

Universidad Nacional
Facultad de Filosofía y Letras
Sistema de Estudios de Posgrado
Escuela de Literatura y Ciencias del Lenguaje
Maestría Profesional en Traducción Inglés-Español

Traducción e Informe de Investigación

**LOS PROCESOS SINTÁCTICOS DE INTERFERENCIA LINGÜÍSTICA EN LA
TRADUCCIÓN INVERSA (ESPAÑOL-INGLÉS) DEL SINTAGMA NOMINAL
EN EL GÉNERO CIENTÍFICO**

Trabajo de graduación para aspirar al grado
Magíster en Traducción
(Inglés – Español)

presentado por

MARTA MARÍA PANIAGUA RAMÍREZ
Cédula 1-1030-0923
Carné # 204909-0

2007

La traducción que se presenta en este tomo se ha realizado para cumplir con el requisito curricular de obtener el grado académico de la Maestría en Traducción Inglés – Español, de la Universidad Nacional.

Ni la Escuela de Literatura y Ciencias del Lenguaje de la Universidad Nacional, ni el traductor, tendrán ninguna responsabilidad en el uso posterior que de la versión traducida se haga, incluida su publicación.

Corresponderá a quien desee publicar esa versión gestionar ante las entidades pertinentes la autorización para su uso y comercialización, sin perjuicio del derecho de propiedad intelectual del que es depositario el traductor. En cualquiera de los casos, todo uso que se haga del texto y de su traducción deberá atenerse a los alcances de la Ley de Derechos de Autor y Derechos Conexos, vigente en Costa Rica.

Dedicatoria

A Jesús, Salvador de la humanidad.

A José Andrés Díaz Loría, mi esposo.

A Humberto Paniagua Baltodano † y María del Carmen Ramírez Martínez, mis
padres, por cimentar mis bases espirituales y académicas

Agradecimientos

A Dios por la salud, la capacidad y los medios para realizar este proyecto.

A José Andrés Díaz Loría, mi esposo, por su constante e incondicional colaboración espiritual, instrumental y afectiva.

A Sherry Gapper por su apoyo académico y moral, y por su pronta y precisa respuesta a las interrogantes sobre mi segunda lengua.

A Judith Tomcsányi por su perspicacia didáctica y agudo consejo lingüístico, y por la confianza y comprensión que me mostró durante el proceso.

A Bianchinetta Benavides por su estímulo académico y afectivo, y por tenderme su mano de forma incondicional.

Resumen

En este proyecto se traducen tres artículos científicos sobre los análisis de ADN humano en el campo forense¹. Este tema es actual y relevante en el mundo por su diversidad de aplicaciones y continuo progreso. La traducción de artículos científicos relacionados con esta área es fundamental para facilitar la comunicación en la comunidad de expertos. Dada la condición de lengua franca del inglés en el área científica-tecnológica, la traducción inversa (español-inglés) desempeña un papel protagonista en este proceso.

La traducción inversa, tal como se plantea en este proyecto, supone el uso de estructuras aceptables en la lengua de llegada, no sólo desde el punto de vista gramatical, sino también estilístico. Un análisis lingüístico descriptivo comparado de la forma, función y distribución del sintagma nominal del corpus de textos científicos en español e inglés permite explorar y clasificar algunas de las causas de las interferencias sintácticas sutiles en la traducción inversa, las cuales se relacionan con la transferencia de estructuras de la lengua materna y con mecanismos de hipercorrección y sobregeneralización. Mediante el análisis de estos fenómenos lingüísticos es posible trazar modelos y representar algunos mecanismos psicotraductológicos sobre los procesos de interferencia sintáctica.

Descriptor: Sintagma nominal, lingüística descriptiva, lingüística comparada, constituyentes inmediatos, estilo, interferencia lingüística, hipercorrección, sobregeneralización, género científico-técnico.

¹ Timken, Mark D., Swango, Katie L., Orrego, Cristián y Martin R. Buoncristiani. "A Duplex Real-Time qPCR Assay for the Quantification of Human Nuclear and Mitochondrial DNA in Forensic Samples: Implications for Quantifying DNA in Degraded Samples". *Journal of Forensic Sciences*. L., setiembre del 2005, vol. 50, No. 5.

Miller, Sasha., King, Sheria y Kalai Mathee. "An Overview of Human DNA Analysis in Forensics". *Revista de Ciencias Forenses*. Asociación Costarricense de Ciencias Forenses (ACCF), noviembre del 2006, vol. 1, No. 1.

Entrala, Carmen. *Técnicas de análisis del ADN en genética forense*. 2000.
<<http://www.ugr.es/~eianez/Biotecnologia/forensetec.htm>>. 4 de agosto del 2006.

Índice General

Resumen	v
Traducciones	1
<i>Visión general sobre el análisis forense del ADN humano</i>	<i>2</i>
<i>Ensayo dúplex de qPCR en tiempo real para cuantificar el adn nuclear y mitocondrial humano en muestras forenses: implicaciones para la cuantificación del adn en muestras degradadas.....</i>	<i>29</i>
<i>DNA Analysis Techniques in Forensic Genetics.....</i>	<i>88</i>
Informe de investigación.....	109
Introducción	110
Capítulo 1: Marco Teórico.....	117
Capítulo 2: Descripción y comparación lingüística de los sintagmas nominales del inglés y el español	132
Capítulo 3: Factores de interferencia lingüística en la traducción inversa (español-inglés) del sintagma nominal	157
Conclusiones	184
Bibliografía.....	195
Apéndice: Textos originales.....	200

Traducciones

VISIÓN GENERAL SOBRE EL ANÁLISIS FORENSE DEL ADN HUMANO

Sasha Miller^{1,2}, Sheria King^{1,3} and Kalai Mathee^{1*},

¹*Department of Biological Sciences, Florida International University,
Miami, FL 33199, USA*

Dirección actual:

²Especialista y analista de ADN, *Forensic Science Center, Culloden Road, St.
Michael, Barbados*

³*Royal Bahamas Police Force, East Street, Nassau, Bahamas.*

Título abreviado: Estudios forenses del ADN humano

Descriptores: VNTR, STR, Y-STR, SNP, ADNmt, PCR

* Correspondencia:

Dr. Kalai Mathee, Ph.D.

Profesora adjunta

Department of Biological Sciences

Florida International University

University Park

Miami, FL 33199

Tel: (305) 348 1261

FAX: (305) 348 2913

Correo electrónico: matheek@fiu.edu

VISIÓN GENERAL SOBRE EL ANÁLISIS FORENSE DEL ADN HUMANO

Resumen: Las variaciones genéticas entre los individuos de una especie particular son un hecho fundamental de la vida; sin embargo, las semejanzas entre los genomas, aun entre especies no relacionadas, resultan muy sorprendentes. Las investigaciones realizadas por Alec Jeffreys identificaron las semejanzas y las diferencias del ADN humano como una huella genética única para cada individuo. A raíz de este hallazgo, los análisis de ADN se introdujeron en el campo legal. Desde entonces, las técnicas de análisis de ADN han continuado evolucionando y han llegado a ser más discriminatorias. Este artículo presenta una visión conjunta sobre las características secuenciales que los analistas forenses suelen emplear para la identificación del ADN humano, tales como (por sus siglas en inglés): el número variable de repeticiones en tándem (VNTR) (*Variable Number of Tandem Repeats*), las repeticiones cortas en tándem (STR) (*Short Tandem Repeats*), el ADN mitocondrial, los cromosomas Y (o los Y-STR) (*Y-Short Tandem Repeats*) y los polimorfismos nucleótidos simples (SNP) (*Single Nucleotide Polymorphisms*). Los estudios forenses de ADN se pueden definir como la utilización de variaciones moleculares medidas con el fin de identificar y distinguir las muestras de distintos individuos. El reconocimiento legal de los perfiles de ADN como pruebas admisibles se ha facilitado mediante i. el establecimiento de bases de datos nacionales de ADN como *Combined DNA Information System* (CODIS) y *Human SNP Reference Database* (DbSNP), y ii. el desarrollo de paquetes comerciales de marcadores

genéticos (*kits*) multiplex extensos y reproducibles. Estos avances han permitido que la técnica de análisis de ADN surja como una herramienta potente e invaluable para las investigaciones criminales, tanto para imputar como para absolver a los implicados.

INTRODUCCIÓN

El análisis forense del ADN se ha utilizado en forma extensa en la identificación humana [1]. A partir de la secuenciación del genoma humano se ha determinado que este comprende alrededor de 3×10^9 de pares de bases de nucleótidos y los individuos tienen en común la gran mayoría del ADN humano (99,9%) [2, 3]. Sin embargo, en el caso de los seres humanos, el 0,1% (un par de bases por cada 1000) de las regiones del ADN contienen polimorfismo (lo permite que un individuo sea único) en cantidad suficiente para utilizarse como una herramienta para la identificación humana [4]. Estos polimorfismos fundamentan el análisis del ADN. Por lo tanto, los estudios forenses sobre el ADN se conceptualizan en la utilización de las medidas de las variaciones moleculares que permitan identificar y distinguir las muestras de cada persona.

Las pruebas forenses de ADN aparecieron por primera vez en el sistema judicial de Estados Unidos en 1986 en el caso *Florida vs Andrew*, y desde entonces se ha convertido en el foco de atención del público [5]. El análisis forense del ADN ha progresado sorprendentemente y ha conformado protocolos sólidos, eficientes y precisos. Como consecuencia, en la actualidad los tribunales suelen aceptar estos ensayos como pruebas concluyentes. Algunas bases de datos como el sistema

CODIS o Sistema Combinado de Índice de ADN son útiles para resolver crímenes y enjuiciar a los culpables, aunque también son de utilidad para liberar personas que han sido inculpadas por equivocación [6]. El análisis del ADN también se utiliza en otros casos de identificación humana tales como la prueba de paternidad, las indagaciones sobre personas desaparecidas, la identificación de soldados desconocidos y las investigaciones de sucesos de destrucción masiva. En este artículo se estudian las técnicas que los analistas forenses suelen utilizar para la identificación humana.

El número variable de repeticiones en tándem (VNTR)

Los genomas eucarióticos están conformados por secuencias de ADN repetidas y de distintos tamaños, las cuales por lo general se determinan con base en la longitud de la unidad de repetición y el número de secuencias adyacentes, o bien, a partir de la longitud de toda la región que se repite [7, 8]. Estas regiones que también se conocen comúnmente como ADN satélite, se encuentran dispersas a lo largo del genoma y se clasifican tanto como mini o microsatélites. Los minisatélites también se conocen como el número variable de repeticiones en tándem (VNTR), y consisten de unidades de repetición que presentan una gama de longitud de bases aproximado de 8-100 pb [9]. Por lo general, la región del VNTR comprende un fragmento de ADN de unos pocos miles de bases de longitud. Por otro lado, las repeticiones cortas en tándem (STR), o los microsatélites son versiones más pequeñas de las secuencias del VNTR y constituyen regiones del ADN que contienen unidades de repetición de 2-7 pb, con una longitud total que comúnmente varía entre 100 y 450 pb [6, 10].

Alec Jeffreys, primer experto en descubrir que algunas regiones del ADN contienen secuencias que se repiten en tándem [11, 12], concluyó que el número de repeticiones que se encuentran en cualquier sitio podía variar entre los individuos, y a partir de estos hallazgos se formularon los análisis del ADN [11, 12]. Como consecuencia, Jeffreys puso en práctica una técnica para examinar la variación de la longitud de esas secuencias repetidas del ADN o del VNTR que se localizan entre dos sitios de restricción. Esta técnica supuso el empleo de enzimas de restricción idóneas para separar las regiones del ADN que rodeaban a los VNTR, lo cual generaba fragmentos de longitud variable que eran distintos entre los individuos. Por lo tanto, esta técnica se denominó polimorfismo de la longitud de fragmentos de restricción o RFLP por sus siglas en inglés (*Restriction Fragment Length Polymorphism*). Los fragmentos del VNTR se separaron por medio de electroforesis en gel de agarosa y se detectaron por medio de la técnica de hibridación con sondas. Tal procedimiento consiste en transferir el ADN desde un gel hacia una membrana de nylon, para luego identificarlo con sondas radioactivas o luminiscentes de detección. De esta manera, la variación del ADN queda expuesta en los fragmentos de longitud variable o alelos, o en las distintas posibilidades para un locus en particular [13].

La tecnología del VNTR ofrece más estabilidad que los métodos proteicos y séricos utilizados con anterioridad en el área forense [6]. La hemogenética forense (el sistema o grupo ABO) fue la primera herramienta genética utilizada para distinguir individuos. Pese a ser un método rápido, no era muy eficaz pues sólo hay cuatro grupos fenotípicos posibles: A, B, AB y O [14]. Además, el 40% de la población de

Estados Unidos es de tipo O [14]. De esta forma, la utilidad de los grupos sanguíneos ABO se limitaba a descartar la posibilidad de que un individuo fuera la fuente de alguna muestra tomada en el lugar de los hechos, y no era muy útil cuando la persona resultaba involucrada, especialmente si la muestra era del tipo O. Por lo tanto, la combinación de varios loci de VNTR permitió establecer mayor capacidad de discriminación entre los individuos, y por ende, los resultados de las pruebas de VNTR adquirieron gran aceptación en los tribunales.

En los análisis de ADN se examinan varios loci de VNTR, pues cuanto más sea la cantidad de marcadores de ADN que se reconozcan y comparen, mayor será la probabilidad de encontrar diferencias genotípicas entre personas sin parentesco. Asimismo, cada pieza de información aumenta el grado de certidumbre sobre el carácter vinculante de un mismo individuo en relación con varias muestras. En particular, este tipo de pruebas fueron útiles para los casos de identificación humana, pues la combinación de varios VNTR permite establecer diferencias considerables entre los individuos [6, 1]. Aun así, éste método presenta bastantes limitaciones serias. Dado que las secuencias del VNTR son relativamente largas, es necesario obtener datos confiables de ADN genómico no degradado, el cual es difícil de encontrar en las muestras [15, 6]. Además, como las diferencias de tamaño entre los alelos adyacentes son pequeñas, se requiere que éstos se agrupen en varios paquetes, lo cual complica el análisis estadístico [6, 16]. Por último, se considera que este método tarda mucho tiempo, en comparación con otros.

Repeticiones cortas en tándem (STR)

A lo largo de los años, el análisis forense del VNTR evolucionó hacia otro método que utiliza secuencias de repetición más cortas, las cuales se conocen como STR o microsatélites. El ADN microsatélite se encuentra disperso a lo largo del genoma humano, el cual está compuesto por tres mil millones de pb, y aparece en promedio una vez cada 10 000 nucleótidos [12, 17]. Con el tiempo, se ha caracterizado y compilado un número considerable de marcadores de STR para conformar la base de datos conocida en inglés como *Short Tandem Repeat DNA Internet Database*, la cual se denomina comúnmente como STRBase por su nombre en inglés [18]. Desde su inicio, la STRBase ha sido un recurso informativo sobre los marcadores de STR más comunes para la comunidad que realiza análisis forenses de ADN [18].

Las secuencias STR se denominan sistemáticamente según la longitud de las unidades de repetición. Por ejemplo, las unidades de repetición de los dinucleótidos y trinucleótidos constan de dos o tres nucleótidos repetidos en forma recurrente y consecutiva. Actualmente, entre otros tipos de sistemas de STR, las secuencias de repetición de tetranucleótidos son las más utilizadas para fines forenses [6]. Estas secuencias son más fáciles de analizar debido a que son muy polimórficas, mientras que las secuencias de los penta y hexanucleótidos son menos comunes en el genoma humano [1, 19, 20, 21, 22].

El empleo de la reacción en cadena de la polimerasa (PCR por sus siglas en inglés) (*Polymerase Chain Reaction*) ha revolucionado el campo de la biología molecular. Desde que Kary Mullis la describió inicialmente en 1986, los laboratorios

forenses y de análisis de ADN se han beneficiado mucho por su capacidad de crear millones de copias de una secuencia específica de ADN en tan sólo unas pocas horas [23]. La tecnología de amplificación del ADN conocida como PCR es muy apta para el análisis forense de muestras de ADN, debido a que es muy sensible y rápida, y además el análisis no está limitado por la calidad del ADN, tal como sucede con los métodos de RFLP [1, 6, 12]. Este método es muy útil porque el ADN que se recolecta en las escenas de crimen suele ser limitado tanto en cantidad como en calidad. Al principio, los fragmentos de STR se detectaban sobre un gel de agarosa por medio de un método de tinción de plata. Sin embargo, esta técnica fue desplazada por el método de detección por fluorescencia, el cual es más sensible y apropiado [24, 25]. En este método, el producto que se amplifica en la PCR se identifica al aplicar un tinte fluorescente en el amplicón a lo largo del extremo 5' de un oligonucleótido cebador (conocido en inglés como *primer*), y después los diversos fragmentos de STR identificados por este método se analizan según la longitud que los separa. Los métodos de separación actuales incluyen la electroforesis en geles verticales de agarosa o poliacrilamida y la electroforesis capilar, conocida también como CE por sus siglas en inglés (*Capillary Electrophoresis*) [24].

Es posible analizar varios fragmentos de STR en forma múltiple; es decir, amplificarlos simultáneamente, pues la PCR se lleva a cabo en un tubo que contiene distintos cebadores que a su vez presentan identificadores fluorescentes diferentes [6, 26, 27]. El análisis múltiple permite detectar esos fragmentos en forma consecutiva y simultánea a lo largo de una sola línea [6]. Los fragmentos de STR también se pueden analizar fuera de la línea, después de haber realizado los

ensayos, ya sean los de electroforesis en geles verticales de poliacrilamida por medio de escáneres de fluorescencia, o bien, los que emplean secuenciadores basados en geles o en electroforesis capilar [20, 28, 29]. En los Estados Unidos, las dos plataformas principales de instrumentos utilizadas para la detección de los alelos STR por fluorescencia se conocen por sus nombres en inglés como la ABI 310 *Genetic Analyzer* y la Hitachi FMBIO II [6].

Los análisis múltiples de STR son ensayos sensibles de detección por fluorescencia. Este sistema es valioso porque se puede llegar a automatizar, lo cual sería una importante ventaja ante la creciente demanda de pruebas de ADN [20]. Cuando se analizan los sitios de varios cromosomas, los STR arrojan resultados que permiten establecer criterios de diferenciación muy considerables para distinguir a individuos no emparentados y aun a aquellos con un parentesco muy cercano [16, 27]. Además, los alelos discretos causan que el análisis de resultados sea más fácil de interpretar y comparar respecto a los sistemas de RFLP debido a que se utilizan bases de datos de ADN computarizados, mientras que en el caso de los RFLP se requiere agrupar o clasificar el ADN por tamaños similares [1, 6, 16].

Para que el análisis del ADN se llegara a aceptar entre un gran número de jurisdicciones fue necesario caracterizar una serie común de marcadores estandarizados. El laboratorio del FBI estableció 13 loci de STR centrales para que se incluyeran en la base de datos estadounidense de ADN conocida como CODIS (Cuadro 1) [6, 30, 31]. Cuando todos los 13 loci de STR centrales de la CODIS se examinan, la probabilidad de coincidencia aleatoria promedio para dos individuos no emparentados resulta ser menor que uno en un billón [16, 27, 31, 32]. En la

actualidad, las compañías estadounidenses *PE Applied Biosystems* y *Promega Corporation* cuentan con paquetes comerciales de marcadores de STR que les permite atender las necesidades de la comunidad de análisis forense del ADN, y abarcan una colección común de loci de STR (Cuadro 2) [9, 33, 34].

Con la invención de la tecnología PCR, que requiere de cantidades pequeñas de ADN de plantilla o de molde (por lo general 1 ng), se puede analizar ADN degradado parcialmente, dado que los fragmentos que se utilizan son pequeños y propensos a permanecer intactos en comparación con los del VNTR [6]. Además, en el análisis de STR se elimina la marginación de los alelos más largos en los marcadores minisatélites, producto de la amplificación preferencial de los alelos más pequeños [1]. Esto sucede porque en los individuos heterocigotos ambos alelos de STR tienen dimensiones similares, pues el tamaño de las unidades de repetición suele ser pequeño.

Asimismo, la tecnología para detectar fragmentos de longitud variable es cada vez más precisa. Actualmente, la electroforesis capilar con detección por fluorescencia y la electroforesis en geles verticales de poliacrilamida brindan una resolución de pares de bases para medir fragmentos de longitud superior a los 500 nucleótidos [24]. Por lo tanto, por razones biológicas y tecnológicas, los STR son más ventajosos que los minisatélites del VNTR debido a su menor tamaño. Además, los análisis de STR son la herramienta más utilizada hoy en día para la identificación forense del ADN, y probablemente seguirán realizándose durante muchos años porque forman parte de las bases de datos del ADN, las cuales se encuentran en constante crecimiento en todo el mundo.

Análisis del ADN mitocondrial (ADNmt)

En los casos en que el análisis múltiple de STR convencional no brinda resultados convincentes debido a que la muestra se encuentra muy degradada, se puede emplear el ADN mitocondrial (ADNmt) para obtener información útil. Esto sucede porque las células contienen mayor número de copias de ADNmt que de ADN nuclear, y por lo tanto las primeras tardan más en degradarse a lo largo de la muestra [1, 6]. La secuencia del ADNmt humano se determinó por primera vez en el laboratorio de Frederick Sanger en 1981 [35]. Este tipo de ADN contiene un genoma circular pequeño de 16569 pb, que al igual que el genoma nuclear, consiste de una región codificante y otra no codificante [35]. En la actualidad la secuencia original suele denominarse la secuencia de Anderson o la secuencia de referencia de Cambridge (CRS por sus siglas en inglés) (*Cambridge Reference Sequence*), la cual es útil para comparar nuevas secuencias [36]. No fue sino hasta 1993, que el ADNmt se utilizó por primera vez como una herramienta para la identificación forense y desde entonces su análisis se ha aceptado ampliamente en los tribunales de justicia [37, 38].

El ADNmt es un marcador útil para las pruebas de identificación humana porque posee una “región de control” no codificada de 1100 pb, la cual también se conoce como el lazo D (*D-loop* en inglés), y se caracteriza por presentar un grado razonable de variación de secuencias entre los individuos [1, 6]. Se han identificado dos regiones hipervariables dentro del lazo D, y suelen denominarse como la Región Hipervariable I y la Región Hipervariable II (HVI y HV2 respectivamente por sus iniciales en inglés). Debido a que la mayor parte de los polimorfismos del ADNmt se

encuentran dispersos a lo largo de las regiones hipervariables del lazo D, los científicos forenses examinan tanto el HV1 como el HV2 mediante amplificación de PCR y luego mediante el análisis de secuencias [1, 6, 39].

El ADNmt es muy útil para los casos forenses en los que las muestras recolectadas se encuentran degradadas en extremo o cuando se requiere asociar individuos con parentesco materno [39, 6]. Como se mencionó anteriormente, las células contienen varias copias de ADNmt, y sólo existe una copia del ADN nuclear, lo cual hace que su utilización sea ventajosa [1, 6]. Por lo tanto, el uso del ADNmt como un marcador alternativo permite que los científicos forenses obtengan valiosa información a partir de muestras que contienen cantidades ínfimas de ADN, o bien, que se han expuesto a condiciones ambientales extremas [40].

El ADNmt lo hereda la madre [6, 39], pues el espermatozoide no aporta ninguna mitocondria al cigoto durante la fertilización. Por eso, salvo en casos donde ocurra una mutación, el tipo de ADNmt no es único para cada persona. Los hermanos y familiares de parentesco materno tienen en común una secuencia idéntica de ADNmt. Aunque esto disminuye la posibilidad de identificar individuos en los casos forenses, es muy útil para resolver investigaciones sobre personas desaparecidas y sucesos de destrucción masiva. En la actualidad, las pruebas de ADNmt han desempeñado un papel importante para asociar parentescos familiares y resolver enigmas históricos como la identificación de los soldados desconocidos vietnamitas y el caso de la familia Romanov en los Estados Unidos [9, 40, 41].

