and add your sample (20.0 μl of each sample prepared in step 8.2.4) into its appropriate well in the gel. Use the gel map available in the worksheet as reference.

NOTE: Always include a positive and a negative control, to verify that amplification reagents and equipment worked properly.



- 8.2.8. Include at least one Molecular Weight Marker per run: Add 20.0 µl of the 1 Kb DNA Ladder (10.0 – 12.0 ng/µl).
- 8.2.9. Use the Up/Down (▲\▼) buttons to select program: “RUN E-Gel 0.8-2.0%”.
- 8.2.10. Place the UV protector (orange cover) and press “Go” button. Run the gel for a maximum time of 15 min.

NOTE: You can monitor the run by pressing the red button on the bottom part of the iBase™. Make sure the orange cover is placed before activating the UV light or that you are wearing the orange goggles.

- 8.2.11. Document the results in the worksheet, results will be noted with “+” if a band is present and “-” if the band is absent. Multiple “+” signs denote the intensity of the band.

8.3. Cycle Sequencing using BigDye® Direct Cycle Sequencing Kit:

NOTE: You will use M13 forward and reverse primers included in BigDye® Direct Cycle Sequencing kit. Add M13 reverse primer for the “reverse” reaction; and M13 forward primer, for the “forward” reaction for each sample.

- 8.3.1. Calculate the required amount of components as shown in next table:

Component	Volume
BigDye® Direct Sequencing Master Mix	2.00 µL
One sequencing primer:	1.00 µL
BigDye® Direct M13 Reverse primer OR	
BigDye® Direct M13 Forward primer	
Total volume of Master Mix for each reaction:	3.00 µL

NOTE: The FAFG worksheets automatically include one extra reaction to account for handling variation. The analyst may change this value as needed.

- 8.3.2. Vortex the master mix at medium speed for 5 sec and quick-spin to remove liquid from the cap.
- 8.3.3. For each sequencing reaction, add 3 µL of the sequencing reaction mix to the appropriate well in the respective forward or reverse reaction well
- 8.3.4. Seal the reaction plate with adhesive film or caps, then spin the plate briefly.

DNA0070_mtDNA_SOP
 Printed copies of this document could be out of date and should be considered uncontrolled.



- 8.3.5. Place the tray in the thermal cycler.
- 8.3.6. Close the heated cover.
- 8.3.7. When prompted, select the program “BigDyeDir CycleSeq” and change the reaction volume to 13 µL.
- 8.3.8. Start the thermal cycler.
- 8.3.9. The PCR amplification conditions are as following:

Initial steps			Denature	Anneal	Extend	Final Step
Hold	Hold	Hold	25 cycles			Hold
37 °C	80 °C	96 °C	96 °C	50 °C	60 °C	4°C
1.5 min	2 min	1 min	10 sec	5 sec	75 sec	∞

NOTE: This is a stopping point, you might store the reaction plate at 4°C overnight or at -15 °C to -25 °C for long-term storage.

8.4. Purification of sequencing products using BigDye®XTerminator™ Purification Kit:

- 8.4.1. Spin the reaction plate at 100x g for 1 minute, then remove the seal.
- 8.4.2. Vortex the Xterminator® Solution bulk container at maximum speed for at least 10 seconds, until it is homogeneous.
- 8.4.3. Prepare a premix with SAM™ Solution and XTerminator® Solution in an appropriately sized tube accordingly to the following volumes:

Component	Volume
SAM™ Solution	61.40 µL
XTerminator® Solution	13.60 µL
Total volume of Master Mix for each reaction:	75.00 µL

NOTE: Make sure there are no particulates in the SAM solution before pipetting. If there are particulates, heat the SAM Solution to 37°C and mix to resuspend. Cool to room temperature before sing.

NOTE: Manufacturer recommends to use a wide-bore pipette tips to dispense XTerminator® Solution. If wide-bore pipette are not available, 1,000 µL tips can be cut previously.

NOTE: The FAFG worksheets automatically include a 5% overfill to account for handling variation. The analyst may change this value as needed.