Dado que el análisis de ADNmt requiere mucho más trabajo y dinero que el análisis de STR, las comparaciones de ADNmt se utilizan principalmente para casos

forenses en los que los marcadores de ADN nucleares fallan [6, 37, 42, 43]. La heteroplasmia, es decir, la presencia de más de un tipo de ADNmt por cada individuo es uno de los factores principales que limitan la implementación generalizada de la prueba del ADNmt en la resolución de casos [6,44]. La elección de tejidos así como del número de muestras complica aun más este tipo de prueba. Por ejemplo, las raíces de los distintos elementos pilosos de los individuos presentan grados variables de heteroplasmia [45, 46].

Se compiló una base de datos de secuencias de HVI y HVII de ADNmt a partir de más de mil personas sin relación de parentesco y procedentes de diversos grupos de población [47]. Además, con el objetivo de sistematizar el análisis de ADNmt, se desarrolló un material de referencia normativo sobre el ADNmt humano conocido como SRM por sus siglas en inglés (*Standard Reference Material for Quality Control in Human Mitochondrial DNA Sequencing*) [48, 49]. Estas bases de datos propician control de calidad entre los laboratorios cuando se examinan las secuencias de ADNmt para realizar identificaciones forenses. El continuo crecimiento de estas bases de datos es considerable puesto que la secuencia del ADNmt se considera como un único locus. En consecuencia, la tarea actual de reportar la rareza de algún tipo de ADNmt entre personas no emparentadas consiste en contar el número de veces en el que una secuencia específica aparece en la base de datos [50]. Por lo tanto, cuanto mayor sea el número de personas no emparentadas en la base de datos, mejores serán las estadísticas para calcular la probabilidad de coincidencia aleatoria entre dos individuos. Se ha demostrado que en el ADNmt existe heteroplasmia tanto de longitud como de secuencia, y esto se puede utilizar para

aumentar la capacidad de discriminación del análisis, y por ende, para fortalecer las pruebas de ADNmt [6, 51].

Sistemas del cromosoma Y

El ADNmt no es el único marcador genealógico disponible para el analista forense del ADN. Los marcadores del cromosoma Y también se transmiten de generación en generación sin cambio alguno excepto cuando ocurre una mutación [1, 6], por lo cual resultan útiles para obtener información sobre el parentesco paterno. Los polimorfismos Y-STR se describieron por primera vez a principios de la década de 1990 y su análisis llegó a ser un método popular para la identificación de casos forenses [52, 53]. El cromosoma Y permanece invariable cuando se transmite de padre a hijo, y varía menos entre los hombres en comparación con el STR autosómico. Por lo tanto, los marcadores múltiples de Y-STR se utilizan para establecer con gran precisión las diferencias entre los individuos. Hasta hace poco tiempo, la utilización del cromosoma Y para fines forenses estaba limitada ante la falta de marcadores polimórficos [6]. Sin embargo, en la actualidad se ha identificado gran cantidad de STR de este cromosoma y se están elaborando paquetes comerciales de marcadores genéticos multiplex reproducibles y extensos [54, 55, 56, 57]. En un análisis detallado reciente, Hanson y Ballantyne informaron sobre el uso potencial de al menos 389 marcadores separados de Y-STR [58+]. La corporación Promega Corporation cuenta con paquetes comerciales de marcadores de Y-STR que amplifican 12 loci de Y-STR (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 y DYS439) [9, 33, 34]. Dado

que los paquetes comerciales de marcadores para el análisis de los Y-STR acaban de introducirse al mercado, su utilización en los casos forenses rutinarios es limitada [59]. Los Y-STR son indicados particularmente para investigar los casos de violación sexual, dado que aportan perfiles masculinos específicos de ADN, lo cual soluciona aquellos casos de interpretación en los que las muestras se encuentran mezcladas [1, 6]. Por eso, el análisis de los Y-STR es una técnica prometedora para los estudios forenses que se ocupan del parentesco paterno, tales como la identificación de víctimas en los desastres masivos [1, 6].

Además, los marcadores Y-STR son muy útiles para las pruebas de paternidad en las que estudia la descendencia masculina [52, 60]. De hecho, la capacidad de los marcadores Y-STR para detectar el parentesco paterno se demostró en 1998 cuando se vinculó a los descendientes actuales de Thomas Jefferson y Eston Hemmings [61]. La prestigiosa revista científica *Nature* publicó un informe titulado como “Jefferson es el padre del hijo menor de su esclava”. Este estudio se basó en los marcadores de ADN del cromosoma Y con el fin de detectar la línea de parentesco masculina de Jefferson en uno de los descendientes del hijo menor de Sally Hemming, cuyo nombre es Eston Hemmings.

El Cuadro 3 presenta un resumen sobre los diez Y-STR que más se han estudiado en genética forense [55]. Estos marcadores difieren entre sí en cuanto a el poder de discriminación y la capacidad para distinguir dos individuos masculinos no emparentados. Además, en la Internet hay acceso a una base de datos que contiene información sobre diversos marcadores de Y-STR [62].

Polimorfismos nucleótidos simples (SNP por sus siglas en inglés)

Entre distintas personas, la variación de secuencias en un único sitio del genoma se suele denominar como polimorfismos nucleótidos simples. Venter y otros determinaron que existe un millón o más de SNP por cada individuo [3]. Dado que el análisis de diversos SNP permite crear un perfil único para cada persona, el estudio de los SNP tiene el potencial de convertirse en una herramienta vital para la discriminación de individuos en el futuro y es de interés particular para la comunidad forense que analiza el ADN.

En primer lugar, según Li y otros, puesto que la interpretación de resultados es más fácil, los SNP han demostrado la capacidad de diferenciarse más fácilmente respecto a los STR, tanto en las muestras que se encuentran entremezcladas como en aquellas que proceden de una única fuente [63]. En segundo lugar, diversas tecnologías están progresando para miniaturizar y automatizar el procedimiento de análisis de los SNP [1]. Por ejemplo, se ha descrito que los ensayos de microchips permiten examinar más de mil SNP en forma simultánea [64]. Por lo tanto, el análisis múltiple de los SNP se podría realizar a mayor escala respecto a los STR. En tercer lugar, el procesamiento de las muestras y el análisis de datos se automatizaría en forma más completa dado que la separación por tamaños no es necesaria [63, 65]. Por último, los marcadores SNP de PCR son más resistentes a la degradación respecto a los amplicones de STR debido a que son de longitud más corta [1]. Sin embargo, cada locus de SNP contiene sólo dos alelos posibles, y se requieren más marcadores para aumentar el grado de certidumbre en el proceso de discriminación [31]. Por eso, pese a que los SNP son más comunes en el genoma humano, su falta

de polimorfismo crea la necesidad de caracterizarlos a gran escala para lograr una probabilidad de coincidencia aleatoria equivalente a los trece loci centrales de STR [1, 31]. Se han identificado diversos marcadores de polimorfismos nucleótidos simples (SNP) del cromosoma Y, que junto con los marcadores Y-STR, se están investigando actualmente para ser empleados en las ciencias forenses [66]. Además, los SNP pueden caracterizar a una población específica dado que presentan una baja tasa de mutación [67, 68, 69]. Por lo tanto, con el análisis de pocos loci de los SNP específicos de una población, en el futuro podría predecirse el origen étnico de los que cometen crímenes. El Cuadro 4 compara los marcadores SNP y STR.

En noviembre del 2001, el centro de información biotecnológica *National Center for Biotechnology Information* (NCBI) divulgó una base de datos de referencia para el SNP humano conocida como dbSNP, la cual se puede consultar en la Internet por medio de la página <<http://www.ncbi.nlm.nih.gov/SNP/>>. Además, la base de datos *The Human Gene Mutation Database* (HGMD) contiene información detallada sobre los SNP y se encuentra disponible para la comunidad científica en la dirección de Internet <<http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html>>.

CONCLUSIÓN

Existen diversas técnicas innovadoras en el horizonte de los estudios de genética humana, que serían útiles para que el científico forense especializado en ADN pueda realizar pruebas de identidad humana. Sin embargo, según el informe *Report of National Commission on the Future of DNA Evidence*, en el futuro aumentará la utilización de los marcadores de ADNmt, el cromosoma Y, el SNP y el STR [70].

Algunas perspectivas indican que el análisis de los STR por electroforesis seguirá siendo de momento el método de escogencia para la comunidad forense, aunque los procedimientos podrían variar con la incorporación de microdispositivos más compactos y portátiles [1, 6]. El futuro del análisis forense de ADN humano se inclina hacia un aumento en el número de muestras procesadas dado que los investigadores forenses están llegando a un punto en el que revisar, interpretar y realizar informes de pruebas de ADN les toma más tiempo que la prueba de laboratorio propiamente dicha [1].

Desde que el análisis del ADN apareció en el campo de las ciencias forenses, las pruebas de ADN han resultado ser una herramienta confiable e invaluable para descifrar información en las investigaciones sobre crímenes. Los laboratorios forenses de ADN investigan constantemente nuevas técnicas para la identificación humana. Sin embargo, la invención de nuevos métodos representa sólo la mitad del esfuerzo. Una vez que se ha comprobado que una nueva técnica es fiable desde el punto de vista científico y suele aceptarse en la comunidad de expertos, debe ser admitida por los tribunales de justicia [38, 71]. Hasta hace poco, las pruebas de ADN humano habían sido el objeto de estudio principal, pero recientemente las pruebas del ADN de plantas, animales y virus ya llegaron por primera vez a los tribunales [72, 73, 74]. Es claro que los laboratorios forenses de ADN se proponen aprovechar al máximo el potencial de esta molécula; sin embargo, el análisis del ADN no humano ha sido desatendido.

Cuadro 1. Resumen de los trece loci STR básicos de la CODIS [6, 30, 31].

Nombre del locus	Motivo de repetición	Acceso ^a en la GenBank	Alelo en la GenBank	Rango de los alelos ^b	Número de alelos vistos
CSF1PO	TAGA	X14720	12	6-16	15
FGA	CTTT	M64982	21	15-51,2	69
TH01	TCAT	D00269	9	3-14	20
TPOX	GAAT	M68651	11	6-13	10
VWA	[TCTG][TCTA]	M25858	18	10-24	28
D3S1358	[TCTG][TCTA]	-	-	9-20	20
D5S818	AGAT	G08446	11	7-16	10
D7S820	GATA	G08616	12	6-15	22
D8S1179	[TCTA][TCTG]	G08710	12	8-19	13
D13S317	TATC	G09017	13	5-15	14
D16S539	GATA	G07925	11	5-15	10
D18S51	AGAA	L18333	13	7-27	43
D21S11	[TCTA][TCTG]	AP000433	29	24-38	70

^a La información sobre las secuencias para un locus STR particular en la GenBank se puede acceder en la Internet en <http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank> con sólo ingresar el número de acceso que se muestra.

^b Los números de esta columna indican el número de unidades de repetición presentes en los alelos.

Cuadro 2. Multiplexes de STR disponibles en el comercio (etiquetados con fluorescencia) [9, 33, 34].

Nombre	Fuente	Loci STR que se incluye	Poder de discrimination
TH01, TPOX, CSF1PO monoplexes	<i>Promega</i>	TH01, TPOX, CSF1PO	1:410
AmpFISTR Blue	<i>PE Applied Biosystems</i>	D3S1358, VWA, FGA	1:5000
AmpFISTR Green I	<i>PE Applied Biosystems</i>	Amelogenina, TH01, TPOX, CSF1PO	1:410
CTTv	<i>Promega</i>	TH01, TPOX, CSF1PO, VWA (vWF)	1:6600
FFFL	<i>Promega</i>	F13A1, FES/FPS, F13B, LPL	1:1500
GammaSTR	<i>Promega</i>	D16S539, D13S317, D7S820, D5S818	1:1,8x104
PowerPlex (versión 1.1 & 1.2)	<i>Promega</i>	TH01, TPOX, CSF1PO, VWA, D16S539, D13S317, D7S820, D5S818	1:1,2x108
AmpFISTR Profiler	<i>PE Applied Biosystems</i>	Amelogenina, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820, D3S1358, FGA, VWA	1:3,6x109
AmpFISTR Profiler Plus	<i>PE Applied Biosystems</i>	Amelogenina, D5S818, D13S317, D7S820, D3S1358, FGA, VWA, D8S1179, D21S11, D18S51	1:9,6x1010
AmpFISTR COfiler	<i>PE Applied Biosystems</i>	D3S1358, D16S539, Amelogenina, TH01, TPOX, CSF1PO, D7S820	1:8,4x105
AmpFISTR SGM Plus	<i>PE Applied Biosystems</i>	D3S1358, VWA, D16S539, D2S1338, Amelogenina, D8S1179, D21S11, D18S51, D19S433, TH01, FGA	1:3,3x1012
PowerPlex 2.1	<i>Promega</i>	D3S1358, TH01, D21S11, D18S51, VWA, D8S1179, TPOX, FGA, Penta E	1:8,5x1010
PowerPlex 16	<i>Promega</i>	CSF1PO, FGA, TPOX, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, Amelogenina	1:1,8x1017

Cuadro 3. Resumen de los diez Y-STR más estudiados en el campo forense [55].

Nombre del marcador	Acceso en la GenBank	Motivo de repetición	Rango de los alelos	Tamaños de los productos de PCR
DYS19	X77751	TAGA	8-16	178-210 pb
DYS385	Z93950	GAAA	10-22	252-300 pb
DYS388	G09695	ATT	12-17	128-143 pb
DYS389 I	G09600	(TCTG) (TCTA)	I: 7-13	239-263 pb
DYS389 II	G09600	(TCTG) (TCTA) (TCTG) (TCTA)	II: 23-31	353-385 pb
DYS390	G09611	(TCTG) (TCTA) (TCTG) (TCTA)	18-27	191-227 pb
DYS391	G09613	TCTA	8-13	275-295 pb
DYS392	G09867	TAT	7-16	236-263 pb
DYS393	G09601	AGAT	9-15	108-132 pb
YCAIII	AC006370	CA	19-25	192-204 pb

Cuadro 4. Comparación de los marcadores de STR y SNP

Características	STR	SNP
Frecuencia en el genoma humano	~1 en cada 10 kb	~1 por cada 1 kb
Capacidad general de proporcionar información	Alta	Baja; sólo un 20-30% proporciona la misma cantidad de información que los STR
Tipo de marcador	Marcadores de repeticiones de di-, tri- y tetranucleótidos	Marcadores bialélicos
Número de alelos por marcador	Por lo general más de 5	Por lo general 2
Métodos de detección	Gel/electroforesis capilar	Análisis de secuencias: hibridación con <i>microchip</i>
Capacidad múltiple	>10 marcadores con múltiples tintes fluorescentes	Un potencial de 1000s en un <i>microchip</i>

REFERENCIAS

- [1.] V.W. Weedn, *Deoxyribonucleic acid: Future analytical techniques.*, in *Encyclopedia of Forensic Sciences*. Academic Press. 2000 pp 490-499.
- [2.] B.S. Shastri, SNP alleles in human disease and evolution. *J. Hum. Genet.* 47 (2002) 561-566.
- [3.] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, et al., The sequence of the human genome. *Science* 291 (2001) 1304-1351.
- [4.] J.G. Taylor, E.H. Choi, C.B. Foster and S.J. Chanock, Using genetic variation to study human disease. *Trends Mol. Med.* 7 (2001) 507-512.
- [5.] D.J. Reeder, Impact of DNA typing on standards and practice in the forensic community. *Arch. Pathol. Lab. Med.* 123 (1999) 1063-1065.
- [6.] L. Carey and L. Mitnik, Trends in DNA forensic analysis. *Electrophoresis* 23 (2002) 1386-1397.
- [7.] N. Hardman, Structure and function of repetitive DNA in eukaryotes. *Biochem. J.* 234 (1986) 1-11.
- [8.] D. Tautz and M. Renz, Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* 12 (1984) 4127-4138.
- [9.] J.M. Butler, *Forensic DNA typing: biology and technology behind STR markers*: Academic Press. 2001.
- [10.] D. Tautz, Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17 (1989) 6463-6471.
- [11.] A.J. Jeffreys, V. Wilson and S.L. Thein, Hypervariable 'minisatellite' regions in human DNA. *Nature* 314 (1985) 67-73.
- [12.] A. Raina and T.D. Dogra, Application of DNA fingerprinting in medicolegal practice. *J. Indian Med. Assoc.* 100 (2002) 688-694.
- [13.] P.M. Schneider, S. Meuser, W. Waiyawuth, Y. Seo and C. Rittner, Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. *Forensic Sci. Int.* 97 (1998) 61-70.
- [14.] AABB, *Facts about blood and blood banking*. 2003, <http://www.aabb.org/>.
- [15.] K.L. Richie, M.D. Goldsborough, M.M. Darfler, E.A. Benzinger, M.L. Lovekamp, D.J. Reeder, et al., Long PCR for VNTR analysis. *J. Forensic Sci.* 44 (1999) 1176-1185.
- [16.] H.A. Hammond, L. Jin, Y. Zhong, C.T. Caskey and R. Chakraborty, Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am. J. Hum. Genet.* 55 (1994) 175-189.
- [17.] T.J. Hudson, Joe Doupe Young Investigators Award. The Human Genome Project: tools for the identification of disease genes. *Clin. Invest. Med.* 21 (1998) 267-276.
- [18.] C.M. Ruitberg, D.J. Reeder and J.M. Butler, STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res.* 29 (2001) 320-322.
- [19.] P. Gill, C.P. Kimpton, A. Urquhart, N. Oldroyd, E.S. Millican, S.K. Watson, et al., Automated short tandem repeat (STR) analysis in forensic casework--a strategy for the future. *Electrophoresis* 16 (1995) 1543-1552.

- [20.] C.P. Kimpton, P. Gill, A. Walton, A. Urquhart, E.S. Millican and M. Adams, Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Meth. Appl.* 3 (1993) 12-22.
- [21.] M. Litt, X. Hauge and V. Sharma, Shadow bands seen when typing polymorphic dinucleotide repeats: some causes and cures. *BioTechniques* 15 (1993) 280-284.
- [22.] A. Urquhart, N.J. Oldroyd, C.P. Kimpton and P. Gill, Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. *BioTechniques* 18 (1995) 116-121.
- [23.] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich, Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51 (1986) 263-273.
- [24.] E. Buel, M.B. Schwartz and M.J. LaFountain, Capillary electrophoresis STR analysis: comparison to gel-based systems. *J Forensic Sci* 43 (1998) 164-170.
- [25.] J. Fregeau and R.M. Fourney, DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification. *BioTechniques* 15 (1993) 100-119.
- [26.] J.F. Anderson, M.J. Greenhalgh, H.R. Butler, S.R. Kilpatrick, R.C. Piercy, K.A. Way, et al., Further validation of a multiplexing STR system for use in routine forensic identity testing. *Forensic Sci. Int.* 78 (1996) 47-64.
- [27.] P. Gill, Role of short tandem repeat DNA in forensic casework in the UK- past, present and future perspectives. *BioTechniques* 32 (2002) 366-385.
- [28.] C.A. Crouse, S. Rogers, E. Amriott, S. Gibson and A. Masibay, Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. *J. Forensic Sci.* 44 (1999) 87-94.
- [29.] A.M. Lins, C.J. Sprecher, C. Puers and J.W. Schumm, Multiplex sets for the amplification of polymorphic short tandem repeat loci--silver stain and fluorescence detection. *Biotechniques* 20 (1996) 882-889.
- [30.] B. Budowle and R.C. Allen, Analysis of amplified fragment-length polymorphisms (VNTR/STR loci) for human identity testing. *Methods Mol. Biol.* 98 (1998) 155-171.
- [31.] R. Chakraborty, D.N. Stivers, B. Su, Y. Zhong and B. Budowle, The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis* 20 (1999) 1682-1696.
- [32.] B. Budowle, B. Shea, S. Niezgodna and R. Chakraborty, CODIS STR loci data from 41 sample populations. *J. Forensic Sci.* 46 (2001) 453-489.
- [33.] <http://www.appliedbiosystems.com/products>.
- [34.] www.promega.com/geneticidentity/.
- [35.] S. Anderson, A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, et al., Sequence and organization of the human mitochondrial genome. *Nature* 290 (1981) 457-465.
- [36.] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull and N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23 (1999) 147.

- [37.] M.M. Holland, D.L. Fisher, L.G. Mitchell, W.C. Rodriguez, J.J. Canik, C.R. Merrill, et al., Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War. *J. Forensic Sci.* 38 (1993) 542-553.
- [38.] E. Marchi and R.J. Pasacreta, Capillary electrophoresis in court: the landmark decision of the People of Tennessee versus Ware. *J. Capillary Electrophor* 4 (1997) 145-156.
- [39.] J.M. Butler and B.C. Levin, Forensic applications of mitochondrial DNA. *Trends Biotechnol.* 16 (1998) 158-162.
- [40.] M.M. Holland and T.J. Parsons, Mitochondrial DNA sequence analysis - validation and use for forensic casework. *Forensic Sci. Rev.* 11 (1999) 21-50.
- [41.] P. Gill, P.L. Ivanov, C. Kimpton, R. Piercy, N. Benson, G. Tully, et al., Identification of the remains of the Romanov family by DNA analysis. *Nat. Genet.* 6 (1994) 130-135.
- [42.] M.N. Gabriel, E.F. Huffine, J.H. Ryan, M.M. Holland and T.J. Parsons, Improved mtDNA sequence analysis of forensic remains using a "mini-primer set" amplification strategy. *J. Forensic Sci.* 46 (2001) 247-253.
- [43.] K.M. Sullivan, R. Hopgood and P. Gill, Identification of human remains by amplification and automated sequencing of mitochondrial DNA. *Int. J. Legal Med.* 105 (1992) 83-86.
- [44.] K. Bender, P.M. Schneider and C. Rittner, Application of mtDNA sequence analysis in forensic casework for the identification of human remains. *Forensic Sci. Int.* 113 (2000) 103-107.
- [45.] K.E. Bendall, V.A. Macaulay and B.C. Sykes, Variable levels of a heteroplasmic point mutation in individual hair roots. *Am. J. Hum. Genet.* 61 (1997) 1303-1308.
- [46.] K. Sekiguchi, K. Kasai and B.C. Levin, Inter- and intragenerational transmission of a human mitochondrial DNA heteroplasmy among 13 maternally-related individuals and differences between and within tissues in two family members. *Mitochondrion* 2 (2003) 401-414.
- [47.] B. Budowle, M.R. Wilson, J.A. DiZinno, C. Stauffer, M.A. Fasano, M.M. Holland, et al., Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci. Int.* 103 (1999) 23-35.
- [48.] B.C. Levin, H. Cheng, M.C. Kline, J.W. Redman and K.L. Richie, A review of the DNA standard reference materials developed by the National Institute of Standards and Technology. *Fresenius J. Anal. Chem.* 370 (2001) 213-219.
- [49.] B.C. Levin, H. Cheng and D.J. Reeder, A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics* 55 (1999) 135-146.
- [50.] M.R. Wilson, M. Stoneking, M.M. Holland, J.A. DiZinno and B. Budowle, Forensic analysis of mitochondrial DNA: application of multiplex solid-phase - fluorescent minisequencing to high throughput analysis. *Crime Lab. Digest.* 20 (1993) 68-77.
- [51.] T. Grzybowski, Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots. *Electrophoresis* 21 (2000) 548-553.

- [52.] M. Kayser, A. Caglià, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, et al., Evaluation of Y chromosomal STRs: a multicenter study. *Int. J. Legal Med.* 110 (1997) 125-133.
- [53.] A. Perez-Lezaun, F. Calafell, M. Seielstad, E. Mateu, D. Comas, E. Bosch, et al., Population genetics of Y-chromosome short tandem repeats in humans. *J. Mol. Evol.* 45 (1997) 265-270.
- [54.] A.J. Redd, S.L. Clifford and M. Stoneking, Multiplex DNA typing of short-tandem-repeat loci on the Y chromosome. *Biol. Chem.* 378 (1997) 923-927.
- [55.] L. Gusmao, A. Gonzalez-Neira, C. Pestoni, M. Brion, M.V. Lareu and A. Carracedo, Robustness of the Y STRs DYS19, DYS389 I and II, DYS390 and DYS393: optimization of a PCR pentaplex. *Forensic Sci. Int.* 106 (1999) 163-172.
- [56.] M. Prinz, K. Boll, H. Baum and B. Shaler, Multiplexing of Y chromosome specific STRs and performance of mixed samples. *Forensic Sci. Int.* 85 (1997) 209–218.
- [57.] E.K. Hanson and J. Ballantyne, A highly discriminating 21 locus Y-STR "megaplex" system designed to augment the minimal haplotype loci for forensic casework. *J. Forensic Sci.* 49 (2004) 40-51.
- [58.] E.K. Hanson and J. Ballantyne, Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications. *Leg. Med.(Tokyo)* (2005) 1-11.
- [59.] D. Corach, L.F. Risso, M. Marino, G. Penacino and A. Sala, Routine Y-STR typing in forensic casework. *Forensic Sci. Int.* 118 (2000) 131-135.
- [60.] L. Roewer, M. Kayser, P. de Knijff, K. Anslinger, D. Corach, S. Füredi, et al., A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Sci. Int.* 114 (2000) 31-43.
- [61.] E.A. Foster, M.A. Jobling, P.G. Taylor, P. Donnelly, P. de Knijff, R. Mieremet, et al., Jefferson fathered slave's last child. *Nature* 396 (1998) 27-28.
- [62.] S. Willuweit, L. Roewer, M. Krawczak, M. Kayser and P. de Knijff, *Y-STR haplotype reference database*. 2003, <http://www.ystr.org>.
- [63.] J. Li, J.M. Butler, Y. Tan, H. Lin, S. Royer, L. Ohler, et al., Single nucleotide polymorphism determination using primer extension and time-of-flight mass spectrometry. *Electrophoresis* 20 (1999) 1258-1265.
- [64.] D.G. Wang, J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, et al., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280 (1998) 1077-1082.
- [65.] A. Linacre and D. Graham, Role of molecular diagnostics in forensic science. *Expert Rev. Mol. Diagn.* 2 (2002) 346-353.
- [66.] P.A. Underhill, L. Jin, A.A. Lin, S.Q. Mehdi, T. Jenkins, D. Vollrath, et al., Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* 7 (1997) 996-1005.
- [67.] M.J. Daly, J.D. Rioux, S.F. Schaffner, T.J. Hudson and E.S. Lander, High-resolution haplotype structure in the human genome. *Nat. Genet.* 29 (2001) 229-232.
- [68.] L.B. Jorde, W.S. Watkins and M.J. Bamshad, Population genomics: a bridge from evolutionary history to genetic medicine. *Hum. Mol. Genet.* 10 (2001) 2199-2207.
- [69.] R. Lewis, SNPs as windows on evolution. *The Scientist* 16 (2002) 16.

- [70.] Report, *National Institute of Justice Report: The future of forensic DNA testing*. 2000, National Commission on the Future of DNA Evidence: Washington, D.C.
- [71.] J.A. Auxier and H.M. Prichard, The role of the expert witness: an update. *Health Phys.* 81 (2001) 269-271.
- [72.] M. Menott-Haymond, Pet cat hair implicates murder suspect. *Nature* 386 (1997) 774.
- [73.] G. Sensabaugh and D.H. Kaye, Non-human DNA evidence. *Jurimetrics J* 38 (1998) 1-16.
- [74.] C.K. Yoon, Forensic science - Botanical witness for the prosecution. *Science* 260 (1993) 894-895.

**ENSAYO DÚPLEX DE qPCR EN TIEMPO REAL PARA CUANTIFICAR EL ADN
NUCLEAR Y MITOCONDRIAL HUMANO EN MUESTRAS FORENSES:
IMPLICACIONES PARA LA CUANTIFICACIÓN DEL ADN
EN MUESTRAS DEGRADADAS***

Mark D. Timken,² Ph.D.; Katie L. Swango,¹ Ph.D.; Cristián Orrego,¹ Ph.D.; and Martin R. Buoncristiani,¹ M.P.H.

RESUMEN: Se desarrolló un ensayo dúplex de reacción en cadena de la polimerasa cuantitativa (qPCR por sus siglas en inglés correspondientes a *quantitative Chain Reaction Polymerase*) en tiempo real para cuantificar el ADN nuclear y mitocondrial humano en muestras forenses. La porción nuclear del ensayo utilizó la amplificación de una secuencia de análisis de ~170-190 pb que abarca la región de repetición de las repeticiones cortas en tandem (STR por sus siglas en inglés que corresponden a *Short Tandem Repeats*) del locus TH01. La porción mitocondrial del ensayo utilizó la amplificación de una secuencia de análisis de 69 pb en la región del ND1. Los estudios de validación llevados a cabo en un equipo ABI 7000 SDS que utiliza detección TaqMan[®] demostraron que ambas porciones del ensayo dúplex suministran una adecuada sensibilidad de cuantificación y una precisión menor a la de 10-15 copias por cada genoma de interés y ninguna porción muestra reactividad cruzada

* Este trabajo en parte fue apoyado por el número de identificación de financiamiento 2002-IJ-CX-K008 que otorgó el Instituto de Justicia: *National Institute of Justice, Office of Justice Programs, US DOJ*. Los puntos de vista que se exponen corresponden a los de los autores y no necesariamente representan la posición o las políticas oficiales del Departamento de Justicia de los Estados Unidos: *US Department of Justice*.

Revisado el 2 de octubre del 2004; presentado con las revisiones el 30 de marzo del 2005; aceptado el 9 abril del 2005; publicado el 3 de agosto del 2005.

² *California Department of Justice Jan Bashinski DNA Laboratory, 1001W. Cutting Blvd., Suite 110, Richmond, CA 94804.*

hacia los genomas no humanos que habitualmente se encuentran. Como parte de los estudios de validación, se cuantificó un conjunto de muestras cuya DNasa se encontraba degradada. Para ello se utilizaron tres métodos distintos: el ensayo dúplex de qPCR nuclear y mitocondrial, el ensayo de qPCR con el paquete comercial de marcadores genéticos ABI Quantifiler™ para la cuantificación del ADN humano, el cual amplifica y detecta una secuencia nuclear de 62 pb, y finalmente, el ensayo de hibridación *slot-blot*. Para el conjunto de muestras que no están degradadas o que presentan un grado moderado de degradación, los tres métodos muestran una precisión adecuada para cuantificar el ADN nuclear y lograr amplificaciones de STR satisfactorias para obtener perfiles completos, mediante el paquete comercial de marcadores genéticos ABI AmpF/STR® Identifiler™. Sin embargo, para las muestras muy degradadas, el ensayo dúplex de qPCR proporciona mejores estimaciones del molde o la plantilla nuclear para la amplificación de STR en comparación con el ensayo comercial de qPCR o la hibridación *slot-blot*. Por un lado, el ensayo comercial de qPCR sobrestima la cantidad de los fragmentos de ADN de longitud STR, lo cual aumenta la proporción de alelos no detectados en los loci más largos de STR. Por otro lado, la hibridación *slot-blot* desestima la cantidad de ADN nuclear al amplificar un número excesivo de moldes o plantillas, lo que aumenta la proporción de artefactos de amplificación de STR.

Descriptor: ciencias forenses, cuantificación de ADN, reacción en cadena de la polimerasa cuantitativa, ADN degradado, TH01, ND1, transición de fusión.

La necesidad de aumentar la cantidad de ensayos de genotipificación forense de ADN ha despertado creciente interés para la creación de nuevos métodos de cuantificación de ADN que sean más eficientes que la hibridación *slot-blot*, que actualmente se utiliza en muchos laboratorios de análisis forense de ADN (1). A pesar de que el método de hibridación *slot-blot* ofrece una excelente especificidad y buena sensibilidad (2, 3), supone un protocolo de larga duración y labor intensiva que aún no está listo para automatizarse. Ante esta situación, numerosas publicaciones recientes han descrito enfoques alternativos para cuantificar el ADN en muestras forenses, entre los cuales están el ensayo de hibridación líquida (4, 5) y diversos ensayos de PCR de punto final (6-8). Otro método que está resultando útil para la cuantificación de ADN forense es el ensayo de PCR cuantitativa en tiempo real (qPCR), el cual se utiliza extensivamente en la investigación biomédica y en el diagnóstico molecular (9-13). Los ensayos cuantitativos de PCR se han elaborado para diversas aplicaciones forenses como la cuantificación del ADN nuclear (14-16), mitocondrial (17-19) y del cromosoma Y humano (16, 19). Estos ensayos se valen al máximo de las propiedades de la qPCR, lo que permite que esta técnica sea de peculiar atracción para las aplicaciones forenses: i. los ensayos de qPCR pueden diseñarse para cuantificar genomas específicos de interés; ii. los ensayos presentan una sensibilidad adecuada para detectar tan sólo unas pocas copias e incluso una sola copia del ADN que se analiza; iii. los rangos de detección dinámica de la qPCR abarcan los tres órdenes de magnitud ordinarios (por ejemplo, de 30 pg a 30 ng de ADN nuclear) que se necesitan en la mayoría de las aplicaciones forenses; iv. además, los protocolos experimentales para la cuantificación de qPCR en tiempo real son sencillos, prácticos y aptos para

automatizarse. También, el empleo de sondas de detección de sustancias específicas (por ejemplo, TaqMan[®] (20) o las sondas moleculares de Beacon (21)) permiten que se realicen ensayos múltiple de qPCR en tiempo real que pueden cuantificar en forma simultánea más de una unidad por muestra, lo cual ofrece la posibilidad de ahorrar tiempo, trabajo y ADN extraído.

En este informe describimos un ensayo dúplex de qPCR en tiempo real para cuantificar simultáneamente el ADN nuclear y mitocondrial humano en muestras forenses. Este fue elaborado para emplearse en forma general en las cuantificaciones forenses de ADN, pero en especial para el análisis posterior a la extracción de muestras que contienen ADN muy degradado. Aunque éstas no dejan de ser usuales en los casos habituales, se encuentran con mayor frecuencia en sucesos de desastres masivos, sepulturas masivas y casos de personas desaparecidas (22, 23). Debido al alto poder de discriminación para identificar a los humanos, la genotipificación de las repeticiones cortas en tandem (STR) suele ser el método analítico que más se utiliza. Sin embargo, la calidad o la cantidad de ADN nuclear extraído de las muestras degradadas impide a menudo la genotipificación satisfactoria de STR, y como consecuencia, se obtienen perfiles parciales o nulos de STR. Por lo tanto, tales muestras se pueden analizar mediante métodos menos discriminantes, los cuales se basan en los polimorfismos de las subregiones hipervariables I y II (HVI y HVII) de la región de control mitocondrial humana. Actualmente, para decidir en un inicio si el procedimiento de análisis se realizará por medio del STR nuclear o el mitocondrial, los analistas se suelen fundamentar en un enfoque cuantitativo dado por la hibridación *slot-blot*, el cual: i. se ha encontrado que desestima la cantidad del ADN nuclear en las

muestras degradadas (24) o en aquellas que contienen mucha contaminación microbiana (23); ii. no ofrece información sobre la calidad (longitud de los fragmentos) del ADN nuclear que se cuantifica; iii. no permite conocer directamente la cantidad de ADN mitocondrial humano presente en la muestra; y iv. no brinda información que permita prever la presencia de inhibidores que interfieran en la PCR. Ante tales deficiencias de cuantificación, los análisis forenses de muestras complicadas a menudo comienzan por obtener resultados inadecuados sobre los STR, y luego continúan con la utilización de cualquier residuo de ADN extraído para intentar el análisis mitocondrial. Anteriormente se mencionó (17) que la eficiencia y calidad del procedimiento de este análisis mejoraría notablemente si antes de comenzar se obtuvieran estimaciones confiables sobre las cantidades de ADN nuclear y mitocondrial humano en las muestras. Con base en esas estimaciones, desde el principio se seleccionaría un enfoque analítico óptimo, lo cual produciría resultados de genotipificación o haplotipificación recomendables junto con un ahorro de tiempo, trabajo, costos de reactivos y de marcadores genéticos, así como de ADN extraído.

Diversas publicaciones recientes han tratado con exactitud estos temas aplicados a las muestras forenses, y han hecho hincapié en las muestras degradadas. Por ejemplo, von Wurmb-Schwark y otros (25) desarrollaron un ensayo dúplex de PCR de punto final para detectar unidades nucleares (amplicón de 164 pb en el gen beta globina) y unidades mitocondriales (260 pb en el ND1). En dicho ensayo, los productos de la PCR de punto final se resolvieron y revelaron mediante la electroforesis capilar con detección por fluorescencia, y aunque los resultados no resultaron ser cuantitativos, la intensidad de las señales nucleares y mitocondriales demostraron que

este método es una herramienta útil y adecuada para el análisis forense de muestras óseas antropológicas. Alonso y otros (19, 26) elaboraron diversos ensayos de qPCR para muestras degradadas y que presentan complicaciones, tal como el ensayo dúplex que cuantifica los cromosomas X y Y humanos en forma simultánea (unidades de 106 pb y 112 pb en el gen de amelogenina). También llevaron a cabo dos ensayos qPCR uniplex para cuantificar el genoma mitocondrial humano, el primero para un HVI de 113 pb y el segundo para un HVI de 287 pb. Dada la diferencia de tamaño entre las dos secuencias, se demostró que si se comparan los resultados de cuantificación de estos ensayos de qPCR mitocondriales se podría obtener información sobre el nivel de degradación del ADN. Recientemente, Adreasson y otros (17) describieron un ensayo dúplex de qPCR para ADN nuclear y mitocondrial. La porción nuclear de este ensayo cuantifica una unidad de 79 pb en el gen retinoblastoma (*RB1* por su abreviatura en inglés), y la porción mitocondrial cuantifica una unidad de 135 pb que abarca la unión de los genes *tRNA^{lys}* y *ATP8*. Los resultados demostraron que el enfoque dúplex de qPCR hace posible la realización cuantificaciones de ADN útiles y sensibles, y a la vez permite que el analista ahorre tiempo y cantidad de ADN, la cual muchas veces es escasa. Sin embargo, se ha notado (1) que no es conveniente escoger la secuencia del gen *RB1* para el análisis de qPCR debido a que algunas investigaciones sobre la homología de secuencias entre especies indican que la secuencia nuclear del *RB1* se conserva relativamente. Por lo tanto, aunque se espera que este ensayo brinde resultados precisos respecto a las muestras forenses de origen humano conocido, no se ha demostrado que sea completamente específico para los primates, y que por ende pueda utilizarse en forma generalizada en el campo forense.

En el presente estudio describimos un nuevo ensayo dúplex de qPCR para ADN nuclear y mitocondrial que se llevó a cabo según el mismo enfoque general de Andreasson y otros, (17), salvo que escogimos secuencias de análisis nuclear y mitocondrial alternativas para realizar la amplificación y la cuantificación. Para la porción nuclear del ensayo de qPCR escogimos una secuencia de análisis que abarca la región de repetición STR del locus TH01, el cual es exclusivo de los primates y se ha utilizado extensamente para las aplicaciones forenses. Esta secuencia de análisis es de interés particular para la cuantificación si se considera que la razón principal para cuantificar el ADN nuclear humano en las muestras forenses es determinar la cantidad que se debe extraer para realizar la amplificación, mediante la utilización de un paquete comercial de marcadores genéticos de PCR múltiplex de STR (8). Los resultados indican que si se escoge la secuencia relativamente larga del TH01 (~170-190 pb) para los casos que presentan muestras degradadas, esto permite mejorar los resultados de análisis de STR en comparación con los estudios que parten de la cuantificación de secuencias cortas (por ejemplo, 62 pb con base en el paquete de marcadores genéticos de qPCR conocido en inglés como *Applied Biosystems QuantifilerTM* (16)), o en la hibridación *slot-blot*. Para la porción mitocondrial del ensayo, seleccionamos una secuencia de análisis relativamente corta (69 pb) en el gen mitocondrial *ND1*. Esta secuencia representa un método sensible para determinar la presencia de ADN mitocondrial humano en las muestras humanas, ya sean conservadas o en estado de descomposición. Aparte de describir los resultados de la cuantificación de ADN en las muestras degradadas, también incluimos aspectos sobre la elaboración de ensayos dúplex de qPCR y algunos detalles acerca de la validación

de estudios forenses como los de precisión, reproducibilidad, sensibilidad, especificidad de especies y aplicaciones para las muestras de trabajo.

Materiales y métodos

Las muestras y los estándares

En términos generales, los fragmentos extraídos del ADN genómico humano previamente cuantificados y de alto peso molecular que se obtienen del *Promega* (Female-#G1521) y del *Applied Biosystems* (TaqMan[®] Control DNA) se utilizaron como estándares de cuantificación de qPCR. Los fragmentos previamente cuantificados de ADN que se obtuvieron del *Applied Biosystems* (estándares de ADN genómico de los paquetes comerciales de marcadores genéticos Quantifiler[™] y QuantiBlot[™]), *Promega* (K562 y #G1471-Masculino), y el ATCC (HL60) también se utilizaron como muestras de control en distintos estudios.

A no ser que se haya indicado lo contrario, los fragmentos de ADN extraídos como referencia y las muestras simuladas no probatorias se obtuvieron con la utilización del protocolo de extracción orgánica del Departamento de Justicia de California (*California Department of Justice*), el cual incluye la digestión de ProK/SDS, la extracción de fenol/cloroformo, y finalmente la concentración de Centricón 100 (Milliporo) en el *buffer* Tris-EDTA (TE⁻⁴). Para las muestras simuladas de violación sexual se prepararon extractos diferenciales de espermatozoides con fragmentos epiteliales mediante un protocolo de digestión de DDT (27). Para los trozos de elementos pilosos que no procedían de la raíz (fragmentos de 2 cm) se realizaron extracciones mediante i. el protocolo de un agitador magnético (28) ii. el protocolo de

un molidor de tejido que se utilizó primeramente para homogenizar el fragmento de cabello con 150 μL de TE^{-4} , para después verterlo en 50 μL de Chelex 20% (Bio-Rad) hirviendo. Los fragmentos extraídos de ADN se guardaron a -20°C , con excepción de las muestras sacadas del Chelex, las cuales se almacenaron a $+4^{\circ}\text{C}$.

Para los estudios que involucraban degradación, las muestras se fragmentaron mediante el tratamiento con DNasa I (Invitrogen). Los tubos separados que contenían ADN HL60 de alto peso molecular, cada uno con 1,4 μg de ADN en 5 μL de TE^{-4} , se trataron con 0,5 U de DNasa I, un *buffer* de reacción de DNasa I de 10x y agua esterilizada para que la reacción alcanzara el volumen de 10 μL . Se alcanzaron niveles crecientes de fragmentación al permitir que la digestión ocurriera en los tubos a temperatura ambiente y por periodos de 1, 2, 3, 4, 5, 15, 30, 45, 60 min o durante toda la noche. Al final de cada periodo de digestión, la actividad de la DNasa culminó al añadir 2 μL de EDTA de 25 mM y al calentar cada tubo a 65° durante 15 min. Las muestras no se purificaron ni se concentraron más. Una muestra de ADN intacto de “0 min” que contenía todos los componentes de reacción para la digestión menos la DNasa I se trató en forma idéntica a las muestras en la serie de degradación. El grado de fragmentación del ADN se evaluó mediante electroforesis en geles (2% agarosa, detección de bromuro de etidio) al utilizar *HyperLadder* I (Biolina) y marcadores de tamaño *Ready-Load λ -ADN/Hind III* (Invitrogen).

Las muestras de ADN no humano se adquirieron como fragmentos extraídos de ADN genómico previamente cuantificados, a partir de distintos proveedores: de *E.coli*, *C. perfringens* (Sigma), *B. subtilis*, *S. epidermidis*, *C. albicans* (ATCC), de ratón

(Promega) y de gato, pollo, res, perro, pescado, caballo, mono, cerdo, rata (Laboratorios Zyagen Labs, San Diego, CA).

La genotipificación de STR

Para la genotipificación de STR se utilizó el paquete comercial de marcadores genéticos requerido para la amplificación de PCR AmpF/STR® Identifiler™ (Applied Biosystems). Se realizaron las amplificaciones de PCR de 1 ng de ADN nuclear en un volumen de reacción de 25 µL según las instrucciones del fabricante especificadas para un termociclador de PCR de marca GeneAmp® 9700 (Applied Biosystems). Los STR se detectaron y determinaron en un analizador genético conocido como Prism® 3100 Genetic Analyzer (Applied Biosystems) según las instrucciones del fabricante, con la excepción de que en vez de utilizarse los 10 segundos de tiempo recomendados para la inyección electrocinética de la muestra, se emplearon cinco segundos, y los datos se analizaron con una línea de base de 35, y no con una de 51 como predeterminó el fabricante. Para el análisis de datos, los alelos se identificaron con un umbral mínimo de 100 RFU por medio del GeneScan® (v.3.7.1/NT) y el Genotyper® (v.3.7/NT) (Applied Biosystems).

La PCR mitocondrial de HVI/HVII

Las fracciones de HVI y HVII de la región de control del ADN mitocondrial se amplificaron en un ensayo PCR dúplex mediante la utilización de reactivos del paquete comercial de marcadores genéticos de la secuencia de la región mitocondrial ADN

HVI/HVII conocido como LINEAR ARRAY™ (Roche Applied Science) (29, 30). Este paquete de marcadores produce amplicones de 444 pb (HVI) y 416 pb (HVII) nominales. Para cuantificar el genoma mitocondrial, las cantidades de molde o plantilla para la PCR se determinaron por medio de nuestro ensayo dúplex nuTH01 de qPCR (que utiliza 100 pg de ADN nuclear por amplificación dúplex de HVI/HVII), o mediante nuestro ensayo dúplex mtND1 de qPCR (que utiliza ~14 000 copias mitocondriales por amplificación dúplex de HVI/HVII). El rendimiento y la pureza posteriores a la amplificación de los productos HVI y HVII de PCR se evaluaron con electroforesis de geles verticales de poliacrilamida (4% de gel de agarosa NuSieve 3:1 (FMC), tinción de bromuro de etidio) en 5 µL del producto de la PCR mediante un *Low DNA Mass Ladder* (Invitrogen) que actúa como un marcador del tamaño y la cantidad.

La cuantificación por hibridación slot-blot

El paquete comercial de marcadores genéticos para la cuantificación de ADN humano de la marca QuantiBlot™ (*Applied Biosystems*) se utilizó según las instrucciones del fabricante. Las sondas hibridadas se detectaron mediante quimioluminiscencia (*SuperSignal West Femto* (Pierce)) y la utilización de un sistema de cámara CCD (CCDBio 16SC (Hitachi/MiraiBio)). Los datos de la CCD se analizaron de forma semiautomática mediante un programa de computación denominado SlotQuant que se ejecuta mediante el paquete de análisis genético *Gene Tools* (SynGene). Los estándares es de cuantificación (el estándar A de ADN proporcionado

por el paquete comercial de marcadores genéticos) oscila entre 20 ng y 25 pg para cada serie.

La cuantificación de qPCR con Quantifiler™

El paquete comercial de marcadores genéticos para la cuantificación del ADN humano Quantifiler™ (*Applied Biosystems*) que a continuación se denominará “qPCR de Quantifiler™” se utilizó según las instrucciones del fabricante para obtener datos a partir de un equipo de qPCR conocido como *Applied Biosystems 700 Prism®* SDS.