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0070	mtDNA Sequencing	Revision: Original	Effective date:05/03/2019	Page 9 of 15



8.4.4. Vortex the XTerminator® Solution bulk container at maximum speed for at least 10 seconds, until the solution is homogeneous.

8.4.5. Add 75.0µL of SAM™ /XTerminator® solution premix to each well.

NOTE: Avoid pipetting from the top of the liquid and vortex the premix constantly, approximately every 4 wells.

8.4.6. Seal the plate with a septa mat and vortex the reaction plate for 30 minutes at 1800 – 2000 rpm.

8.4.7. Spin for 2 minutes at 1000xg.

NOTE: It is recommended to organize your time to place the plate for CE afterwards. If this is not possible, store the sample plates sealed with heat seal film or adhesive film for up to 48 hours at room temperature (20 to 25°C) or up to 10 days at 4°C or -20°C.

8.5. Capillary Electrophoresis:

8.5.1. After BigDye XTerminator purification, place the plate in the 3500 instrument

NOTE: Before starting the run make sure the instrument is set to perform sequencing applications; this includes: spatial and spectral calibration (for Dye Set Z), POP-7 installation, Instrument protocols, Assay, Run Module, Mobility and Calibration files specific for BigDye® Direct Cycle sequencing kit.

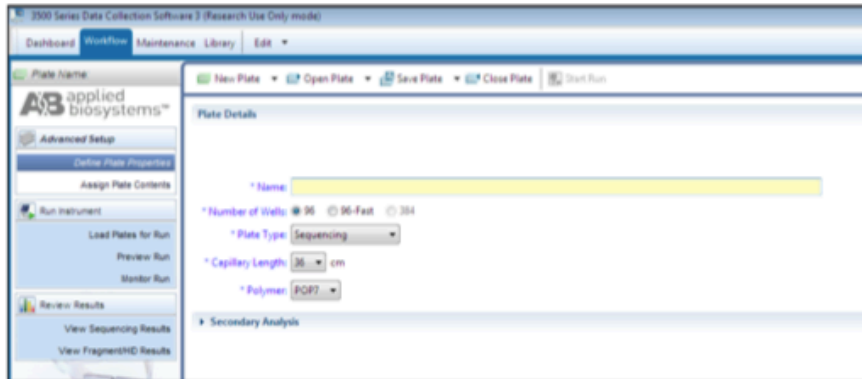
- Assay: "FastSeq_36_POP-7"
- Instrument Protocol: "BDXFast Seq 36_POP7_1"
- Basecalling Protocol: "BDTv3.1_PA_Protocol_POP-7"
- Dye: Z

Instrument Protocol Properties	Oven Temperature: 60 °C	Run Voltage: 13.4 kV	PreRun Voltage: 13.4 kV	Injection Voltage: 1.5 kV
	Run time: 2520 sec	PreRun Time: 180 sec	PreRun Time: 180 sec	Injection Time: 3 sec

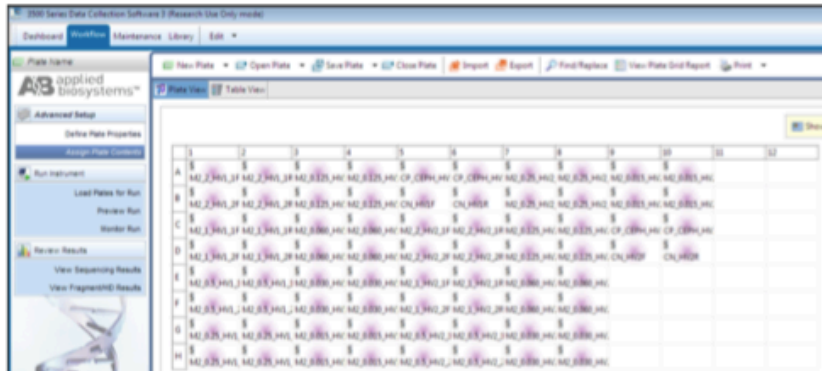
8.5.2. In the instrument, AB 3500, start by pre-heating the oven at 60°C on the "Dashboard" menu.