El diseño del ensayo qPCR

Los cebadores y las sondas (Cuadro 1) para los ensayos uniplex de qPCR TaqMan® de nuTH01 y el TaqManMGB® de mtND1 se diseñaron mediante el programa computacional *Applied Biosystems' PrimerExpress v.2.0*. En general, las instrucciones de diseño se siguieron según la recomendación de *Applied Biosystems* (31), aunque los ajustes preestablecidos del programa se ampliaron con frecuencia para permitir que las longitudes del amplicón excedieran la máxima longitud recomendada de 150 pb. Las secuencias del ADN utilizadas para el diseño de trabajo se cargaron a partir del GenBank, disponible en la Internet por medio del centro estadounidense de biotecnología *National Center for Biotechnology Information (NCBI)* (32). La información sobre la secuencia también se obtuvo de los sitios de Internet STRBase (33) y MitoMap (34).

El ensayo nuclear de qPCR (nuTH01) se diseñó para abarcar la secuencia STR en el gen humano tirosina hidroxilasa (TH01) en el cromosoma 11 (11p15.5) mediante

la utilización de la secuencia del locus AF536811 del GenBank. En este trabajo de diseño configuramos el PrimerExpress™ para ajustar las posiciones de los extremos 3' de los cebadores del qPCR de tal manera que coincidieran con cebadores conocidos o deducibles a partir de los paquetes comerciales de amplificación del STR (35, 36); esto con el fin de que las amplificaciones de qPCR pudieran predecir cuantificaciones de la genotipificación del STR en forma más precisa. Para la optimización realizada por el PrimerExpress™, las posiciones de la sonda fluorogénica así como los extremos 5' de los cebadores se trabajaron usualmente sin restricciones. El ensayo qPCR mitocondrial (mtND1) se diseñó para el gen ND1 del genoma mitocondrial humano a partir de la secuencia del locus HUMMTCG del GenBank (37). El gen ND1 presenta la subunidad 1 de la proteína mitocondrial NADH deshidrogenasa.

CUADRO 1—Secuencias de oligonucleótidos para los cebadores y la sonda de nuTH01, y para los cebadores, la sonda y el estándar numérico de copias mitocondriales del mtND1

Oligonucleótido	Secuencia {5' → 3'}
nuTH01-sentido	AGG GTA TCT GGG CTC TGG
nuTH01-antisentido	GGC TGA AAA GCT CCC GAT TAT
nuTH01-sonda	FAM-ATT CCC ATT GGC CTG TTC CTC CCT T-BHQ
mtND1- sentido	CCC TAA AAC CCG CCA CAT CT
mtND1- antisentido	GAG CGA TGG TGA GAG CTA AGG T
mtND1-sonda	VIC-CCA TCA CCC TCT ACA TC-MGB-NFQ
mtND1-estándar	GAG CGA TGG TGA GAG CTA AGG TCG GGG CGG TGA TGT AGA GGG TGA TGG TAG ATG TGG CGG GTT TTA GGG

Tanto para el ensayo nuclear como para el mitocondrial, antes de ordenar los oligonucleótidos sintéticos para el trabajo experimental, se compararon secuencias de sondas y de cebadores potenciales respecto a las secuencias de ADN disponibles por

medio del sitio de Internet del NCB al utilizar las herramientas de búsqueda denominadas en inglés *Basic Local Alignment Search Tool* (BLASTn por sus siglas en inglés) (38). El propósito de estas comparaciones fue identificar y evitar homologías inadvertidas entre los cebadores o sondas respecto a los genomas que no son objeto de estudio, lo cual causaría reactividades cruzadas entre especies. Además, se realizaron pruebas para evitar polimorfismos de nucleótidos simples (SNP por sus siglas en inglés) (*Single Nucleotide Polymorphisms*) en las secuencias de los cebadores o las sondas. La información sobre los SNP se obtuvo a partir de diversas fuentes encontradas en los siguientes sitios de Internet: *NCBI SNPdB* (39), *On-Line Mendelian Inheritance in Man (OMIM)* (40), y para las secuencias mitocondriales: *MitoMap* (34). Una vez que se elaboraron los ensayos uniplex y dúplex de qPCR, éstos se optimizaron según las recomendaciones publicadas (41, 42).

Las cuantificaciones de qPCR para los genes nuTH01 y mtND1

Las secuencias de los cebadores y las sondas para los ensayos de qPCR de nuTH01 y mtND1 aparecen en el Cuadro 1. Para las amplificaciones dúplex de qPCR de nuTH01 y mtND1, cada ensayo se efectuó como una amplificación de 20 μL que incluía 10 μL de *Universal Master Mix 2x* no UNG de la marca TaqMan[®] (*Applied Biosystems*), 4 μL de muestra y el 6 μL restante compuesto para alcanzar las concentraciones finales de: 0,16 $\mu\text{g}/\mu\text{L}$ de BSA no acetilada (Sigma); 600 nM en cada cebador de nuTH01; 200 nM en la sonda de nuTH01; 50 nM en cada cebador de mtND1; 100 nM en la sonda de mtND1. Los cebadores (*Qiagen-Operon* (de Alameda,

CA)) se diluyeron en agua esterilizada y desionizada. Las sondas (*Qiagen-Operon* para el nuTH01, *Applied Biosystems* para el mtND1) se diluyeron en TE⁻⁴.

Para las amplificaciones uniplex de qPCR de TaqMan[®]/ TaqManMGB[®] de los genes nuTH01 y mtND1, los ensayos se llevaron a cabo con los mismos reactivos que se describieron para el ensayo dúplex de qPCR, salvo que se reemplazaron las combinaciones indeseables de cebadores y sondas por el agua esterilizada o el TE⁻⁴. Los ensayos uniplex SYBR[®] Green I por lo general se realizaron como amplificaciones de 20 µL que incluían 10 µL de *Green Master Mix* de SYBR[®] 2x (*Applied Biosystems*) o 10 µL de *QPCR Master Mix* (*Stratagene*) de Brilliant[®] 2x, 4 µL de muestra, y el 6 µL restante compuesto para alcanzar las concentraciones finales de 300nM en cada uno de los cebadores correspondientes.

La lectura de los datos de la qPCR en tiempo real se llevó a cabo en un equipo de *Applied BioSystems Prism[®] 7000 SDS*, el cual se controló mediante la versión 1.0 del programa computacional de 7000 SDS Collection. El equipo se configuró para las siguientes condiciones de trabajo: volúmenes de muestras de 20 µL; modo de emulación 9600; una porción de 10 min para la activación de la polimerasa a 95°, seguida por 45 ciclos de qPCR de dos pasos (15 s de desnaturalización a 95°C y 60 s de hibridación y extensión combinadas a 60°). Las variaciones pronunciadas en la fluorescencia de fondo se corrigieron mediante una referencia pasiva de ROX, la cual se incluye dentro del *Master Mix* de qPCR de *Applied Biosystems* para cada una de las muestras. Con el objetivo de realizar una curva de fusión después del ciclo final de extensión de PCR de las pruebas que utilizan detección de SYBR[®] Green, el

programa computacional *SDS Collection* se configuró con el fin de controlar la fluorescencia durante el aumento de temperatura (a razón de $\sim 1,8^{\circ}\text{C}/\text{min}$) desde 60°C hasta 95°C .

Al analizar las curvas de amplificación se utilizaron ajustes empíricos de las líneas de base y el umbral de los ciclos para cada uno de los tipos de ensayo (para el nuTH01, umbral=0,15; línea de base de 6-18 ciclos), (para el mtND1, umbral= 0,06; línea de base de 3-13 ciclos). Para cada prueba de qPCR, el programa *SDS Collection* realizó una gráfica de calibración lineal de CT (umbral del ciclo) vs. $\log C_0$ (concentración estándar inicial de ADN) al utilizar los resultados de amplificación que se obtuvieron de un conjunto de disoluciones frescas preparadas a partir del estándar precuantificado de ADN genómico humano de alto peso molecular (*Promega Female* o *Applied Biosystems TaqMan*[®]). Las cuantificaciones de ADN para las muestras desconocidas se interpolaron a partir de la curva de calibración lineal obtenida. Estos pasos de calibración e interpolación son atributos semiautomáticos que ofrece el programa *SDS Collection*. Para el ensayo nuclear de qPCR, las gráficas de calibración se construyeron a partir de los datos de las diluciones estándares de ADN, las cuales contienen 25, 5, 1, 0,5, 0,1, y 0,05 (en duplicado) ng de ADN total por cada muestra. Para el ensayo qPCR mitocondrial, las gráficas de calibración se extendieron para incluir menores cantidades de cebadores por muestra (0,010, 0,001, y 0,0001 (en duplicado) ng). En cada una de las pruebas se incluyó al menos un control negativo, $4\mu\text{L}$ de TE^{-4} o agua esterilizada.

Los números de copias de ADN nuclear se estimaron a razón de una copia nuclear haploide por cada 3,3 pg de ADN genómico (43). Los números de las copias de ADN mitocondrial se estimaron a razón de 400 copias mitocondriales por cada 3,3 pg de ADN genómico estándar de *Promega* o 450 copias mitocondriales por cada 3,3 pg de ADN genómico estándar de HL60. Estas proporciones de números de copias mitocondriales se estimaron en forma empírica al pasar las muestras de ADN estándar de HL60 y de *Promega* contra un conjunto de diluciones del estándar oligonucleótido sintético de mtND1 cuantificado (por absorción UV-vis) (ver el Cuadro 1 para la secuencia oligonucleotídica del estándar numérico de copias mitocondriales). Utilizamos series de disoluciones de ADN genómico de alto peso molecular para construir las dos gráficas de calibración lineal correspondientes a la porción nuclear y mitocondrial del ensayo dúplex de cada prueba de qPCR.

En los casos apropiados, la eficiencia de la amplificación qPCR se determinó a partir de la pendiente de las curvas de calibración lineal (% de eficiencia de la PCR = $100[(10^{(-1/\text{pendiente})} - 1)]$ (44).

Cálculos del perfil de fusión de los oligonucleótidos

Los perfiles de fusión para los alelos TH01 seleccionados se calcularon mediante el programa MELT94, el cual funciona con DOS y se encuentra disponible en la Internet (45). Este programa utiliza la teoría y las ecuaciones de Poland (46) así como las de Fixman y Freire (47) para calcular las variaciones de la estabilidad termal en la secuencia de los fragmentos de ADN (48). Se construyeron cantidades considerables de perfiles de fusión al exportar los datos de salida del MELT94

(temperaturas para el 50% del helical: el 50% de los estados de fusión vs. número de secuencia) hacia una hoja de cálculo de Excel[®].

Resultados y discusión

El diseño del ensayo dúplex de qPCR para los genes nuTH01 y mtND1

En la sección titulada 'Materiales y métodos', presentamos un breve bosquejo sobre los procedimientos que se utilizaron para diseñar nuestros ensayos uniplex de qPCR nuclear (nuTH01) y mitocondrial (mtND1). En esta sección se muestra una discusión más detallada de los aspectos seleccionados sobre el diseño y elaboración del ensayo, lo que incluye: i. selección de las secuencias de ADN que se analizan; ii. los resultados experimentales que demuestran que los ensayos uniplex de qPCR para los nuTH01 y mtND1 funcionan satisfactoriamente en las amplificaciones dúplex; y iii. la observación atípica pero predecible de las curvas de fusión de SYBR[®] Green para los TH01 de la qPCR.

Selección y diseño de la unidad de análisis para la amplificación nuclear de TH01

Dado que la razón principal para cuantificar el ADN nuclear en las muestras forenses es determinar la cantidad de cebadores que se deben emplear como materia prima para la genotipificación de STR, anticipamos que la cuantificación de una unidad de STR sería satisfactoria para la genotipificación de STR. Además, debido a su uso generalizado en los análisis forenses de ADN y las investigaciones de paternidad, se han realizado bastantes caracterizaciones de los loci STR en relación con la

reactividad cruzada entre las especies (49, 50) y las tasas de mutación del ADN (51). Aunque al principio de la elaboración del ensayo no nos centramos únicamente en el locus TH01, al examinar las secuencias de ADN en cada loci STR de la base de datos CODIS y al realizar el diseño del ensayo mediante el PrimerExpress™, este locus de inmediato se identificó como una unidad de análisis prometedora. En ambas regiones flanqueantes, el locus TH01 contiene secuencias aptas para el diseño o ubicación de una sonda de detección TaqMan®. La unidad de análisis de STR del TH01 también resultó interesante porque los amplicones tienen una longitud suficiente para ubicarlos de forma aproximada en el rango medio de tamaños que producen los paquetes comerciales de marcadores de STR, y tampoco alcanzan una longitud que comprometa mucho la eficiencia de la qPCR de TaqMan®. Durante la etapa de elaboración, diseñamos y probamos varios ensayos preliminares de qPCR en el STR del locus TH01. La evaluación de estos ensayos se basó en la comparación de las propiedades de qPCR (por ejemplo, la eficiencia de PCR (>90%), la sensibilidad (CT bajo), y precisión (escasa variación en el CT, especialmente cuando hay poca cantidad de cebadores)) con el objetivo de identificar algún ensayo óptimo, el cual aquí se designa como nuTH01. La secuencia que se analiza para el ensayo qPCR de nuTH01 se muestra en la Fig. 1, donde se presentan las posiciones relativas de los cebadores de la PCR, la región de repetición del STR y la sonda fluorogénica FAM/BHQ para el extremo de 5'.

Selección y diseño de la unidad de análisis para la amplificación del ND1 mitocondrial

Para diseñar el ensayo de qPCR mitocondrial seleccionamos una región del gen ND1 (NADH subunidad 1 de deshidrogenasa) que corresponde con las bases 3485-3553 de la secuencia de referencia *Cambridge Reference Sequence* (CRS por sus siglas en inglés) (52). Esta unidad de análisis se escogió a partir de estudios anteriores sobre la homología de secuencias cruzadas entre especies (53, y comunicación personal con Cummings MP), los cuales indican que esta subregión del gen ND1 se conserva muy escasamente. Esta característica es prometedora para la elaboración de un ensayo específico para la especie. Además, la unidad de análisis mencionada se escogió porque esta región del ND1 contiene pocos SNP conocidos, en especial cuando se compara con las subregiones de control de HVI/HVII. Diseñamos y evaluamos diversos ensayos para la unidad de análisis de ND1 mediante herramientas de búsqueda en la Internet y evitamos cualquier SNP conocido por su relación con alguna enfermedad, tal como el caso del SNP del CRS 3460, el cual se asocia al LHON. Con base en el mismo criterio del ensayo de qPCR de nuTH01, seleccionamos un ensayo óptimo para la qPCR mitocondrial, y se denomina mtND1 (Cuadro 1).

```

GGGCAAAATT CAAAGGGTAT CTGGGCTCTG GGGTGATTCC CATTGGCCTG
CCCGTTTTAA GTTTCCCATATA GACCCGAGAC CCCACTAAGG GTAACCGGAC

TTCCTCCCTT ATTTCCCTCA TTCATTCAATT CATTCAATTCA TTCATTCAACC
AAGGAGGGAA TAAAGGGAGT AAGTAAGTAA GTAAGTAAGT AAGTAAGTGG
                                     ↑
ATGGAGTCTG TGTTCCCTGT GACCTGCACT CGGAAGCCCT GTGTACAGGG
TACCTCAGAC ACAAGGGACA CTGGACGTGA GCCTTCGGGA CACATGTCCC

GACTGTGTGG GCCAGGCTGG ATAATCGGGA GCTTTTCAGC CCACAGGAGG
CTGACACACC CGGTCCGACC TATTAGCCCT CGAAAAGTCG GGTGTCCTCC

```

FIG. 1—Información de la secuencia para la porción nuTH01 del ensayo dúplex de qPCR. Se muestran las posiciones relativas de los cebadores con sentido y antisentido (flechas horizontales), la sonda de detección TaqMan® (aparece subrayada en la cadena superior), y la región STR del (CATT)₇ (aparece subrayada en la cadena inferior). La flecha vertical pequeña marca la posición de la secuencia que separa aproximadamente el amplicón nuTH01 en rangos de temperatura de fusión altos y bajos. La posición de esta flecha corresponde a la secuencia número 96 en la Fig. 5.

El ensayo de mtND1 cuantifica una cantidad relativamente corta (69 pb) de la secuencia de análisis, una cadena que aparece en el Cuadro 1 como la secuencia para nuestro estándar numérico de copias de mtND1. Una de las ventajas de seleccionar ensayos con secuencias de análisis cortas fue que permitió la compra de un oligonucleótido sintético que podría usarse como un estándar numérico de copias para el ensayo mitocondrial, en vez de preparar un estándar a partir de la clonación o la purificación del producto de la PCR. También anticipamos que un ensayo corto de qPCR detectaría con más sensibilidad la presencia de ADN mitocondrial humano, aun en muestras de ADN degradado.

Elaboración del ensayo dúplex de qPCR para los genes nuTH01 y mtND1

Con el fin de elaborar un ensayo dúplex de qPCR satisfactorio se requiere identificar condiciones de reacción que permitan realizar efectivamente dos amplificaciones que ocurran en forma independiente en el mismo tubo. El dilema que se pretende evitar es que una de las amplificaciones reduzca la eficiencia de la PCR durante la segunda amplificación. Esta situación produciría que la amplificación se retarde para el segundo ensayo, y además, que se obtengan valores aumentados de CT y se desestime la cantidad del ADN que se mide en el ensayo. Una estrategia para evitar esta situación es elaborar ensayos dúplex de qPCR que trabajen en condiciones que limiten la amplificación de PCR de la secuencia de análisis más abundante. Por ejemplo, que se restrinjan las concentraciones del cebador o de la sonda para esta amplificación (42).

Para el ensayo dúplex de qPCR de los genes nuTH01 y mtND1 realizado en muestras forenses, esperamos que la proporción de copias del genoma mitocondrial respecto a las copias del genoma nuclear haploide excedieran normalmente en 100. Esta expectativa tiene su origen en informes bibliográficos sobre las proporciones entre los números de copias mitocondriales y nucleares en varios tipos de tejidos (54, 55). Por lo tanto, nuestro enfoque para elaborar un ensayo dúplex se basó en encontrar condiciones restrictivas adecuadas para la porción del ensayo de qPCR que analiza el mtND1, mientras que se buscaron condiciones optimizadas para el análisis del nuTH01.

La elaboración de este ensayo se llevó a cabo en varias etapas (42). Con el objetivo de determinar las condiciones óptimas, primero se examinó el ensayo uniplex

de qPCR para el nuTH01 con un rango de concentraciones de 50-900 nM para el cebador y 50-300 nM para la sonda. Estos experimentos indicaron que la sensibilidad y la precisión apropiadas son un CT bajo y poca desviación estándar del CT respectivamente, lo cual se logra al realizar el ensayo de nuTH01 con concentraciones de 600 nM para el cebador y 200 nM para la sonda de TaqMan[®]. En forma similar, examinamos el ensayo uniplex de qPCR para el mtND1 con un rango de concentraciones de 25-900 nM para el cebador y 25-200 nM para la sonda para saber si se podía obtener una sensibilidad y precisión adecuadas al realizar este ensayo con una concentración de 50nM de cebador y de 100 nM de sonda de TaqManMGB[®].

Nuestra meta de realizar el ensayo de mtND1 bajo condiciones restringidas es evidente al observar las cantidades contrastantes de concentración de cebadores que se utilizaron para los dos ensayos uniplex de qPCR. Con el objetivo de observar si esas condiciones de reacción son adecuadas para las amplificaciones dúplex de qPCR, comparamos los resultados obtenidos al cuantificar muestras idénticas, tanto en el formato uniplex como dúplex de qPCR (Fig. 2). A partir de esta comparación se obtuvo que ambos ensayos de qPCR uniplex o dúplex mostraron cuantificaciones y eficiencias casi idénticas. La prueba qPCR de nuTH01 no perdió eficiencia ni mostró un CT retardado al realizarse de manera dúplex vs. uniplex. Debido a una pequeña cantidad de escape en la emisión de señales de fluorescencia de VIC hacia el canal de detección de la sonda fluorogénica FAM, la amplificación dúplex del ensayo qPCR de nuTH01 presentó un CT ligeramente menor que cuando se realizó de manera uniplex. Como un paso final, pusimos a prueba el ensayo dúplex al reforzar una muestra de control de ADN de 3 ng con cantidades crecientes del estándar numérico

de copias de un oligonucleótido sintético de mtND1. Para ello se agregó hasta mil millones de copias mitocondriales efectivas adicionales a la muestra de ADN original. Como se muestra en la Fig. 3, las gráficas de amplificación de nuTH01 y los valores correspondientes de CT no cambiaron en forma notoria, ni siquiera con la presencia de un número irreal y excesivo de copias mitocondriales.

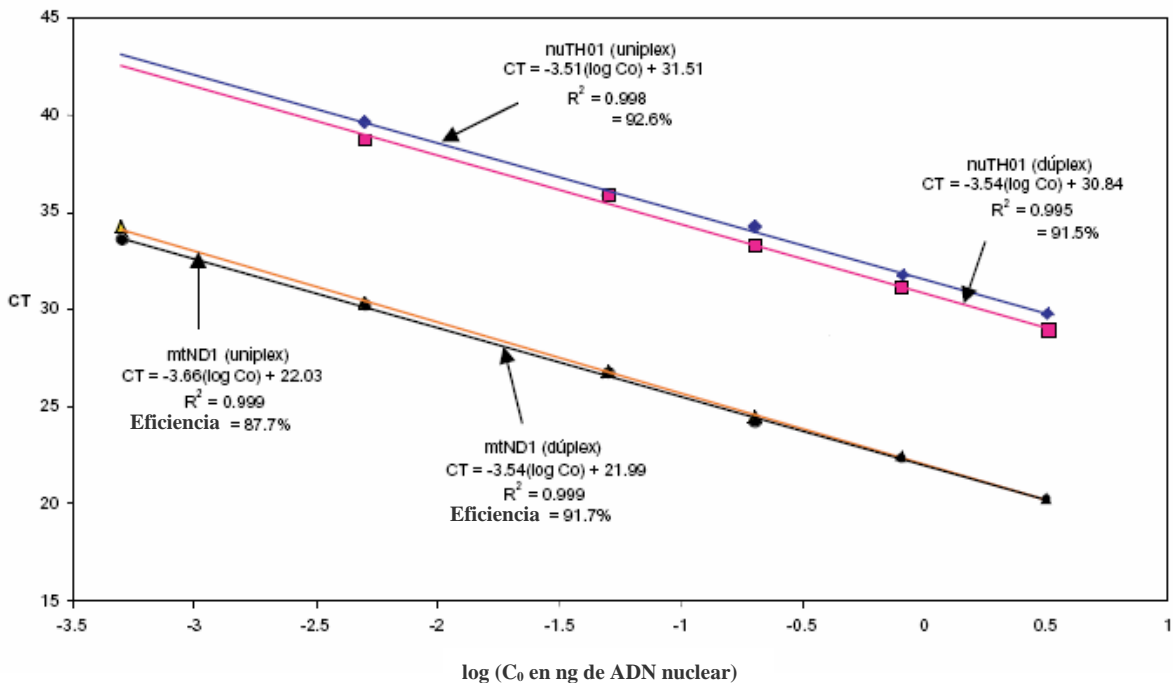


FIG. 2— Las curvas estándares para los ensayos de qPCR de nuTH01 y mtND1 aparecen en la modalidad uniplex y dúplex. Para los ensayos de qPCR del nuTH0, los puntos de los datos que se muestran son los promedios de las mediciones duplicadas de las muestras que contienen 3,2; 0,8; 0,2; 0,05; y 0,005 ng de ADN genómico Promega (femenino). En los ensayos de qPCR de mtND1 también se incluyeron las medidas duplicadas de las muestras que contienen 0,0005 ng de ADN. La eficiencia de la PCR de cada ensayo se calculó con la pendiente de la curva estándar CT v. log C_0 .

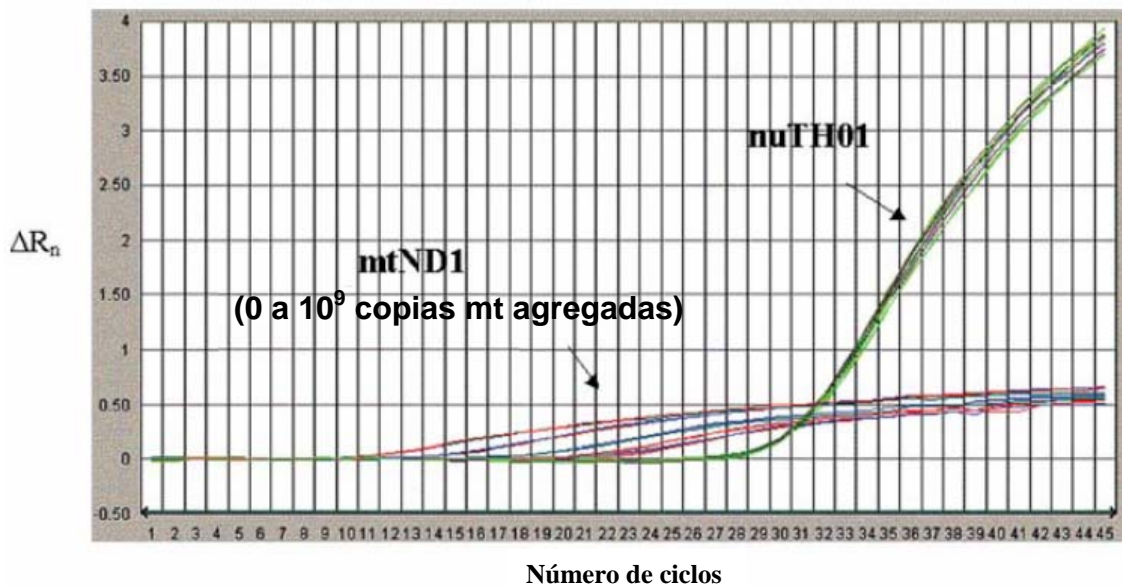


FIG. 3—Curvas de amplificación de qPCR para los ensayos dúplex de nuTH01 y mtND1 con 3 ng de ADN de control junto con copias adicionales de 0, 10^6 , 10^7 , 10^8 , y 10^9 del estándar de mtND1 de una sola cadena. Los ensayos se realizaron en duplicado.

Dado que el ensayo dúplex de qPCR se planeó para limitar las condiciones de amplificación del mtND1 y por otro lado optimizar las del nuTH01, el ensayo dúplex no se optimizó para las muestras cuyo número de copias del genoma nuclear es cercano o mayor al número de copias del genoma mitocondrial. Para estas muestras, cuanto más se refuerce la amplificación del nuTH01, menor será la eficiencia de la amplificación del mtND1, lo cual causaría una desestimación de la cantidad de ADN mitocondrial. Hasta el momento hemos encontrado solamente un tipo de muestra forense (las células de espermatozoides largos y sin cola que se obtienen de glóbulos de extracción diferencial) en las que la proporción de copias del genoma mitocondrial no es grande en comparación con las del genoma nuclear. Aunque este tipo de muestras no se suelen emplear en la haplotipificación mitocondrial, esta limitación del ensayo se debe tener en cuenta al interpretar los datos de cuantificación del mtND1.

Curvas de fusión de SYBR® Green para el TH01

La selección de una secuencia de análisis para la qPCR que abarque la región de repetición de un locus STR introduce la posibilidad de que el ensayo dependa en cierta forma del genotipo de STR, por ejemplo, que el ensayo de nuTH01 arroje resultados de medidas distintas para un genotipo 6,6 de TH01 respecto a un genotipo 9,9.3 de TH01. Dentro del grado de precisión y exactitud de los experimentos de cuantificación, no hemos visto efectos de este tipo en el ensayo de qPCR para el nuTH01. El único efecto observado que depende del alelo fue la presencia de curvas atípicas de fusión de SYBR® Green. Aunque este efecto no altera la calidad del ensayo de qPCR de TH01, es nuevo y requiere de una explicación.

Con los experimentos iniciales de desarrollo, cada ensayo preliminar se evaluó por medio de pruebas de qPCR mediante detección de SYBR® Green I. Este método de detección permite la identificación temprana de combinaciones deficientes de cebadores, lo que evita el empleo de sondas de detección TaqMan® que aplican métodos de tinción, las cuales son más caras. Como parte de esas valoraciones iniciales, evaluamos la especificidad de las amplificaciones de PCR mediante electroforesis de geles (4% de gel de agarosa 3:1 de NuSieve con tinción de EtBr) de las mezclas de reacción de la qPCR posteriores al ensayo, y mediante el equipo de qPCR que se utiliza para elaborar las curvas de fusión de SYBR® Green. Durante el desarrollo observamos curvas atípicas de fusión de SYBR® Green para cada uno de nuestros ensayos tentativos de qPCR de TH01 (Fig. 4).

Para los ensayos de qPCR que se detectan con SYBR® Green, se puede producir una curva de fusión (o disociación) al final del último ciclo de PCR si se

configura el equipo de qPCR para controlar la fluorescencia de SYBR[®] Green de cada muestra mientras se aumenta lentamente la temperatura desde ~60°C (la temperatura de extensión del último ciclo de PCR) hasta ~95°C. Para las temperaturas que corresponden a la fusión de los productos de la PCR, habrá descensos concomitantes en la fluorescencia de SYBR[®] Green debido a la desintercalación del marcador fluorescente o tinte indicador. Para visualizar fácilmente estos cambios en la fluorescencia, se suelen graficar curvas de fusión como la primera derivada negativa del cambio en la fluorescencia ($-dF/dT$) para que cada “transición de fusión” aparezca como un solo pico en la gráfica. Se espera que un ensayo de qPCR bien diseñado produzca un solo amplicón y por lo general, una sola transición en la curva de fusión de SYBR[®] Green.

La Figura 4 muestra las curvas de fusión de SYBR[®] Green tomadas para diversas muestras, mediante el ensayo de qPCR de TH01. Esta gráfica indica que en vez de una sola, cada muestra presenta dos transiciones de fusión (un pico principal de alta temperatura de fusión ($T_f \sim 81^\circ\text{C}$) y un “hombro” de temperaturas de fusión más bajas ($T_f \sim 78\text{-}79^\circ\text{C}$). La observación de estas dos transiciones de fusión no fue anticipada dado que el ensayo de qPCR de TH01 no ha sido muy reconocido. Anteriormente, este ensayo había dado resultados sólo para un producto de la PCR por medio de geles aplicados posteriormente a la amplificación, y no se había encontrado evidencia de transiciones de fusión falsas en ningún control negativo.

La apariencia bifásica de las transiciones de fusión indica la presencia de dos dominios que dependen de la secuencia en cada amplicón de TH01. Este hecho se sustenta en los cálculos de perfiles de fusión que dependen de los alelos, los cuales

se realizan para los amplicones por medio de un paquete de programas computacionales denominado MELT94, el cual emplea información de las secuencias para calcular la temperatura de fusión de un fragmento de ADN con su longitud de banda doble (Fig.5). Dichos cálculos dividen al amplicón del TH01 en dos dominios distintos de fusión (un dominio de baja temperatura de fusión que incluye la región de STR (número de secuencia 0-~96) y otro dominio de alta temperatura de fusión (número de secuencia ~96-~180). Estos cálculos predicen que la temperatura del dominio de baja fusión desciende cuando aumenta el número de STR (Fig. 5, cuadro adjunto), lo cual concuerda cualitativamente con la tendencia observada para las curvas de fusión experimentales que se muestran en la Fig. 4. Esta tendencia se debe a un aumento en el porcentaje de AT de la secuencia del dominio de baja temperatura cuando el número de repeticiones de CATT aumenta. Se predice que el perfil de fusión del dominio de alta temperatura es independiente del genotipo del TH01 (Fig. 5), lo cual es consistente con las transiciones de fusión traslapadas ($T_f \sim 81^\circ\text{C}$) que son independientes de los alelos que se muestran en la Fig. 4.

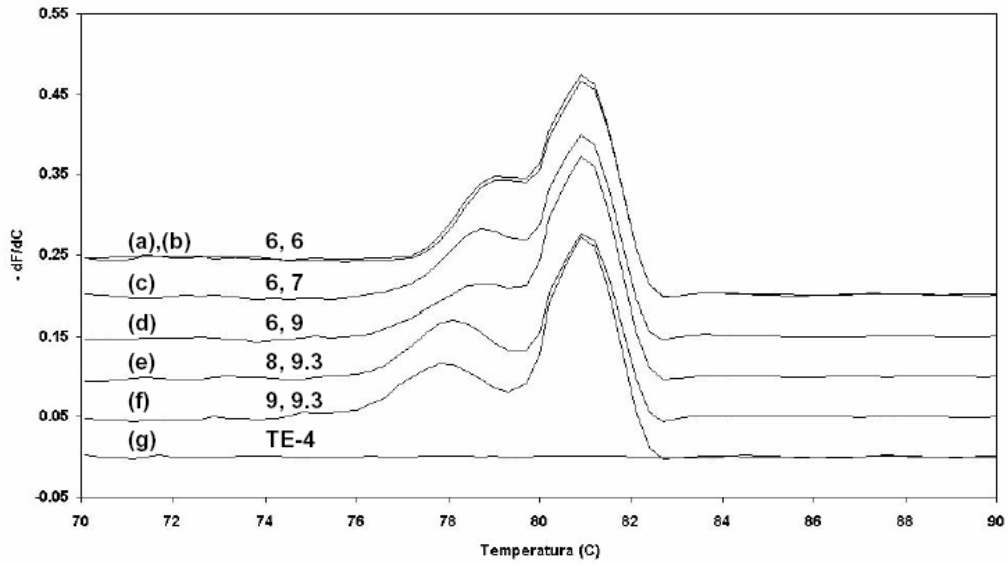


FIG. 4— Las curvas de fusión de SYBR[®] Green que se obtuvieron para los extractos de ADN de seis personas y una muestra de control negativo, mediante el ensayo de qPCR de TH01. Se indica el genotipo de TH01 para cada muestra. Los trazos (a) y (b) corresponden a dos personas distintas, ambas con genotipo 6, 6 en el TH01.

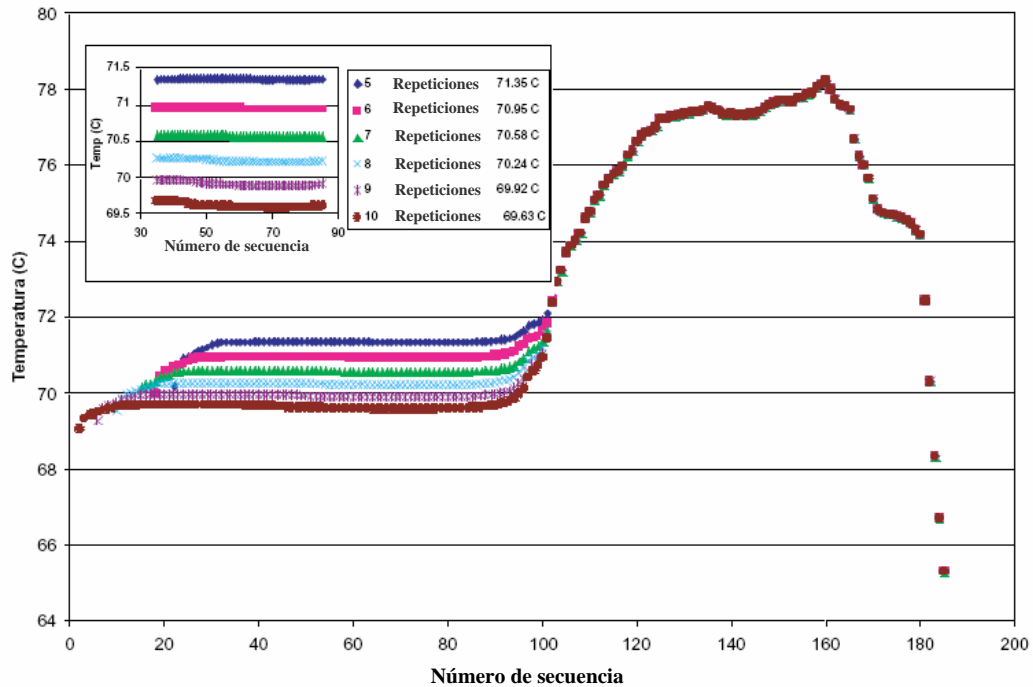


FIG. 5—Perfiles de fusión calculados para los amplicones TH01 (alelos STR 5-10) tal como la genera el programa MELT94. La secuencia número 96 corresponde a la posición de la flecha vertical pequeña de la Fig. 1. El recuadro interno muestra una expansión de los perfiles de fusión en la región STR.

Se ha reportado (56) la fusión de fragmentos cortos de ADN en pasos sucesivos debido a dominios de fusión que dependen de la secuencia. Esto junto con algunos fenómenos de desnaturalización similares que dependen de la secuencia, fundamentan la realización de separaciones mediante técnicas como la electroforesis en geles con gradiente desnaturalizante (DGGE por sus siglas en inglés) (*Denaturing Gradient Gel Electrophoresis*) (57). Sin embargo, hasta donde sabemos, para los amplicones de STR anteriormente no se han registrado transiciones de fusión en pasos sucesivos que dependen del genotipo. Es interesante que las curvas de fusión de SYBR[®] Green para el ensayo de qPCR de TH01 permitan realizar una genotipificación de STR bastante sencilla y rápida, aunque con baja resolución, tanto en este locus como en otros loci STR que presentan curvas similares de fusión que dependen de los alelos.

Validación de la qPCR dúplex para los genes nuTH01 y mtND1: Precisión, sensibilidad y reproducibilidad

La precisión del ensayo dúplex de qPCR para los genes nuTH01 y mtND1 se valoró al realizar 96 ensayos idénticos en un solo plato portador de muestras con 4 ng de ADN estándar previamente cuantificado (ADN genómico humano y de origen femenino de *Promega*) para cada cuantificación. Para la porción del ensayo dúplex correspondiente al nuTH01, medimos un CT promedio de 28,94 ciclos con una desviación estándar (DE) de 0,13 ciclos y un rango de CT de 0,59 ciclos. Luego, para la porción del mtND1 medimos un CT promedio de 22,63 ciclos con una DE de 0,21

ciclos y un rango de CT de 0,98 ciclos. No se observaron desviaciones sistemáticas en el CT de ninguna de las dos porciones del ensayo dúplex a lo largo de las filas y columnas de los 96 ensayos realizados en el plato portador de muestras. Aunque no tomamos en cuenta una serie estándar para la dilución del ADN en esta prueba, la desviación estándar del CT se puede emplear para estimar la desviación estándar relativa (DER o %CV) del número de copias del genoma como $100(2^{(DE \text{ de CT})}-1)$. Esta estimación supone una eficiencia del 100% para el ensayo de PCR. Mediante este método, estimamos que la DER intraplaca del número de copias del genoma es ~10% para el ensayo de nuTH01 y de ~16% para el ensayo de mtND1.

CUADRO 2—Resultados de sensibilidad y precisión obtenidos durante el proceso del ensayo dúplex de qPCR de nuTH01 y mtND1. Los resultados se basan en cinco cuantificaciones duplicadas de los estándares de Promega de ADN genómico femenino diluidos en serie. Las cantidades son de 4µL por muestra. Las desviaciones estándares están entre paréntesis.

Cantidad inicial de ADN (ng)	qPCR de nuTH01		Cantidad inicial aprox. (copias mt)	qPCR de mtND1	
	Cantidad promedio (ng)	(%) DER		Cantidad promedio (copias mt)	(%) DER
100	95,8 (7,7)	8,0	$1,2 \times 10^7$	$8,6 (0,97) \times 10^6$	11
10	10,7 (0,90)	8,4	$1,2 \times 10^6$	$1,3 (0,15) \times 10^6$	12
5	4,7 (0,088)	1,9	$6,1 \times 10^5$	$7,6 (0,95) \times 10^5$	12
1	1,2 (0,034)	3,0	$1,2 \times 10^5$	$1,3 (0,24) \times 10^5$	18
0,5	0,48 (0,057)	12	$6,1 \times 10^4$	$7,1 (1,4) \times 10^4$	20
0,1	0,10 (0,021)	21	$1,2 \times 10^4$	$1,0 (0,20) \times 10^4$	20
0,05	0,051 (0,013)	26	$6,1 \times 10^3$	$7,2 (1,5) \times 10^3$	21
0,01	0,0057* (0,0017)	30*	$1,2 \times 10^3$	$1,3 (0,21) \times 10^3$	17
0,005	—	—	$6,1 \times 10^2$	$7,1 (2,0) \times 10^2$	28
0,001	—	—	$1,2 \times 10^2$	$1,1 (0,39) \times 10^2$	35
0,0001	—	—	$1,2 \times 10^1$	$1,1 (0,41) \times 10^1$	39

*Datos obtenidos a partir de cuatro duplicados porque una de las cuantificaciones se perdió al aplicar esta cantidad de ADN de molde o plantilla.

Para evaluar la sensibilidad del ensayo dúplex de qPCR, se cuantificaron entre 100 ng y 1 fg de ADN de molde o plantilla correspondientes al ADN genómico humano femenino estándar de *Promega*. Todas las cuantificaciones se realizaron en réplicas de cinco para estimar el nivel mínimo de cantidad suministrada de ADN en el que los efectos estocásticos cobran importancia (por ejemplo, interrupción de la señal de qPCR y precisión de cuantificación reducida). Los resultados de tales pruebas se presentan en el Cuadro 2. Como se esperaba, dichos resultados indican que cuando el número de copias del ADN de molde o de plantilla disminuye, ocurre una disminución general en la precisión de ambas porciones del ensayo dúplex de nuTH01 y mtND1. Para la porción del ensayo qPCR correspondiente al nuTH01, se observan niveles adecuados de precisión (<30% DER) para cantidades aproximadas de ADN de molde o plantilla de 50 pg (~15 copias nucleares haploides), y hemos observado también niveles de precisión similares para cantidades aproximadas de 32 pg ó ~10 copias (los datos no se muestran). Al trabajar con cantidades menores de ADN nuclear de plantilla se perdió precisión y aumentó la proporción de curvas de amplificación interrumpidas (por ejemplo, el 20% de los ensayos de qPCR de nuTH01 se detuvieron y no se detectaron cuando se utilizó 10 pg de ADN de plantilla; asimismo sucedió con el 30% de los que se realizaron con 5 pg, el 80% con 1 pg y el 100% con <1 pg). Esto se debe probablemente a los efectos estocásticos del muestreo y en las amplificaciones que utilizan pocas cantidades de ADN de plantilla o de molde. Se observaron niveles adecuados de precisión en la cuantificación (<40% DER) al menos en 100 fg de ADN nuclear de plantilla para la porción mtND1 del ensayo de qPCR, lo cual corresponde con ~12 copias de ADN mitocondrial de plantilla

para el ADN genómico femenino estándar de *Promega*. Todas las curvas de amplificación de la qPCR de mtND1 se interrumpieron al emplear 10 fg de ADN nuclear de plantilla.

Se evaluó la reproducibilidad del ensayo dúplex al comparar las cuantificaciones de seis muestras diferentes de ADN de control con alto peso molecular obtenidas a partir de tres pruebas independientes de qPCR y realizadas por dos analistas diferentes en dos días distintos. En cada prueba las muestras se cuantificaron como réplicas simples. Esos resultados se presentan en el Cuadro 3, donde estimamos la “reproducibilidad” como la DER efectiva para cada muestra al reunir los resultados de la cuantificación de las tres pruebas.

CUADRO 3—Resultados de reproducibilidad obtenidos entre las cuantificaciones dúplex de qPCR de nuTH01 y mtND1 de seis estándares comerciales de ADN genómico de peso molecular elevado (muestras A-F). Las cantidades de slot-blot se basan en una sola medición duplicada para cada muestra. Las cantidades de qPCR se reportan en promedios, al agrupar los resultados de los tres procesos de qPCR, tanto simples como duplicados. Las desviaciones estándares están entre paréntesis. El ADN genómico femenino de *Promega* se utilizó como el estándar de cuantificación para los tres procesos de qPCR.

Muestra	Cantidad para slot-blot (ng/μL)	qPCR de nuTH01		qPCR de mtND1	
		Cantidad promedio (ng/μL)	DER (%)	Cantidad promedio (1000 copias mt /μL)	DER (%)
A	0,85	0,64 (0,08)	12	130 (79)	62
B	0,80	0,96 (0,10)	10	880 (150)	17
C	0,34	0,53 (0,10)	20	310 (35)	11
D	1,31	1,58 (0,31)	19	430 (31)	7
E	0,97	1,51 (0,17)	11	790 (170)	21
F	0,40*	0,45 (0,05)	12	360 (73)	21

*La muestra F representa el patrón de calibración que se utiliza para las cuantificaciones de slot-blot; el valor para la muestra F (0,40 ng/μL) no es una cantidad medida, sino previamente definida.

Estos valores de DER_interplaca en promedio concuerdan razonablemente con los valores de DER intraplaca reportados en nuestra discusión de los ensayos de precisión y sensibilidad. Una notable excepción es la DER simple de longitud larga (62%) que se observó para la cuantificación del mtND1 de la muestra A.

Validación de la qPCR dúplex para los genes nuTH01 y mtND1: La especificidad de las especies

Para evaluar la especificidad de las especies en las amplificaciones de qPCR de nuTH01 y mtND1, el ensayo dúplex se realizó al utilizar ADN de plantilla de cada una de las quince especies diferentes no humanas. Para cada una se utilizaron ensayos duplicados, cada uno con 1 ng de ADN de plantilla. No se registró reactividad cruzada en el ADN no humano, tanto para la porción nuTH01 como para la mtND1 del ensayo dúplex de qPCR. También, retamos al ensayo dúplex al amplificar porciones de 100 ng de cada muestra de ADN microbiana, y no observamos amplificaciones por encima del nivel base en ninguna de las cuantificaciones de nuTH01 ó mtND1. Estos resultados indican que el ensayo dúplex de qPCR para los genes nuTH01 y mtND1 es suficientemente específico para las aplicaciones forenses. Aunque no probamos el ensayo con el ADN de primates mayores, según el trabajo previo realizado con el STR del locus TH01 (49, 50) y las investigaciones de homología de secuencia, anticipamos que el ensayo qPCR para el nuTH01 amplificaría y detectaría esas muestras.

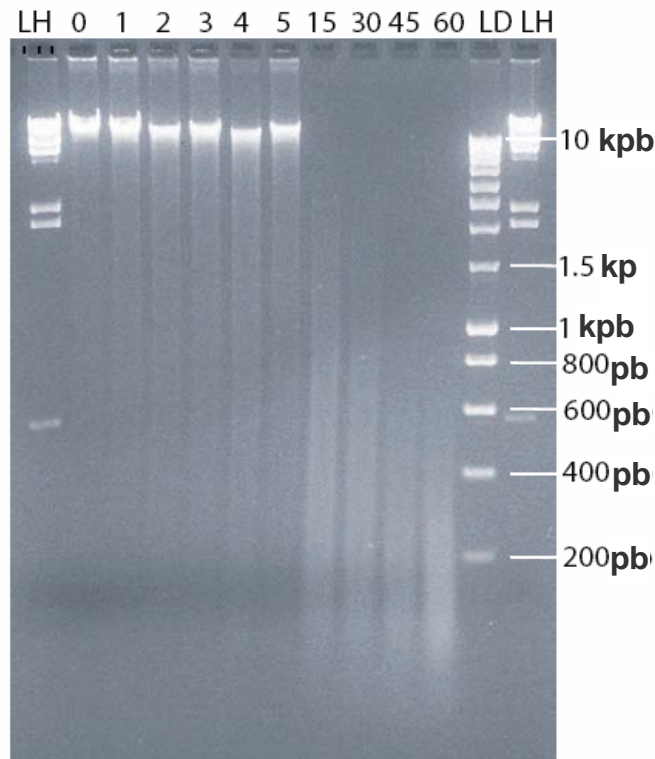


FIG. 6— Serie de degradación de la DNasa para las muestras HL60 (tinción de EtBr). Los carriles etiquetados del 0 al 60 indican el tiempo en minutos del tratamiento con la DNasa. Los carriles etiquetados con LH y LD representan a la λ -HindIII y al marcador estándar “low-mass DNA ladder”.