DNA0070_mtDNA_SOP
Printed copies of this document could be out of date and should be considered uncontrolled.

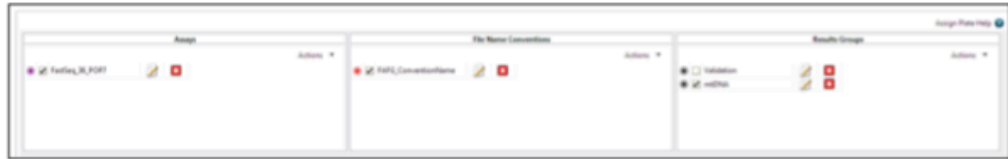
8.5.3. Go to the "Workflow" menu and on the "Define Plate Properties" option create your plate (you can also import it). Make sure that all Plate details, such as: "Plate type", "capillary length" and "Polymer" fields are filled out as the following screenshot:



8.5.4. Move forward to the "Assign Plate Contents" option, on this window create the plate set-up accordingly to the plate set-up in your worksheet (manually or automatically if available a plate converter was used).



8.5.5. Make sure to select the options as shown in the following screenshot:



8.5.6. Press "Link Plate for Run" and move forward to inject your samples

8.5.7. Once the injection has finished press the option "View Sequencing Results" and the list of all the samples that were sequenced from that run will appear as shown in the following screenshot:

Sample File Name	Lane#	Readout Version	Intensity File	Base Calling	First Peak Scan #	Start Scan #	End Scan #	Overall Data Time
1 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
2 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
3 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
4 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
5 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
6 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
7 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
8 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
9 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
10 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
11 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
12 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
13 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
14 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
15 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
16 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
17 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
18 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
19 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
20 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
21 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
22 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
23 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
24 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
25 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
26 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45
27 H01_A_2014_001_401	Pa	48.1.0.1.0	48_2000_P107_2071-0-mock	1.00	106	100	1071	20-Feb-2019 08:45

8.5.8. Select the "Sample File name" for each sample run (usually 4 files will be selected = 2 per region + 2 primers) and press the button "View Trace Reports" from the top options. A report will be generated with quality information about your samples. Save this report. Do the same with each sample from your run:



8.5.9. Open the File where your raw data is stored and save all the injection documents on your USB. Proceed to analyze sequencing information according to the SOP DNA0071_Use of Sequencer_SOP .

9.0 REFERENCES

1. Sanger, F., Nicklen, S., & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 1977; 74(12): 5463-5467.
2. Berglund, E. C., Kiiialainen, A., & Syvänen, A. C. Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet* 2011; 2: 23.
3. Applied Biosystems, Life Technologies (2011) BigDye® Direct Cycle Sequencing Kit Protocol. Carlsbad, CA.
4. Applied Biosystems, Life Technologies (2007) BigDye® XTerminator™ Kit Protocol. Foster City, CA.
5. Applied Biosystems, Life Technologies (2014) E-Gel Technical Guide.
6. Applied Biosystems, Life Technologies (2010) Applied Biosystems 3500/3500xL Genetic Analyzer User Guide. Foster City, CA.

DNA0070_mtDNA_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0070	mtDNA Sequencing	Revision: Original	Effective date:05/03/2019	Page 13 of 15



10.0 APPROVAL

This SOP was approved and released for use on 05/03/2019 by:

Mishel Stephenson, DNA Chief of Forensic Genetics

11.0 DOCUMENT REVISION HISTORY

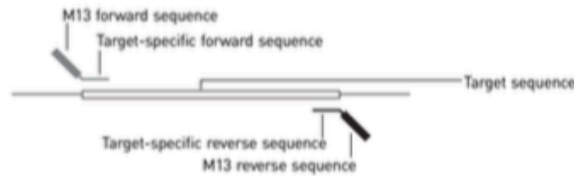
Revision	Revision Date	Reason for change	Section(s)	Reviser (s)
Original	05/03/2019	Initial release	N/A	Alicia Mazariegos y Wendy Argueta

END

DNA0070_mtDNA_SOP
 Printed copies of this document could be out of date and should be considered uncontrolled.