Validación de la qPCR dúplex de nuTH01 y mtND1: ADN degradado

Las muestras de las pruebas forenses a menudo contienen ADN degradado por la exposición al ambiente o a los microbios. Se preparó una serie de degradación del ADN para medir la capacidad del ensayo dúplex de qPCR de los genes nuTH01 y mtND1, mediante el tratamiento de cantidades conocidas de muestras de ADN genómico de alto peso molecular (HL60) con DNasa I para extender los períodos de tiempo que varían entre un minuto y toda una noche. La Figura 6 muestra un gel que ilustra el grado de fragmentación del ADN para las muestras tratadas con DNasa. El aumento en el grado de fragmentación del ADN durante los primeros cinco minutos del

tratamiento con la DNasa es notorio por la intensificación del “manchado” en la zona de bajo peso molecular de las filas tres a la siete del gel. Durante este período, el nivel de degradación se describe mejor como “moderado” dada la aparición consistente de una banda bastante intensa de alto peso molecular. Sin embargo, no hay evidencia en la banda de alto peso molecular para las muestras que son tratadas con DNasa durante al menos 15 min. Para efectos de esta discusión describimos estas muestras como “muy degradadas”.

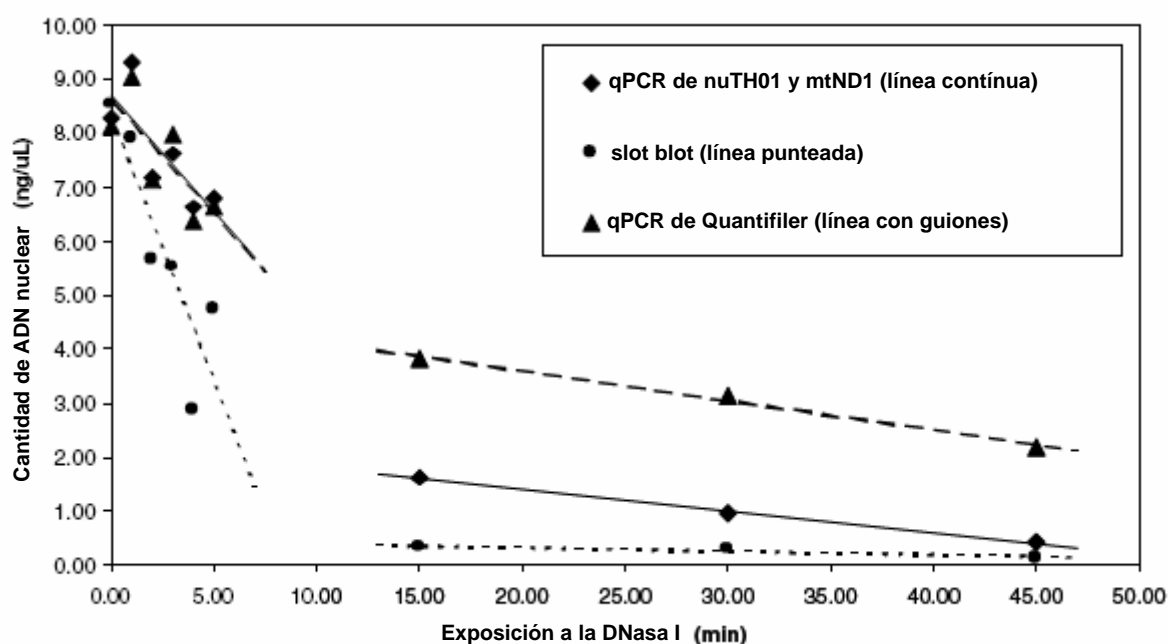


FIG. 7—Cuantificaciones nucleares de las muestras HL60 degradadas con la DNasa mediante qPCR dúplex de nuTH01 y mtND1, qPCR de Quantifiler™, y la hibridación de slot-blot. Para los ensayos de qPCR, cada punto representa los promedios de las cuantificaciones duplicadas. Para la hibridación slot-blot, cada punto representa una sola cuantificación. Las líneas representan ajustes lineales de mínimos cuadrados para los puntos iniciales (0-5 min) y finales (15-45 min) de tiempo en la serie de degradación.

Las muestras provenientes de las series de degradación del ADN se cuantificaron mediante tres métodos diferentes: hibridación de *slot blot*, qPCR dúplex de nuTH01 y mtND1, y mediante el paquete comercial de marcadores genéticos de

cuantificación del ADN humano de *Applied Biosystems Quantifiler*TM. Mientras el ensayo de nuTH01 amplifica una secuencia de análisis de ~170-190 pb, el equipo de *Quantifiler*TM amplifica una de 62 pb. En nuestra comparación incluimos el ensayo de *Quantifiler*TM como medio para estudiar la importancia del tamaño del amplicón para cuantificar el ADN degradado. Para los ensayos de qPCR, se cuantificaron en duplicado 2 μ L de muestra diluida (1:20 en TE⁻⁴) para cada uno de los puntos de la serie de degradación. Para los ensayos de *slot blot*, se obtuvieron cuantificaciones simples a partir de 4 μ L de las mismas muestras diluidas. Los resultados de la cuantificación (sólo cuantificaciones nucleares) para los tres métodos se presentan de forma gráfica en la Fig. 7, a partir de la cual se pueden deducir diversas conclusiones. En primer lugar, para el ADN no digerido de alto peso molecular (la muestra de 0 min), los tres métodos aportan cuantificaciones que concuerdan bastante entre sí en cuanto a los resultados y a las concentraciones (~6 ng/ μ L) estimadas para una dilución 1:20 de ADN de reserva. En segundo lugar, para las muestras moderadamente degradadas (1-5 min de tratamiento con DNasa), los dos métodos de cuantificación de qPCR presentan resultados concordantes, pero ambos arrojan valores de cuantificación mucho más altos de los que se obtienen a partir del método de *slot blot*. En tercer lugar, para las muestras muy degradadas (>15 min de tratamiento con DNasa), ninguno de los métodos de cuantificación concuerdan entre sí. En cuanto a los dos métodos de qPCR, el procedimiento de *slot blot* detecta mucho menos cantidad de ADN. Aunque el ensayo dúplex de nuTH01 y mtND1 no concuerda con el de qPCR de *Quantifiler*TM, éste último detecta mucho más ADN que el primero.

El hecho de que la cantidad de ADN medido en las muestras degradadas dependa del método de cuantificación que se utiliza es un resultado no completamente inesperado. Hay informes previos que indican que el método de *slot blot* desestima la cantidad de ADN nuclear de las muestras degradadas o disfuncionales (14, 23, 24) y se ha planteado que este efecto se debe a una unión deficiente de los fragmentos de ADN degradado a la membrana de *slot blot*. Además, si se toma en cuenta las diferencias de tamaño entre las secuencias de análisis para los ensayos de nuTH01 y *Quantifiler*TM, era de esperar que hubiera algún grado de diferencia en la cuantificación de qPCR del ADN degradado. El ensayo qPCR de *Quantifiler*TM efectivamente mide la concentración de fragmentos de ADN nuclear de al menos 62 pb de longitud, mientras que el ensayo qPCR de nuTH01 mide fragmentos de al menos 170-190 pb de longitud. Para las muestras de ADN de alto peso molecular, estas dos concentraciones son prácticamente las mismas, pero para las muestras muy degradadas, los fragmentos más pequeños presentan mayor concentración respecto a los más largos. Esta diferencia de concentración se indica mediante el gel aplicado posteriormente a la amplificación que se muestra en la Fig. 6 y se refleja en las cantidades divergentes de ADN medido por los ensayos qPCR de *Quantifiler*TM y de nuTH01.

La razón principal para cuantificar el ADN nuclear con exactitud en las muestras forenses es garantizar que se utilice la cantidad correcta de ADN de plantilla o de molde en la amplificación de STR. Cuando se amplifica una cantidad muy pequeña de ADN de plantilla, los resultados de los perfiles de STR podrían estar incompletos, y cuando se amplifica una cantidad excesiva se producen problemas como la falta de

balance interlocus, aumento en las intensidades de pico interrumpidas, adición incompleta de nucleótidos no dirigidos al ADN de plantilla, señales fuera de escala, y señales de detenimiento por contaminación cruzada, los cuales podrían complicar e impedir una genotipificación exacta de STR. Como hemos visto, para las muestras degradadas existen distintos métodos de cuantificación que aportan estimaciones muy diferentes respecto a la cantidad de ADN. Para investigar aspectos relacionados con tales diferencias, intentamos determinar cuál de esos tres métodos ofrece las estimaciones más adecuadas de cantidad de ADN nuclear requeridas para el análisis de STR de las muestras degradadas. Para este fin, utilizamos los resultados de cuantificación que se muestran en la Fig. 7 para preparar partes nominales de 1 ng de ADN nuclear para cada uno de los puntos de tiempo de la serie de degradación de la DNasa según los tres métodos de cuantificación. Estas partes de 1 ng fueron luego amplificadas y genotipificadas mediante el paquete de marcadores genéticos de STR de *Identifiler*TM AmpF/STR[®].

La Figura 8 nos ofrece una visión cualitativa y gráfica de los resultados de genotipificación de STR para la totalidad de las muestras degradadas por la DNasa. El panel (A) de esta figura sugiere que para muestras de ADN no tratadas y degradadas moderadamente (0-5 min de tratamiento con DNasa), los tres métodos fueron apropiados para cuantificar el ADN nuclear para la genotipificación de STR. En estos niveles moderados de degradación, sin importar el método de cuantificación, todos los STR se detectaron con señales mayores que 490 RFU por locus, y no hubo indicios de algún artefacto que pudiera complicar o impedir la genotipificación exacta. Sin embargo, se observó que para las muestras con mayor degradación (15-60 min de

tratamiento con DNasa), las tasas de éxito para la genotipificación de STR dependen del método de cuantificación de ADN.

El panel (C) de la Fig. 8 muestra que para las amplificaciones que se basan en las cuantificaciones de *Quantifiler*TM, una gran cantidad de alelos no se pudo detectar mediante ninguno de los marcadores fluorescentes AmpF/STR[®] *Identifiler*TM. A medida que aumentaba el grado de fragmentación del ADN, los alelos más largos de STR para cada marcador fluorescente fueron los primeros que no alcanzaron el umbral analítico de 100 RFU. Las implicaciones para utilizar una secuencia corta de análisis de qPCR son evidentes. En las muestras muy degradadas se presentaron más fragmentos de ADN con 62 pb de longitud que con tamaño de STR más largo. Como consecuencia, el ensayo de *Quantifiler*TM sobreestimaba la cantidad de fragmentos más largos de ADN de molde o de plantilla, lo cual causaba una amplificación menor de los alelos más largos de STR, y por último, producía cantidades inadecuadas del producto de la qPCR para detección. Con este método de cuantificación y con niveles muy altos de degradación (45-60 min de tratamiento con DNasa), se detectó menos de la mitad de los alelos de STR de *Identifiler*TM, y la sobrestimación de la cantidad de ADN fue tan pronunciada que hasta los alelos más pequeños de STR perdieron una intensidad significativa debido a la baja amplificación (Figs. 8C y 9B, y Cuadro 4).

Por otro lado, la cuantificación de esas mismas muestras mediante el método dúplex de qPCR para los genes nuTH01 y mtND1 produjo un progreso en la tasa de éxito para la genotipificación de STR (Fig.8). Aun para las muestras de más alto estado de degradación (60 min de tratamiento con DNasa), se detectó más del 75% de los alelos de HL60. Para la genotipificación de STR en muestras muy degradadas,

estos resultados indican que la unidad de análisis de ~180 pb de nuTH01 para la qPCR es más apropiada para cuantificar el ADN que la unidad más corta de 62 pb. Los datos que se suministran en el Cuadro 4 muestran que al cuantificar el ADN con base en la unidad de análisis de nuTH01 de mediano tamaño, las intensidades de los STR detectados en el TH01 se mantuvieron en un nivel bastante consistente (~1800-3100 RFU/locus) a pesar del grado variable de la fragmentación del ADN. Aun para las muestras muy degradadas, los STR cortos (~121 pb) de D19 no se sobreamplificaron hasta el punto de formar picos fuera de la escala, ni tampoco los STR largos (~293 pb) de D18 se amplificaron en cantidades tan bajas que no se pudieran detectar.

Las amplificaciones de STR de la cuantificación *slot blot* dejaron muy pocos alelos sin detectar, aun para el caso de las muestras con mayor estado de degradación (Fig.8). Aun así, los electroferogramas que se muestran en los paneles (C) y (D) de la Fig. 9 indican la presencia de diversos artefactos debido a la sobreamplificación de un exceso de ADN de plantilla. Estos artefactos producen “hombros” en la gráfica, dado que algunos nucleótidos que no están dirigidos hacia el ADN de plantilla se agregan en forma incompleta. Además, estos artefactos originan señales de “detenimiento” por contaminación cruzada, debido a los picos que aparecen fuera de escala en los canales detectores adyacentes. Por último, los artefactos también producen picos interrumpidos por encima de los 100 RFU. Para el caso de algunas muestras cuantificadas por *slot blot* en la serie de degradación del ADN, los picos que se encontraban fuera de la escala produjeron señales de detenimiento con una intensidad de 2700 RFU (los datos no se muestran). Los picos fuera de la escala son notorios aun para las muestras que se exponen a la DNasa por

sólo 30 minutos (Cuadro 4). El balance interlocus representado con una proporción de D19/D18 fue mucho peor para las muestras cuantificadas por *slot blot* que para aquellas que se cuantificaron mediante la qPCR (Cuadro 4). Aunque algunos de estos artefactos (por ejemplo, picos fuera de escala, señales de detenimiento) se pueden solucionar al reducir el número de inyecciones de la muestra CE, otros artefactos (por ejemplo, “hombros” en la gráfica por la adición incompleta de nucleótidos, poco balance interlocus) sólo se reparan mediante la reamplificación.

CUADRO 4—Intensidades de STR de AmpFISTR® Identifiler™ seleccionadas para las muestras HL60 degradadas con la DNasa. Todas las amplificaciones utilizaron cantidades de 1 ng de molde o plantilla nuclear al cuantificarse mediante la qPCR dúplex de nuTH01 y mtND1, la qPCR de Quantifiler™ y la hibridación slot-blot. Las intensidades de STR corresponden a RFU por locus. Para las muestras expuestas a la DNasa durante 0-5 min, las intensidades de STR obtenidas a partir de la qPCR son idénticas porque sólo se realizó una amplificación de STR según las cuantificaciones casi idénticas de estas muestras mediante los ensayos de nuTH01 y Quantifiler™ (ver Fig. 7). Los picos fuera de escala, tal como los identifica el programa GeneScan® aparecen en negrita. Para HL60, los tamaños promedio de los alelos STR de Identifiler™ para TH01, D19, y D18 son 177 pb, 121 pb, y 293 pb, respectivamente (60).

Exposición a la DNasa (min)	qPCR dúplex de nuTH01 y mtND1				qPCR de Quantifiler™				Hibridación Slot-blot			
	TH01 (RFU)	D19 (RFU)	D18 (RFU)	D19/D18	TH01 (RFU)	D19 (RFU)	D18 (RFU)	D19/D18	TH01 (RFU)	D19 (RFU)	D18 (RFU)	D19/D18
0	1824	2809	3677	0,76	1824	2809	3677	0,76	2724	2671	4780	0,56
1	1998	1851	2730	0,68	1998	1851	2730	0,68	2861	2360	3747	0,63
2	2883	2489	3253	0,77	2883	2489	3253	0,77	3636	2789	3849	0,72
3	2697	2400	3210	0,75	2697	2400	3210	0,75	3122	2465	3183	0,77
4	2417	2043	2519	0,81	2417	2043	2519	0,81	3791	2880	2637	1,09
5	2469	2119	3133	0,68	2469	2119	3133	0,68	2894	2469	3332	0,74
15	2393	3434	1537	2,23	955	1472	584	2,52	6389	7053	1566	4,50
30	3108	5156	1329	3,88	1036	1786	543	3,29	7797	9103	1766	5,15
45	2540	5408	738	7,33	534	1153	115*	...	6563	9011	903	9,98
60	1941	5249	473	11,10	349	825	5916	8845	543	16,29

*Sólo se detectó un alelo STR del par heterocigoto.

A)	Tratamiento con DNasa	Porcentaje de alelos STR detectados	Número de artefactos de STR*	STR (FAM) azul								STR (VIC) verde					STR (NED) amarillo					STR (PET) rojo					
				D8		D21		D7		CSF		D3	THO1		D13		D16	D2	D19	vWA	TPOX		D18	AML	D5	FGA	
				12	13	29	30	11	12	13	14	16	7	8	8	11	11	17	14	16	8	11	14	15	X	12	22
	Todos los ensayos																										
	0 min	100	- ninguno -																								
	1 min	100	- ninguno -																								
	2 min	100	- ninguno -																								
	3 min	100	- ninguno -																								
	4 min	100	- ninguno -																								
	5 min	100	- ninguno -																								
B)	qPCR de nuTH01 y mtND1																										
	15 min	100	- ninguno -																								
	30 min	96	- ninguno -																								
	45 min	84	2D																								
C)	60 min	76	- ninguno -																								
	qPCR de <i>Quantifiler</i>																										
	15 min	84	- ninguno -																								
	30 min	76	- ninguno -																								
	45 min	48	- ninguno -																								
D)	60 min	28	- ninguno -																								
	Hibridización <i>slot-blot</i>																										
	15 min	100	1 F, 2 D, 2 I, 3 A																								
	30 min	100	1 F, 3 D, 2 I, 3 A																								
	45 min	92	2 F, 3 D, 1 I, 3 A																								
	60 min	92	2 F, 3 D, 3 I, 4 A																								

* F = pico fuera de escala; D= señal de detenimiento detectada; I= señal de interrupción detectada; A = "hombro" detectado por la adición incompleta de A

FIG. 8—Visión general de los resultados de STR de AmpFISTR® Identifiler™ obtenidos para las muestras de ADN HL60 degradadas con DNasa. Cada rectángulo sombreado indica la detección de un alelo STR con una intensidad de al menos 100 RFU. Los rectángulos en blanco indican los alelos que no se detectaron en el umbral analítico de 100 RFU. Para cada marcador fluorescente o tinte indicador, los alelos Identifiler™ se muestran de izquierda a derecha, en orden creciente de tamaño. Las amplificaciones de STR utilizaron cantidades iniciales de 1 ng de ADN nuclear, tal como se cuantificó en el ensayo dúplex de qPCR de nuTH01 y mtND1 (paneles A y B), en el ensayo Quantifiler™ (paneles A y C), o en la hibridización slot-blot (paneles A y D).

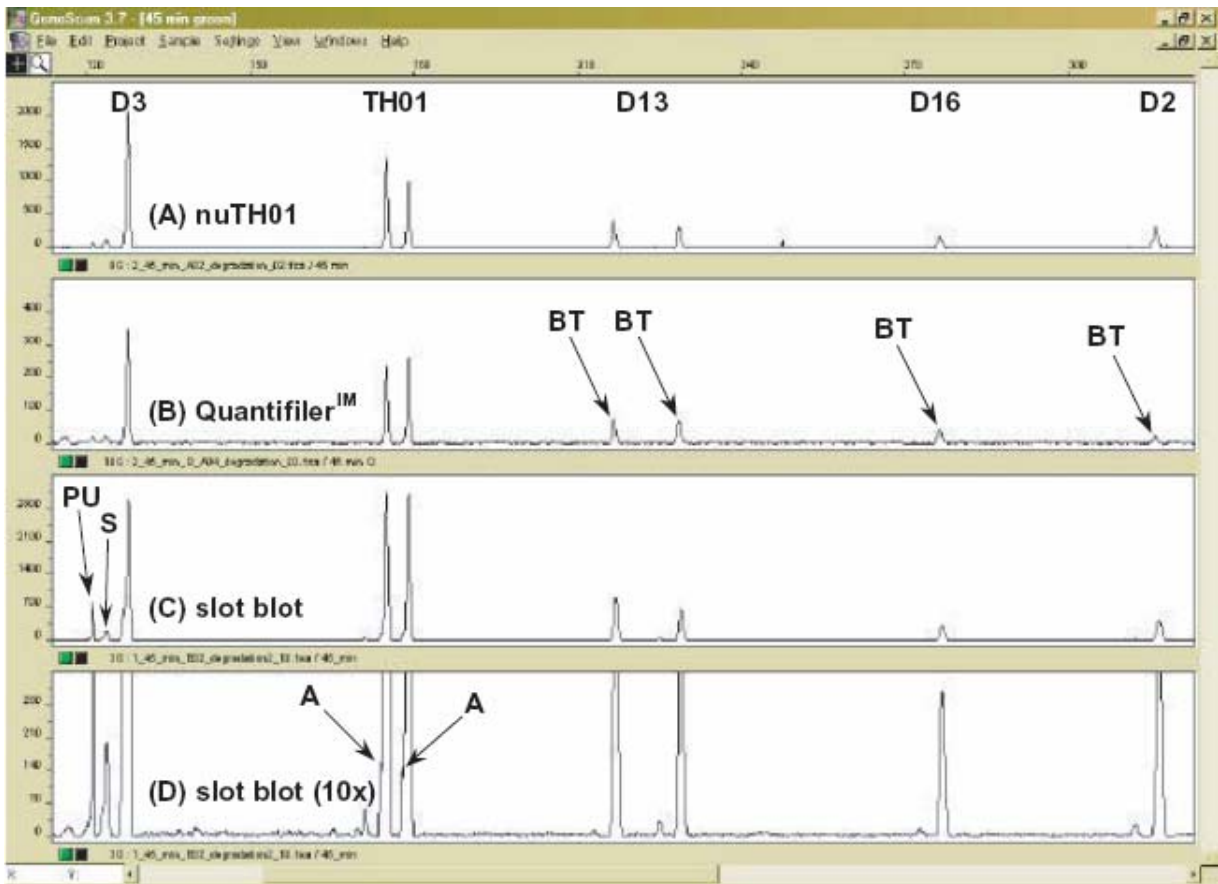


FIG.9—STR con etiqueta VIC de AmpFISTR[®] Identifier[™] obtenidos de las muestras HL60 degradadas con DNasa (45 min). Los paneles muestran alelos STR para amplificaciones de 1 ng de molde o plantilla nuclear, tal como lo determinan: (A) la qPCR dúplex de nuTH01 y mtND1, (B) la qPCR de Quantifiler[™], y (C) la hibridación slot-blot. El panel (D) es una ampliación de diez aumentos de la escala vertical del panel (C). Los artefactos de detección o amplificación se indican como BT= bajo el umbral de detección; PU= señal de detenimiento del canal de detección adyacente; S= señal de interrupción detectada; A= “hombro” detectado debido a la adición incompleta de nucleótidos no dirigidos al molde o plantilla. Los valores completos de RFU en la escala vertical para los paneles son: (A) 2500 RFU; (B) 500 RFU; (C) 3500 RFU; (D) 350 RFU.

Las muestras muy degradadas presentan desafíos analíticos para la cuantificación del ADN y para la subsiguiente genotipificación de STR. Tales muestras suelen constar de fragmentos de una extensa gama de tamaños, y la cantidad y calidad del ADN no se puede representar mediante un único ensayo de qPCR

orientado principalmente hacia la detección y cuantificación de la subpoblación de fragmentos de longitud similar a la secuencia específica de análisis que se amplifica.

En principio, la estimación más exacta de *toda* la cantidad de ADN de las muestras degradadas debería ser proporcionada por los ensayos de qPCR que utilizan secuencias de análisis muy pequeñas. Pero para el fin específico de la genotipificación de STR, no es necesario medir *toda* esa cantidad de ADN. Los resultados indican que si los métodos de qPCR se utilizan para cuantificar el ADN degradado para la genotipificación de STR, es ventajoso seleccionar una secuencia de análisis de qPCR que sea de longitud apropiada para detectar aquellos fragmentos de ADN que son más relevantes para el método de análisis. Esto se cumple para los paquetes comerciales de marcadores genéticos altamente múltiplex para la genotipificación de STR que se utilizan en la mayoría de los laboratorios forenses de ADN. Dichos paquetes recomiendan el uso de un rango bastante estrecho de ADN de plantilla para proporcionar tasas de éxito óptimas para la amplificación y detección de STR. Para el paquete de marcadores de STR *Identifiler*TM de AmpF/STR[®], el cual amplifica STR de longitud que varía entre ~100 pb y ~400 pb, nuestros resultados de genotipificación indican que las cantidades de 1 ng de ADN muy degradado se estiman mejor al utilizar una secuencia de análisis nuclear de qPCR de ~180 pb, tal como en el ensayo dúplex de nuTH01 y mtND1, en vez de una secuencia de 62 pb como la del ensayo qPCR de *Quantifiler*TM. Aunque este estudio se fundamenta en los experimentos que realizamos en un conjunto específico de muestras degradadas por la DNasa, es razonable concluir de forma general que la longitud de la unidad de análisis en el ensayo de qPCR es un aspecto importante para la cuantificación de ADN

en muestras degradadas. Sin embargo, estamos conscientes de que la ventaja de utilizar una secuencia de análisis larga para la qPCR dependerá en particular de la manera en que la muestra se extraiga y purifique.

Además de cuantificar la cantidad de ADN nuclear, también utilizamos la parte del ensayo dúplex de qPCR que analiza el gen mtND1, con el fin de estimar la cantidad de copias del genoma mitocondrial en el mismo grupo de muestras degradadas por la DNasa. Estos resultados de la cuantificación del mtND1 se presentan gráficamente en la Fig. 10. En esta figura, los resultados de la cuantificación mitocondrial que se muestran en cada punto de tiempo de la serie de degradación se normalizan según el resultado de la cuantificación para la muestra no tratada (1,3 millones de copias de mt / μ L). Para establecer una comparación, la figura también muestra cuantificaciones nucleares normalizadas que se estimaron en los ensayos de qPCR de nuTH01 y de *Quantifiler*TM de las mismas muestras degradadas con la DNasa. Nótese que los ensayos de mtND1 y de *Quantifiler*TM indicaron que ~25% del ADN inicial permaneció después de 45 minutos de exposición a la DNasa, mientras que el ensayo de nuTH01 sólo detectó ~5% de la concentración inicial de ADN. Puesto que el ensayo qPCR de mtND1 amplifica una secuencia de análisis (69 pb) relativamente corta, es probable que las cuantificaciones normalizadas para los ensayos de mtND1 y de *Quantifiler*TM sean muy similares a pesar de se trate de genomas completamente distintos. Al observar que la selección de una unidad de análisis nuclear corta produce resultados desfavorables para la genotipificación de

STR, decidimos comprobar si esta misma escogencia para una unidad de análisis mitocondrial afecta la tasa de éxito de la amplificación de los HVI/HVII.

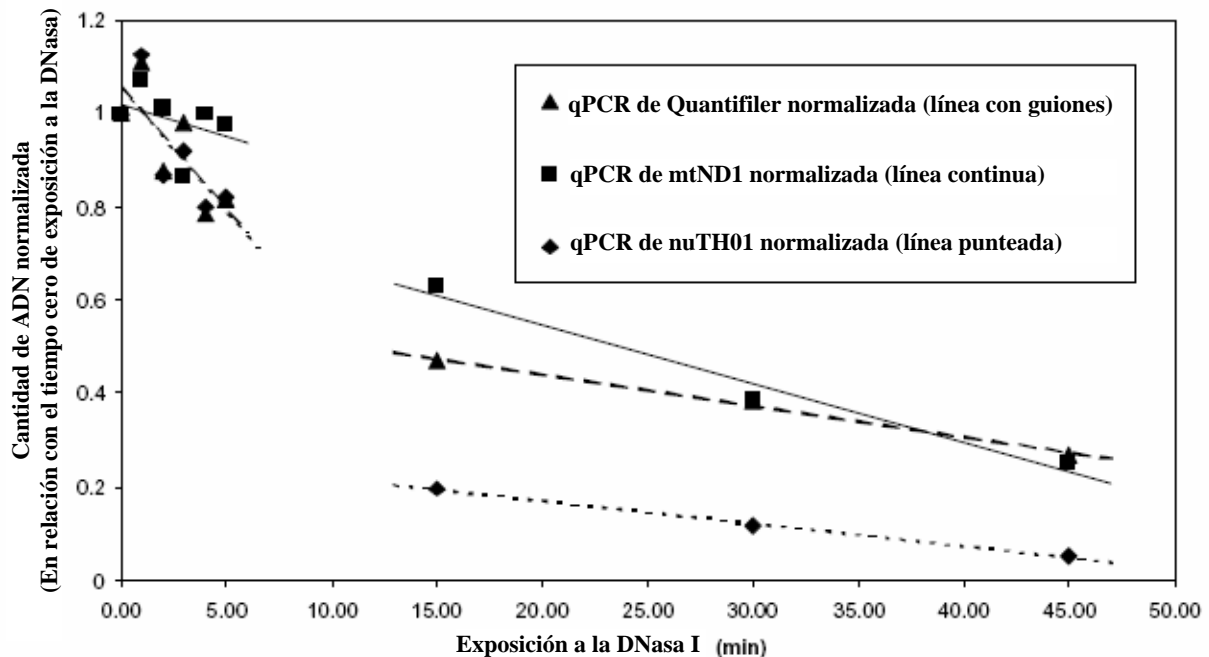


FIG. 10—Cuantificaciones normalizadas de las muestras HL60 degradadas con la DNasa mediante qPCR dúplex de mtND1, qPCR de QuantifilerTM, y qPCR dúplex de nuTH01. Cada punto representa el promedio de las cuantificaciones duplicadas. Para cada ensayo, las cantidades de ADN se normalizaron según la cantidad medida a los 0 min en el tratamiento con DNasa. Por ejemplo, para el ensayo de QuantifilerTM fue de 8,1 ng/ μ L, 8,3 ng/ μ L para el ensayo de nuTH01, y $1,3 \times 10^6$ copias mt / μ L para el ensayo qPCR de mtND1. Las líneas representan ajustes lineales de mínimos cuadrados para los puntos iniciales (0-5 min) y finales (15-45 min) de tiempo en la serie de degradación.

Para tratar este asunto, realizamos dos grupos separados de amplificaciones dúplex de PCR de HVI/HVII en muestras tomadas de la serie de degradación del ADN. En el primer grupo amplificamos 100 pg de ADN nuclear para cada punto de tiempo de la serie, tal y como se midió para la porción de análisis del nuTH01 del ensayo dúplex de qPCR. La estrategia de utilizar la cantidad de ADN nuclear para estimar de forma indirecta la cantidad de ADN de plantilla para las amplificaciones de PCR de los HVI/HVII, se suele emplear en el análisis forense del ADN mitocondrial, siempre que

haya una cantidad suficiente de ADN nuclear para la cuantificación (30, 58). En el segundo grupo amplificamos ~14 000 copias del genoma mitocondrial para cada punto de tiempo, al igual como se estimó para la porción de análisis del mtND1 del ensayo dúplex de qPCR. Para muestras prístinas de HL60, anteriormente determinamos que 100 pg de ADN nuclear representaba ~14 000 copias del genoma mitocondrial. Los resultados obtenidos a partir de los geles aplicados después de la amplificación en los dos grupos de PCR dúplex de HVI/HVII se muestran en la Fig. 11.



FIG. 11—Resultados del gel aplicado después de la amplificación de HVI y HVII para las muestras HL60 degradadas con la DNasa. El gel muestra tinción de EtBr. Cada par de bandas representa los productos de PCR de HVI (~444 pb) y HVII (~416 pb) a partir de 34 ciclos de la amplificación dúplex. La numeración de los carriles del 0 al 60 representa el tiempo en minutos de la digestión con la DNasa. ON representa la digestión con DNasa durante toda la noche. L2 y L4 representan escaleras de bandas (400 pb) de 20 ng y 40 ng respectivamente. En el panel (A), las amplificaciones utilizaron 100 pg de ADN nuclear, tal como se estimó en la porción nuTH01 del ensayo dúplex de qPCR, salvo en la muestra de ON, la cual utilizó 20µL de líquido digestor. En el panel (B), las amplificaciones utilizaron 14000 copias genómicas mitocondriales, tal como se estimó en la porción mtND1 del ensayo dúplex de qPCR.

Se consideró que una amplificación dúplex de qPCR de HVI/HVII es “satisfactoria” si produce suficientes cantidades de productos de la PCR de HVI y HVII para lograr una secuenciación cíclica óptima. Para la secuenciación cíclica, nuestro laboratorio validó un protocolo que utiliza el sistema de secuenciación cíclica de *Applied Biosystems Terminator BigDye*® v 1.1 (30). A pesar de que el fabricante recomienda utilizar 3-10 ng/20 μ L de producto de la PCR en cada reacción de secuencia cíclica de 50 μ L (59), los estudios de validación locales indican que una secuenciación satisfactoria se logra rutinariamente aun con cantidades tan escasas como 250 pg de producto de la PCR por cada reacción. Aun siguiendo la recomendación más conservadora de 3-10 ng/20 μ L, deducimos que una amplificación de HVI/HVII “satisfactoria” es aquella que produce al menos 0,15 ng/ μ L de cada producto de la PCR. Esta concentración es mucho más baja que el límite de detección visual del gel de la tinción de bromuro de etidio. Como consecuencia, nuestro criterio para considerar que una amplificación de HVI/HVII es satisfactoria, consiste simplemente en comprobar que el gel posterior a la amplificación muestre pruebas visibles, o aun pruebas ligeramente visibles para dos bandas de producto de tamaño apropiado. Asimismo, la Fig. 11 muestra que la apariencia visual de dos de esas bandas indican para todas las amplificaciones que ambas unidades de análisis del ensayo dúplex de qPCR cuantificaron satisfactoriamente el ADN mitocondrial para la PCR dúplex de HVI/HVII y que todas las amplificaciones produjeron más de la cantidad necesaria de cada producto de la PCR para lograr una secuenciación cíclica satisfactoria. Aun así, es evidente que hubo ciertas implicaciones para seleccionar una

secuencia de análisis corta para el ensayo de qPCR de mtND1. En particular, la Fig. 11 (panel (B)) revela que para las muestras muy degradadas (15 min de tratamiento con DNasa por la noche) se observa mayor número de bandas débiles de HVI y HVII en el gel aplicado posterior a la amplificación. Esta tendencia señala que el ensayo “corto” de qPCR de mtND1 sobrestimó el número de copias del genoma mitocondrial de ~400 pb de longitud, tal y como el ensayo de qPCR de Quantifiler™ sobrestimó la cantidad de copias del genoma nuclear de 100-400 pb para esas mismas muestras. Sin embargo, en vista de que el protocolo de secuenciación cíclica mitocondrial permite obtener resultados satisfactorios a partir de un rango muy extenso (250 pg-10 ng)) de productos de la PCR, estas sobrestimaciones de la cantidad de ADN mitocondrial en el ensayo de qPCR de mtND1 no causaron errores en la amplificación de HVI y HVII.

Validación de la qPCR dúplex para los genes nuTH01 y mtND1: Muestras de referencia típicas de los casos de trabajo

Para evaluar aun más el ensayo dúplex de qPCR para los genes nuTH01 y mtND1, cuantificamos extractos de ADN a partir de varias muestras de casos trabajo, tanto de referencia como simuladas, para la genotipificación de STR y las amplificaciones de HVI/HVII mitocondriales. En el Cuadro 5 se presenta un resumen de los resultados de un subconjunto de 40 muestras utilizadas en este estudio. Para la genotipificación de STR, cada amplificación utilizó 1 ng de ADN nuclear de plantilla o de molde, tal y como se determinó para la porción del ensayo dúplex de qPCR de nuTH01. Los resultados de STR se representan en el cuadro con los valores de RFU

por cada locus para tres de los quince loci STR de Identifiler™ Amp/STR®, TH01 (163-202 pb), D19 (102-136 pb) y D18 (262-346 pb) (60). Los datos de estos loci particulares se seleccionaron porque sus amplicones representan alelos STR intermedios, cortos y largos respectivamente en el paquete comercial de marcadores genéticos para la genotipificación. Las proporciones de RFU de los D19/D18 se tabulan para indicar la degradación del ADN con base en una métrica simple basada en las STR. Para las amplificaciones de HVI/HVII, se utilizaron ~14 000 copias mitocondriales por reacción, tal y como se cuantifica en la porción del ensayo dúplex de qPCR de mtND1. Para algunas muestras, las cantidades limitadas requirieron una amplificación de menos de 14 000 copias (última columna del cuadro). Las amplificaciones de HVI/HVII se evaluaron mediante electroforesis con geles de post-amplificación, donde los resultados obtenidos a partir del gel se representan con un “+” (para indicar dos bandas de producto de HVI/HVII visibles con intensidades esperadas), un “±” (para indicar dos bandas de producto visibles con intensidades menores a las esperadas), o con un “0” (para indicar que no se observaron bandas de producto). Los resultados de “+” y “±” indican que las amplificaciones proporcionaron más producto de HVI/HVII que el necesario para la secuenciación cíclica.

CUADRO 5—Resumen de los resultados de la cuantificación qPCR de nuTH01 y mtND1 para las muestras de la referencia seleccionada y las del caso de trabajo. Donde se indica, FNE y FE corresponden a las fracciones no espermáticas y a las espermáticas respectivamente, obtenidas de un protocolo de extracción diferencial. Todos los resultados de la cuantificación, ya sean de qPCR o slot-blot, se obtuvieron a partir de experimentos de una sola duplicación. IS indica que la cantidad de muestra fue insuficiente para la amplificación.

Muestra	Número y descripción	qPCR de nuTH01 (ng/4µL)	Slot blot (ng/4µL)	qPCR de mtND1 (1000 copias/4µL)	Proporción de copias genómicas—Mitocondrial: Nuclear	Gel aplicado después de la amplificación de HVI y HVII mitocondrial	Datos del STR de Identifier™ (RFU por locus)				Número aprox. de copias mitocondriales amplificadas para las muestras limitadas
							TH01	D19	D18	D19/D21	
1	sangre líquida	0,42	0,60	200	1600	±	1213	5199	255	20,4	
2	sangre líquida	1,22	0,95	380	1000	+	1349	1065	980	1,1	
3	sangre seca en satén negro	19,0	11,6	1400	240	+	1814	1444	1403	1,0	
4	sangre seca en abono	3,86	9,11	480	410	+	7829	6079	1033	5,9	
5	sangre seca en periódico	0,51	0,45	60	390	+	2522	1586	1363	1,2	
6	sangre seca en mezclilla	0,12	0,10	37	1000	+	4486	3562	775	4,6	
7	sangre seca en lana	2,25	1,50	300	440	+	2096	1335	1230	1,1	
8	sangre seca en alfombra	0,46	0,39	86	620	+	2747	2204	1429	1,5	
9	sangre seca en cuero	0,15	0,10	18	400	+	2562	1468	1292	1,1	
10	sangre seca en mondadientes	1,48	0,66	290	650	+	2466	1399	1145	1,2	
11	femur (en polvo)	0,24	0,05	170	2300	+	2412	3231	378	8,6	
12	pelo sin raíz (Chelex)	0	0	9,1	—	+	IS	IS	IS	IS	
13	pelo sin raíz (DNA-IQ)	0	0	2,0	—	±	IS	IS	IS	IS	5000
14	uña del pie	4,83	0,88	9700	6600	+	2319	3038	459	6,6	
15	uña del pie	1,06	0,22	1400	4400	+	1810	1523	333	4,6	
16	semen en mezclilla (FNE)	0,53	0,13	50	320	+	3459	2335	1323	1,8	
17	semen en calzoncillo (FNE)	0,23	0,09	990	14000	±	2867	4702	456	10,3	
18	tampón vaginal (FNE)	32,1	19,9	880	90	+	3046	1869	927	2,0	
19	afeitado con máquina	1,83	0,80	2900	5200	±	1661	1920	484	4,0	
20	hisopo de máquina de afeitar	0,42	0,07	210	1700	+	1218	915	534	1,7	
21	vendaje ensangrentado	52,4	10,1	1700	110	+	1747	1319	1055	1,3	
21	saliva en colilla de cigarro	0,77	0,29	470	2000	+	1814	1553	671	2,3	
22	saliva en sobre	0,52	0,35	160	1000	+	1890	1308	847	1,5	
24	cepillo de dientes	1,85	1,06	390	700	+	1608	1314	749	1,8	
25	saliva en taza de café	0,84	0,29	270	1100	+	3017	2051	936	2,2	
26	gasa bucal	16,0	5,10	4800	990	+	1677	1182	1130	1,0	
27	gasa bucal	17,8	6,70	4900	910	+	1722	1476	937	1,6	
28	gasa bucal	5,49	2,13	2400	1400	+	2059	1335	1134	1,2	
29	saliva en sobre	0,04	0	3,4	280	±	IS	IS	IS	IS	1300
30	tampón vaginal (FE)	0,04	0	0,036	2,7	0	1204	1417	1005	1,4	300
31	tampón vaginal (FE)	4,68	2,99	0,71	0,5	+	2464	1555	1101	1,4	
32	semen en mezclilla (FE)	2,03	1,61	0,014	0,02	+	1580	1067	1038	1,0	700
33	semen en cuero (FE)	0,27	0,09	0	0	+	3579	2635	1888	1,4	(20 µL)
34	semen en tela de algodón (FE)	39,4	28,5	0,16	0,01	+	2164	1592	1272	1,3	7000
35	semen/sangre en torunda (FE)	0,74	1,06	0,046	0,2	+	1857	1832	1304	1,4	2300
36	semen en calzoncillo (FE)	0,40	0,15	38	310	±	3963	3068	552	5,6	
37	tejido muscular	133	31,4	20000	500	+	1323	620	696	0,9	
38	diente de bebé	0,26	0,23	38	480	±	2017	1846	1354	1,4	
39	diente de bebé	10,6	2,19	800	250	+	2432	1780	1052	1,7	
40	diente	53,1	34,1	5300	330	+	1723	1197	1198	1,0	

Para casi todas las muestras de esos casos, las cuantificaciones obtenidas mediante el ensayo dúplex de qPCR para los genes nuTH01 y mtND1 produjeron amplificaciones satisfactorias de STR y HVI/HVII. Con excepción de una, todas las muestras que contenían una cantidad suficiente de ADN nuclear se amplificaron para obtener un perfil completo de STR. Para la única excepción, la muestra #1, los únicos alelos que no se detectaron fueron los CSF, lo cual es un resultado consistente con el estado de muy alta degradación de ADN que presentaba la muestra ($D19/D18 > 20$). La muestra #4 se inhibió, tal y como se señala en las curvas de amplificación de la qPCR de nuTH01 y mtND1, las cuales presentan pendientes bajas y valores finales de *plateau* reducidos (los datos no se muestran). Dada la presencia de inhibidores de coextracción en la muestra, el ensayo de qPCR de nuTH01 desestimó la cantidad de ADN nuclear, y como consecuencia, se amplificó un exceso de STR del ADN de plantilla y se produjeron señales de intensidad atípica de los alelos TH01 y D19. La muestra #30 provocó una amplificación dúplex de HVI/HVII insatisfactoria; esta muestra se extrajo de la fracción de una célula de espermatozoide que contenía muy poco ADN nuclear o mitocondrial. Varios otros extractos de células de espermatozoides (muestras 31-35) produjeron amplificaciones dúplex de HVI/HVII satisfactorias a pesar de que tenían cantidades muy bajas o casi imperceptibles de ADN mitocondrial, lo cual indica que la porción del ensayo de mtND1 desestimó notablemente la cantidad mitocondrial en esas muestras particulares. Como se discutió anteriormente, esas desestimaciones de ADN que ocurren en el ensayo de qPCR de mtND1 se esperan cuando la proporción entre las copias del genoma mitocondrial y nuclear es baja, tal como el caso de las muestras de fracciones de

espermatozoides. Para las muestras que no provienen de fracciones de células de espermatozoides, se observó que las proporciones empíricas entre las copias del genoma mitocondrial y nuclear ($\#mt:\#nu$) varían entre 90 y 14 000. La amplitud de este rango se debe en parte a las diferencias intrínsecas en la $\#mt:\#nu$ de los distintos tipos de tejidos, y en parte también por las diferencias en el grado de fragmentación del ADN en esas muestras, dado que las secuencias de análisis de ND1 mitocondriales cortos para la qPCR detectan fragmentos de ADN más pequeños que la secuencias de análisis más largas del TH01 nuclear. Es importante notar que para todas las muestras que presentaron una proporción empírica de $\#mt:\#nu$ mayor que dos mil, se encontró una correspondiente degradación del ADN en las razones de intensidad de STR (D19/D18 > ~4), lo cual indica que las proporciones de números de copias del genoma que se estimaron a partir del ensayo dúplex de qPCR pueden suministrar información útil sobre la fragmentación del ADN.

Conclusiones

Desarrollamos un ensayo dúplex de qPCR para la cuantificación específica del genoma humano nuclear y mitocondrial, en una variedad de tipos de muestras forenses. La porción nuclear del ensayo (nuTH01) cuantifica el ADN en muestras con concentraciones que varían desde ~10 pg/ μ L hasta al menos 25 ng/ μ L, mientras que la porción mitocondrial (mtND1) lo hace para un rango de ~4 copias/ μ L hasta ~2 millones de copias/ μ L. Las cuantificaciones fueron adecuadamente exactas y precisas

para determinar las cantidades de ADN de plantilla o de molde para la genotipificación de STR nuclear y las aplicaciones de secuenciación de HVI/HVII mitocondrial.

Para cuantificar el ADN de muestras muy degradadas, los resultados indican que la longitud de la secuencia de análisis de la qPCR es un factor importante cuando se trata de un ensayo para fines forenses. Para la genotipificación de STR, se notó que la porción de análisis del gen nuTH01 durante el ensayo dúplex de qPCR, la cual utiliza una secuencia de ~170-190 pb, es bastante apta para estimar las cantidades de ADN nuclear en muestras muy degradadas. Este ensayo detecta y cuantifica los fragmentos de ADN que son de la misma longitud general que los alelos de STR amplificados en los paquetes comerciales de marcadores genéticos para la genotipificación, y por lo tanto, proporciona un cálculo óptimo de la cantidad de ADN nuclear de plantilla que se debe amplificar para una genotipificación de STR satisfactoria. Al contrario, se obtuvieron resultados deficientes cuando tales muestras muy degradadas se cuantificaron mediante un ensayo de qPCR nuclear que utilizaba una secuencia de análisis mucho más corta (62 pb). Este ensayo que detectaba fragmentos de ADN más pequeños conducía a una sobrestimación de la cantidad de los fragmentos de longitud de STR, y como consecuencia, una gran proporción de alelos de STR no se amplificaban o se detectaban eficientemente. Las cuantificaciones de *slot blot* de las muestras en alto estado de degradación produjeron desestimaciones en las cantidades de ADN nuclear, lo que a su vez provocó que se sobreamplificara un exceso de ADN de plantilla, y ocurrieran picos de STR fuera de la escala así como señales inconvenientes de algunos artefactos.

Menciones

Agradecemos el asesoramiento y la ayuda de varios colegas del Laboratorio de ADN del Departamento de Justicia Jan Bashinski de California (*State of California Jan Bashinski Department of Justice DNA Lab*), entre los cuales mencionamos a Mavos Date Chong, Ram Kishore, Sonja Klein, Jeanette Wallin, Eva Steinberger y Cecilia von Beroldingen. Por el aporte de extractos de ADN a partir de diversos tipos de muestras, agradecemos a los siguientes colegas: Angelynn Moore, Amy Hoover, Christopher Gale, Jessica Battaglia, Dianne Burns, Deanna Kacer y Jeanette Wallin.