APPENDIX A: Sequence of primers for PCR Amplification of mtDNA

Primer design requirements:



- M13 forward primer sequence: 5' TGTA AACGACGGCCAGT 3'
- M13 reverse primer sequence: 5' CAGGAAACAGCTATGACC 3'

M13-tailed forward and reverse specific PCR primers


Region	Primer	Name	Internal Code	Sequence (5'-3')
HV 1	Forward	F15978	A1	TGT AAA ACG ACG GCC AGT CAC CAT TAG CAC CAA AAG CT
		F15989	N/A	TGT AAA ACG ACG GCC AGT CCC AAA GCT AAG ATT CTA AT
	Reverse	R16391	B1	CAG GAA ACA GCT ATG ACC GAG GAT GGT GGT CAA GGG AC
		R16509	N/A	CAG GAA ACA GCT ATG ACC AGG AAC CAG ATG TCG GAT AC
HV 2	Forward	F30	C1	TGT AAA ACG ACG GCC AGT CTC ACG GGA GCTCTC CAT GC
		F11	N/A	TGT AAA ACG ACG GCC AGT CTA TCA CCC TAT TAA CCA CTC
		F109	N/A	TGT AAA ACG ACG GCC AGT GCA CCC TAT GTC GCA GTA TCT GTC
		F312	N/A	TGT AAA ACG ACG GCC AGT CCC CGC TTC TGG CCA CAG CAC
		F317	N/A	TGT AAA ACG ACG GCCAGT CTT CTG GCC ACA GCA CTT AAA C
	Reverse	R408	D1	CAG GAA ACA GCT ATG ACC CTG TTA AAA GTG CAT ACC GCC A
		R99	N/A	CAG GAA ACA GCT ATG ACC AGC GTC TCG CAA TGC TAT CGC
		R180	N/A	CAG GAA ACA GCT ATG ACC AAT ATT GAA CGT AGG TGC GAT AAA T
		R306	N/A	CAG GAA ACA GCT ATG ACC GGG GTT TGG TGG AAA TTT TTT G
		R460	N/A	CAG GAA ACA GCT ATG ACC AGG GGA AAA TAA TGT GTT AGT
		R599	N/A	CAG GAA ACA GCT ATG ACC TTG AGG AGG TAA GCT ACA TA

DNA0070_mtDNA_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

B. ANEXO

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0071	Use of Sequencher	Revision:Original	Effective date: 05/03/2019	Page 1 of 6



Use of Sequencher Software for the Analysis of mtDNA Data from the ABI 3500

1.0 PRINCIPLE

Sequencher® 5.4.6 DNA Sequence Analysis Software from Gene Codes Corporation is a software used for sequencing analysis. For FAFG's purposes this software will specifically be used for mitochondrial DNA (mtDNA) typing. It is required to correctly analyze data that is generated using the BigDye™ Direct Cycle Sequencing Kit.

2.0 BACKGROUND

Biological samples are processed in the Forensic Genetics Laboratory (DGF) with the aim of obtaining a mtDNA haplotype. The DNA is extracted from the samples, quantified, amplified and the mtDNA fragments are separated, detected and sequenced using the ABI 3500 Genetic Analyzer. After electrophoresis, the data collection software stores information from each sample in an .abl file. This allows the analyst to see the mtDNA sequence results and assess the chromatograms, quality, QV20 value and fragment size.

FAFG has a full Sequencher® 5.4.6 license installed in a computer. Client licenses are run in a USB key in one computer.

3.0 SCOPE

This procedure applies to all samples that have been amplified in the FAFG's Forensic Genetics Laboratory for mitochondrial DNA analysis and run on an ABI 3500 Genetic Analyzer.

4.0 POLICY

All staff in the Forensic Genetics Laboratory working under ISO accreditation shall adhere to the procedure on the use of Sequencher software. The biological samples processed in the Forensic Genetics Laboratory with the aim of obtaining a mtDNA haplotype need to be analyzed and interpreted using the Sequencher® 5.4.6.

5.0 SUPPLIES & EQUIPMENT

- Sequencher® 5.4.6
- Windows Systems PC Platform Computer

NOTE: Make sure that your computer meets the following minimum recommended requirements: Windows 8 or later, 64-bits only, 3 GB RAM, 280 MB hard disk space.

6.0 SAFETY

DNA0071_Use of Sequencher_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.



Prolonged computer analysis may cause muscle soreness, eye fatigue, or other discomforts. In order to minimize the occurrence of discomforts, periodic breaks from computer analysis are encouraged.

7.0 QUALITY ASSURANCE

- 7.1. Ensure all worksheets are complete.
- 7.2. Ensure that the negative control fragments are less than or equal to 5 bp size.
- 7.3. Make sure that the positive control, CEPH from the Big Dye Direct kit, fragments are more than or equal to 404 bp length for HVI and 414 bp for HVII, a minimum of 50% of quality, complete comparison range coverage (these means two fragments [forward and reverse] per region [HVI and HVII]), compare consensus to reference rCRS (revised Cambridge Reference Sequence) and selected Consensus to Forensic Standards.
- 7.4. The positive control fragments MUST generate an haplotype equal to the reported for Control DNA CEPH (Table 1).

Table 1. Control DNA CEPH haplotype.

Variant	Reference	Contig[0139]	
1	16,069	C	T
2	16,126	T	C
3	16,291	C	A
4	73	A	G
5	152	T	C
6	185	G	A
7	188	A	G
8	228	G	A
9	263	A	G
10	295	C	T
11	309.1	:	C
12	315.1	:	C

- 7.5. For each sample the comparison range coverage is complete, compare consensus to reference rCRS and selected Consensus to Forensic Standards and a minimum quality of 50%.

NOTE: The quality could be less if the complement fragment has a high quality.

8.0 PROCEDURE

8.1. Open Sequencher and import sample files

- 8.1.1. Connect the hardware key (USB) before opening Sequencher.

- 8.1.2. Open Sequencher with double click at the icon



DNA0071_Use of Sequencher_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.



- 8.1.3. Select File> Save Project, the "Save As" window pop up and save the project as the batch number (example Sequencher_ddmmyyhhmm).
- 8.1.4. Import the reference rCRS by select File>New Project From Template > rCRS. This step can be repeated at is needed for controls and samples.
- 8.1.5. Import sample files by select File>Import>Sequences (select files .ab1 with the controls and samples).

8.2. Compare consensus to references

- 8.2.1. Select Contig>Consensus for Forensic Standards.
- 8.2.2. Select with Ctrl+ the rCRS and the fragments related to one sample or positive control.
- 8.2.3. Select Assemble>Assemble to Reference (Sequencher would align the fragments). A "Contig" is made in the project window. If the quality is too low the software would not take a fragment as part of the "Contig".

NOTE: The software would align fragments even if they not completely cover the region.

NNOTE: You can label the contig by clicking on "Contig" at the project window, it is recommended to change the name to FAFG's sample ID.

- 8.2.4. Double click on "Contig" at the project window will open the Contig window. Select the "Bases" option at the options bar. And then the "Show Chromatograms" option. Now the sequence and the chromatograms are available to edit the base calling.

8.3. Policy for editing the base calling

- 8.3.1. Only edit the base calling at the HVI and HVII regions.
- 8.3.2. Edit just the base called marked with "+".

NOTE: The "+" symbol under the consensus line flags an ambiguous position. The "" symbol under the consensus line flags a position where the base calling is different between the sample and the rCRS.

- 8.3.3. Make sure that the ambiguous position can be resolve by one of the fragments (reverse or forward). If it is not possible identify the fragment with the lowest quality and process it again.
- 8.3.4. If the two fragments present an overlap of peaks (Figure 1) in the same position, this can mean that the sample was contaminated or is heterozygous (event of heteroplasmy). This sample has to be processed again to discard a contamination event.

NOTE: ONLY the region with the peaks overlap should be processed again.



Figure 1. Peaks in case of contamination or heterozygous sample.

8.4. Editing the base calling

- 8.4.1. Check, from the "Edit" menu, that the "Replace When Editing" is selected.

NOTE: This option will just "Replace When Editing" so if an insertion is needed go to the "Edit" menu and select the "Insert When Editing". The insertion would be placed before the selected position.

- 8.4.2. Select the base at the Consensus position (this allow to select the base position in both fragments).

NOTE: From the "Select" menu, choose the command "Next Ambiguous Base" or by pressing "Ctrl+N" so automatically can move to the next ambiguous position.

- 8.4.3. Press the letter "A", "C", "T" or "G" to change the base calling when the problem at the ambiguous position is a base call error or the software identifies a deletion instead of a base.

NOTE: The ":" at the sequence denots a deletion.

- 8.4.4. Press "Backspace" to make a deletion at the sequence.

NOTE: A window will pop up (Figure 2). Select "OK". This option will just "Replace When Editing" so if an insertion is needed go to the "Edit" menu and select the "Insert When Editing".



Figure 2. Window in case of deletion.

8.5. Haplotype report

8.5.1. Return to the project window and select all the Contigs needed for the report.

NOTE: Select multiple Contigs by Ctrl+.

8.5.2. Select from the menu Contig>Compare Consensus to Reference.

8.5.3. At the new window select Comparison Range then check that the "Filter Comparison by:" option is selected, press "Select All" and then press "OK".

8.5.4. For a report: from the menu select "Reports", at the "Variance Table Reports" window select "Entire Table" option and set the "Report format" as "Variance Detail Report".

NOTE: for a simple report select set the "Report format" as "Individual Variance reports".

8.5.5. Then Select the "Save as PDF" to export the report as a .pdf file. Save the Report as the batch number (example: Haplotypes_#).

9.0 REFERENCES

9.1. Gene Codes Corporation. (2017). Tour Guide for Windows and Macintosh Tour Guide. Retrieved marzo 2019, 6, from Gene Codes: www.genecodes.com/sites/default/files/documents/Tutorials/Tour%20Guide.pdf

Fundación de Antropología Forense de Guatemala				
CONFIDENTIAL - Not to be copied or distributed without authorization.				
DNA0071	Use of Sequencher	Revision:Original	Effective date: 05/03/2019	Page 6 of 6



10.0 APPROVAL

This SOP was approved and released for use on 05/03/2019 by:

Mishel Stephenson, DNA Chief of Forensic Genetics

11.0 DOCUMENT REVISION HISTORY

Revision	Revision Date	Reason for change	Section(s)	Reviser (s)
Original	05/03/2019	Original Release	All	Alicia Maria Mazariegos Medina

END

DNA0071_Use of Sequencher_SOP

Printed copies of this document could be out of date and should be considered uncontrolled.

C. ANEXO

Cuadro 11. Haplotipos para diferentes concentraciones de una muestra ósea.

Muestra	N	n	HV1 (16024-16365)	HV2 (73-340)	Frecuencia (%)
7h	7	2	16,111T 16,187T 16,209C 16,223T 16,290T 16,319A 16,362C 16,365T	73G 146C 153G 163G 309.1C 309.2C 315.1C	29

Se evaluaron 7 (N) concentraciones (de 4.0, 1.0, 0.5, 0.1, 0.05, 0.025 y 0.005 ng/ μ L), para las cuales solamente 2 (n) presentaron haplotipos reportables (1.0 y 0.5 ng/ μ L).

Cuadro 12. Haplotipos de muestras procesadas por analistas diferentes.

Muestra	n	HV1 (16024-16365)	HV2 (73-340)	Frecuencia (%)
CEPH	2	16,069T 16,126C 16,291A	73G 152C 185A 188G 228A 263G 295T 309.1C 315.1C	100
8	2	16,111T 16,290T 16,319A 16,362C	73G 146C 153G 235G 263G 309.1C 315.1C	100
9	2	16,223T 16,290T 16,319A 16,362C	73G 146C 153G 182T 184G 235G 263G 309.1C 315.1C	100

Las muestras fueron procesadas por el Analista 1 en octubre del 2018 y por el Analista 2, en febrero del 2019.