Referencias

1. Nicklas JA, Buel E. Quantification of DNA in forensic samples. *Anal Bioanal Chem* 2003;376:1160–7.
2. Waye JS, Presley LA, Budowle B, Shutler GG, Fourney RM. A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. *Biotechniques* 1989;7:852–5.
3. Walsh PS, Varlaro J, Reynolds R. A rapid chemiluminescent method for quantification of human DNA. *Nucleic Acids Res* 1992;20:5061–5.
4. Mandrekar MN, Erickson AM, Kopp K, Krenke BE, Mandrekar PV, Nelson R, et al. Development of a human DNA quantitation system. *Croat Med J* 2001;42:336–9.
5. Hayn S, Wallace MM, Prinz M, Shaler RC. Evaluation of an automated liquid hybridization method for DNA quantification. *J Forensic Sci* 2004;49:87–91.
6. Sifis ME, Both K, Burgoyne LA. A more sensitive method for the quantitation of genomic DNA by Alu amplification. *J Forensic Sci* 2002;47:589–92.
7. Nicklas JA, Buel E. Development of an Alu-based, QSY7-labeled primer PCR method for quantitation of human DNA in forensic samples. *J Forensic Sci* 2003;48:282–91.
8. Fox JC, Cave CA, Schumm JW. Development, characterization, and validation of a sensitive primate-specific quantification assay for forensic analysis. *Biotechniques* 2003;34:314–22.
9. Mackay I, Arden KE, Nitsche A. Real-time PCR in virology. *Nucleic Acids Res* 2002;30:1292–305.
10. Walker NJ. A technique whose time has come. *Science* 2002; 296:557–8.
11. Bustin SA. Quantification of mRNA using real-time reverse transcript PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 2002;29: 23–39.
12. Cockerill FR. Application of rapid-cycle real-time polymerase chain reaction for diagnostic testing in the clinical microbiology laboratory. *Arch Pathol Lab Med* 2003;127:1112–20.
13. Ginzinger DG. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol* 2002;30:503– 12.

14. Nicklas JA, Buel E. Development of an Alu-based, real-time PCR method for quantitation of human DNA in forensic samples. *J Forensic Sci* 2003;48:936–44.
15. Richard ML, Frappier RH, Newman JC. Developmental validation of a real-time quantitative PCR assay for automated quantification of human DNA. *J Forensic Sci* 2003;48:1041–46.
16. Applied Biosystems. Quantifiler™ kits user's manual (part number 4344790B). Foster City, CA: Applied Biosystems, 2003.
17. Andreasson H, Gyllensten U, Allen M. Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. *Biotechniques* 2002;33:402–11.
18. vonWurmb-Schwark N, Higuchi R, Fenech AP, Elfstrom C, Meissner C, Oehmichen M, Cortopassi GA. Quantification of human mitochondrial DNA in real time PCR. *Forensic Sci Int* 2002;126:34–9.
19. Alonso A, Martin P, Albarran C, Garcia P, Primorac D, Garcia O, et al. Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies. *Croat Med J* 2003;44:273–80.
20. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res* 1996;6:986–94.
21. Tyagi S, Kramer F.R. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 1996;14:303–8.
22. Whitaker JP, Clayton™, Urquhart AJ, Millican ES, Downes TJ, Kimpton CP, Gill P. Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. *Biotechniques* 1995;4:670–7.
23. Alonso A, Andelinovic S, Martin P, Sutlovic D, Erceg I, Huffine E, et al. DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. *Croat Med J* 2001;42:260–6.
24. Prinz M, Schmitt C. Effect of degradation on PCR based DNA typing. *Adv Forensic Haemogenetics* 1994;5:375–8.
25. vonWurmb-Schwark N, Schwark T, Harbeck M, Oehmichen M. A simple duplex-PCR to evaluate the DNA quality of anthropological and forensic samples prior to short tandem repeat typing. *Leg Med (Tokyo)* 2004;6: 80–8.
26. Alonso A, Martin P, Albarran C, Garcia P, Garcia O, de Simon LF, et al. Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. *Forensic Sci Int* 2004;139:141–9.
27. Gill P, Jeffreys AJ, Werrett DJ. Forensic application of DNA 'finger[PubMed] prints.' *Nature* 1985;318:577–9.
28. Promega Corporation. Tissue and hair extraction kit protocol (TB307). Madison, WI: Promega Corporation, 2002.
29. Gabriel MN, Calloway CD, Reynolds RL, Primorac D. Identification of human remains by immobilized sequence-specific oligonucleotide probe analysis of mtDNA hypervariable regions I and II. *Croat Med J* 2003;44:293–8.
30. Chong MD, Calloway CD, Klein SB, Orrego C, Buoncristiani MR. Optimization of a duplex amplification and sequencing strategy for the HVI/HVII regions of human mitochondrial DNA for forensic casework. *Forensic Sci Int*. In press.
31. Applied Biosystems. PrimerExpress™ software v2.0 user's manual (part number 4329500A). Foster City, CA: Applied Biosystems, 2001.
32. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank: update. *Nucleic Acids Res* 2004;32:Database issue:D23–6. (<http://www.ncbi.nih.gov/Genbank/GenbankOverview.html>)

33. Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res* 2001;29:320–2. (<http://www.cstl.nist.gov/biotech/strbase/>)
34. Kogelnik AM, Lott MT, Brown MD, Navathe SB, Wallace DC. MITOMAP: A human mitochondrial genome database. *Nucleic Acids Res* 1996;24:177–9. (<http://www.mitomap.org/>)
35. Masibay A, Mozer TJ, Sprecher C. Promega corporation reveals primer sequences in its testing kits. *J Forensic Sci* 2000;45:1360–2.
36. Butler JM, Becker CH. Improved analysis of DNA short tandem repeats with time-of-flight mass spectrometry. *NIJ Science and Technology Research Report*, 2001: Report NCJ 188292. (<http://www.ncjrs.org/pdffiles1/nij/188292c.pdf>)
37. Crews S, Ojala D, Posakony J, Nishiguchi J, Attardi G. Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. *Nature* 1979;277:192–8.
38. <http://www.ncbi.nlm.nih.gov/blast>
39. <http://www.ncbi.nlm.nih.gov/SNP/>
40. <http://www.ncbi.nlm.nih.gov/omim/>
41. Applied Biosystems. TaqManR_ universal PCR master mix protocol (part number 4304449C). Foster City, CA: Applied Biosystems, 2002.
42. Applied Biosystems. User bulletin #5 ABI 7700 sequence detection system: multiplex PCR with TaqManR_ VIC probes (part number 4306236B). Foster City, CA: Applied Biosystems, 2001.
43. Creating standard curves with genomic DNA or plasmid DNA templates for use in quantitative PCR. Applied Biosystems. (http://www.appliedbiosystems.com/support/tutorials/pdf/quant_pcr.pdf)
44. Efficiency of PCR reactions. Stratagene Application Note #10 for the Mx4000 Multiplex Quantitative PCR System. (http://www.stratagene.com/lit_items/appnotes10.pdf)
45. <http://web.mit.edu/osp/www/melt.html>
46. Poland D. Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers* 1974;13:1859–71.
47. Fixman M, Freire JJ. Theory of DNA melting curves. *Biopolymers* 1977;16:2693–704.
48. Lerman LS, Silverstein K. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. *Methods Enzymol* 1984;155:482–501.
49. van Oorschot RAH, Gutowski SJ, Robinson SL. HUMTH01: amplification, species specificity, population genetics and forensic applications. *Int J Legal Med* 1994;107:121–6.
50. Crouse CA, Schumm J. Investigation of species specificity using nine PCR-based human STR systems. *J Forensic Sci* 1995;40: 952–6.
51. Annual report summary for testing in 2002. American Association of Blood Banks, 2003. ([http://www.aabb.org/About the AABB/Standards and Accred/ptannrpt02.pdf](http://www.aabb.org/About%20the%20AABB/Standards%20and%20Accred/ptannrpt02.pdf))
52. Anderson S, Bankier AT, Barrell BG, deBruijn MH, Coulson AR, Drouin IC, et al. Sequence and organization of the human mitochondrial genome. *Nature* 1981;290:457–65.
53. Cummings MP, Otto SP, Wakeley J. Sampling properties of DNA sequence data in phylogenetic analysis. *Mol Biol Evol* 1995;12: 814–22.
54. Szuhai K, van den Ouweland JM, Dirks RW, Lemaitre M, Truffert JC, Janssen GM, et al. Simultaneous A8344G heteroplasmy and mitochondrial DNA copy number quantification

- in myoclonus epilepsy and ragged-red fibers (MERRF) syndrome by a multiplex molecular beacon based real-time fluorescence PCR. *Nucleic Acids Res* 2001;29:e13.
55. Gahan ME, Miller F, Lewin SR, Cherry CL, Hoy JF, Mijch A, et al. Quantification of mitochondrial DNA in peripheral blood mononuclear cells and subcutaneous fat using real-time polymerase chain reaction. *J Clin Virol* 2001;22:241–7.
 56. Li W, Xi B, Yang W, Hawkins M, Schubart UK. Complex DNA melting profiles of small PCR products revealed using SYBR Green I. *Biotechniques* 2003;35:702–6.
 57. Abrams ES, Murdaugh SE, Lerman LS. Intramolecular DNA melting between stable helical segments: melting theory and metastable states. *Nucleic Acids Res* 1995;23:2775–83.
 58. Budowle B, Smith J, Moretti T, DiZinno T. DNA typing protocols: molecular biology and forensic analysis. Natick, MA: Eaton Publishing, 2000.
 59. Applied Biosystems. BigDyeR_ Terminator v1.1 cycle sequencing kit protocol (part number 4337036A). Foster City, CA: Applied Biosystems, 2002.
 60. Applied Biosystems. Application Note: AmpF/STRR_ Identifiler R_ PCR amplification kit (publication number 112AP02–01). Foster City, CA: Applied Biosystems, 2003. Additional information and reprint requests: Martin R. Buoncristiani, M.P.H. California Dept. of Justice Jan Bashinski DNA Laboratory 1001 W. Cutting Blvd., Suite 110 Richmond, CA 94804

DNA ANALYSIS TECHNIQUES IN FORENSIC GENETICS

*Laboratorio de ADN forense, Depto. de Medicina Legal
Universidad de Granada, España**

Carmen Entrala

Index.:

1. Evolution of techniques in the study of DNA polymorphism
2. Polymerase Chain Reaction (PCR) analysis
 - 2.1 Generalities
 - 2.2 PCR components
3. Future technologies: biochips
 - 3.1 Biochip applications

Abstract:

DNA identification or “DNA fingerprinting” is based on the study of a series of DNA fragments present in all individuals; those fragments are characterized by their high variability or polymorphism.

The analysis of a certain number of these DNA sequences or fragments leads to a close to 100% probability of identification between individuals.

In addition to being very polymorphic, DNA used in Forensic Genetics identification is a non-coding or non-expressive DNA, which does not show phenotypic features of individuals; this is very important when considering genetic database creations.

* Forensic DNA Laboratory, Department of Legal Medicine, University of Granada, Spain

In order to analyze DNA polymorphisms, Forensic Genetics laboratories use several techniques that are permanently evolving, thus making DNA identification faster and more accurate.

1. Evolution of techniques in the study of DNA polymorphism

Forensic blood typing started in the early 20th century, when Karl Landsteiner described the ABO blood groups, and Von Dürgen together with Hirschfeld discovered blood group hereditary transmission. This science emerged as a branch of Criminalistics, whose objective was the genetic identification both in cases of criminal investigation and in biological paternity studies. Initially, research was focused on the study of erythrocytary antigens (ABO, Rh, MN groups), serum proteins, erythrocytary enzymes, and the HLA system. The study of those markers enabled the incrimination or exoneration of people as possible suspects due to the similarities or differences of their genetic combination in respect to the biological evidence found in crime scenes.

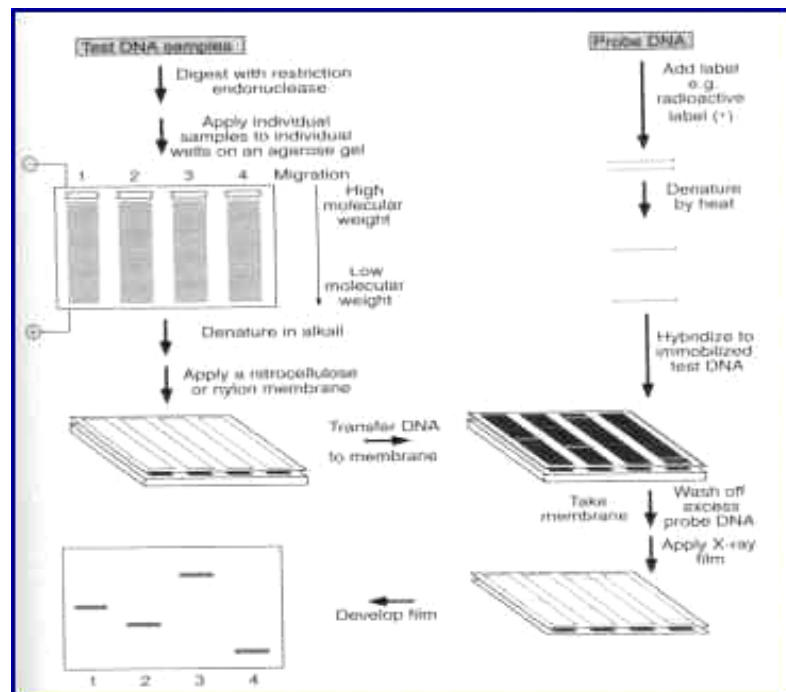
But it was not until the mid-20th century when, thanks to the discovery of DNA and its structure and the later progress in DNA analysis techniques, that Forensic Blood Typing evolved considerably to the extent of becoming today a new Forensic Medicine subspecialty called **Forensic Genetics**. This science basically studies some DNA regions varying between individuals, in other words, DNA polymorphic regions. Therefore, by analyzing a certain number of polymorphic regions, the probability of genetic matching between two individuals becomes practically impossible (except in cases of uniovular twins).

Though science had the necessary tools for DNA study, it was not applied in judicial case resolutions until 1985, when the British Home Office requested the aid of Alec J. Jeffreys, a professor of genetics at the University of Leicester. The first caseworks on Criminalistics were solved thanks to the RFLP (Restriction Fragment Length Polymorphism) technique. Jeffreys discovered the presence of hypervariable minisatellite regions scattered throughout the human genome, which generate variable fragment lengths when treated with restriction enzymes. Later studies also made by Jeffreys showed that length differences between these fragments occurred because the hypervariable regions have a certain number of tandem repeats of a central sequence, and this number varies between individuals.

The first polymorphic DNA locus was discovered in 1980 by Wyman and White by using an arbitrary DNA probe. In this way, they observed fragments of over 15 different lengths in a small sample of individuals. Later, other hypervariable loci were found in human insulin gene sequence, RAS oncogene, zeta-globin pseudogene and myoglobin gene. These hypervariable loci consisted of tandem repeats of an oligonucleotide sequence (11 to 60 bp), such that the different fragment lengths depended on the number of tandem repeats and they were called VNTR (Variable Number of Tandem Repeats).

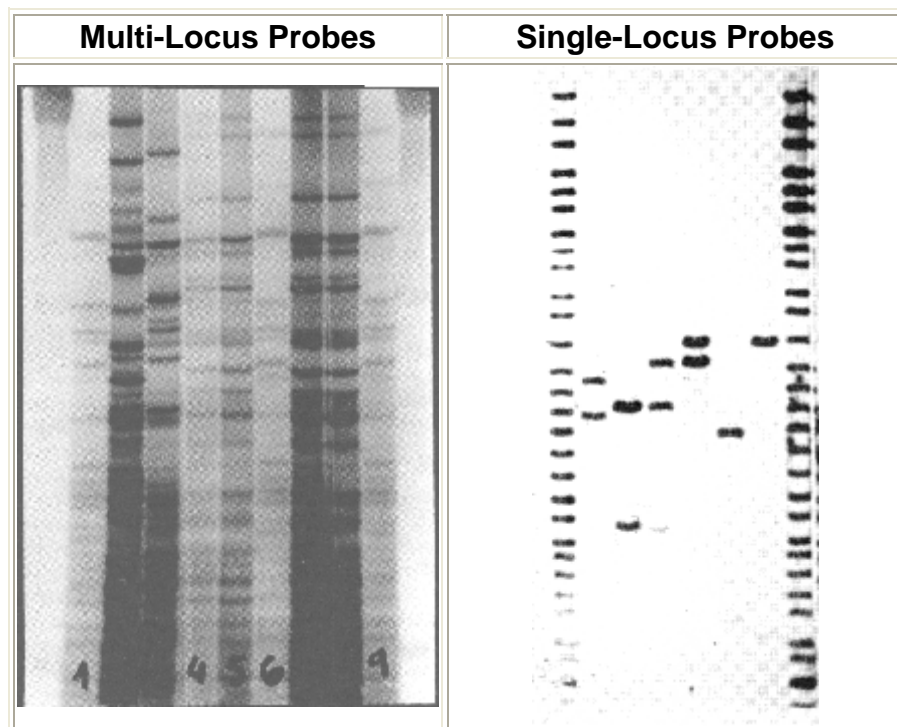
After the first VNTRs were discovered, it was seen that they could be applied to forensic medicine and then replace classic markers. At the beginning, the study of these markers was conducted by means of the Southern blot technique. This technique basically involves the following stages:

1. DNA **digestion** with restriction enzymes after extracting a high molecular weight DNA.
2. Separation of fragments obtained from slab gel **electrophoresis**.
3. **Denaturalization** of cut and separated fragments.
4. **Transfer** and heat-based binding (80°C/176°F) of single strands to a nitrocellulose or nylon membrane.
5. **Prehybridization** with non-specific DNA probes in order to block the non-specific binding sites that could exist in the membrane.
6. **Probe marking** with radioactive nucleotides (normally ^{32}P).
7. **Hybridization** of the marked and denaturalized probe with DNA fragments that are fixed to the membrane, and membrane washing in order to get rid of excess or imperfectly hybridized probes.
8. **Developing** of radiographic films and **interpretation** of results.



The type of probes used may be of two kinds:

- **Single-Locus Probes (SLP):** They are specific for a given chromosome region. They are bound to long nucleotide sequences and have more variability than multi-locus probes. As a result, one or two bands are observed for each individual, depending on the homozygous or heterozygous condition. The band pattern obtained with these probes is called DNA profiling.
- **Multi-Locus Probes (MLP):** They hybridize with minisatellite sequences present at various loci of different chromosomes. They are probes of 10 to 15 nucleotides that are repeated several times, and after developing them, 10 to 20 bands are observed for each person. This multiple-band pattern is known as DNA fingerprinting.



Multi and single-locus probes have several advantages and disadvantages in relation to a series of parameters such as:

- **Information provided:** Since multi-locus probes appear in multiple bands, they have a higher power of discrimination. Nevertheless, single-locus probes are more specific because they hybridize with longer DNA fragments.
- **DNA quantity and quality:** When multi-locus probes are used, approximately 1 µg of non-degraded DNA is required. On the other hand, single-locus probes require less than 100 ng, and it is not necessary for the DNA to be in a perfect condition, as long as the probe complementary fragment is kept intact.
- **Species Specificity:** Multi-locus probes are suitable for DNA from humans and hundreds of superior animals, while single-locus probes are exclusive to human DNA.

Though SLP analysis is and has been useful for paternity studies, the same cannot be said about its application to Criminalistics because it has a number of disadvantages:

- The amount of DNA required ranges from 20 to 100 ng; this quantity is difficult to obtain in criminalistic cases in which biological evidence found at the crime scene is minimal.
- Regarding DNA quality, finding enough non-degraded DNA needed for a single-locus probe analysis is very difficult in forensics.
- The required time for such analysis is two or three days.

- Since this analysis requires large amounts of DNA, the whole sample is usually used up in the first analysis, a fact that complicates counter forensic studies and further case reconsiderations.

All of these limitations were overcome due to the application of a technique in Forensic Genetics, the Polymerase Chain Reaction (PCR), which revolutionized many fields of biology and medicine.

The study of biological evidence from PCR analysis has enabled the resolution of a large number of criminalistic caseworks that were not considered before due to the lack of a large enough sample for a RFLP analysis. With the PCR implementation, minimal samples like a root hair, a tiny blood or semen spot and even dandruff are sufficient evidence to carry out a genetic identification analysis.

2. Polymerase Chain Reaction (PCR) analysis

2.1 Generalities

As long as a nucleotide sequence portion is known, PCR enables DNA from a region selected from genome to be amplified more than a million times. This technique was implemented by Kary B. Mullis, who was awarded the 1993 Chemistry Nobel Prize for this invention.

For PCR, two synthetic oligonucleotides are needed; they require a length of 15-20 nucleotides and have to be complementary to the flanking region that is intended for amplification. These oligonucleotides act as primers for the *in vitro* DNA synthesis, which is typically catalyzed by the Taq Polymerase enzyme. This enzyme is isolated from *Thermus Aquaticus*, a thermophile bacterium, which is able to grow

at high temperatures (79-85 °C/174-185 °F). At this temperature, the Taq Polymerase is able to keep an average extension over 60 nucleotides per second in GC-rich regions. The optimum temperature for Taq Polymerase action makes it possible to use high temperatures for primer binding and extension, so that the reaction efficiency level is increased and the extension of unspecifically-bound-to-DNA primers is reduced.

The reaction is carried out in several cycles, each of which includes three steps:

- 1) DENATURALIZATION:** Single-stranded template DNA is needed in order to start the reaction. This condition is obtained by applying temperatures of 90-95 °C/194-203 °F that break intercatenary hydrogen bridges, and thus separate both strands. In order to achieve full strand separation of the entire sample, that temperature has to be constant for several minutes. If the DNA is only partially denaturalized, it will have the tendency to naturalize again in a fast way, avoiding efficient primer hybridization and a later extension.

- 2) HYBRIDIZATION:** This step is also called annealing. Once the DNA is denaturalized, temperature is decreased to a range of 40-60°C/104-140°F so that primers can be bound to the fragment flanking sequences intended for amplification. Fusion or annealing temperature (T_m , melting temperature) depends on several factors and is relatively specific for each primer. Primer

and sequence lengths are critical when designing amplification parameters; a simple formula to calculate T_m is the following:

$$T_m = 4(G+C) + 2(A+T).$$

In addition, each primer requires a number of experimental studies to determine its specific annealing temperature; if temperature is too low, binding will be made unspecifically, and if it is too high, it will not be completed.

3) EXTENSION: During this step, the Taq Polymerase incorporates nucleotides in the 3'-end primer by using the previous denaturalized DNA strand as template. The temperature for this step is usually 72°C/161.6°F because this is the temperature required for the Taq Polymerase to reach its maximum activity. Normally, a 20-second extension is enough for fragments under 500 bp, and a 40-second extension for fragments over 1.2 Kb.

One important factor in the course of the different steps is the ramp time. This time is defined as the time spent when going from one temperature to another; it depends on the design and features of the equipment where this process is automatically performed, the thermal cycler. In new generations of thermal cyclers, this factor has been optimized to become minimal.

2.2 PCR Components

2.2.1 AMPLIFICATION BUFFER

Typically used PCR buffers include KCl, Tris and $MgCl_2$. The $MgCl_2$ is the most influential component for reaction specificity and efficiency because Mg^{2+} ions are

necessary for Taq Polymerase activity; in other words, they act as polymerase cofactors.

The optimum MgCl_2 concentration is around 1.5 mM, if 200 mM concentrations are used for each one of the dNTPs. Even so, sometimes it is necessary to try different quantities of Mg, since too much of this element leads to the accumulation of non-specific products, and too little to a decrease in amplification efficiency.

2.2.2 PRIMERS

When selecting primers for a DNA fragment amplification, there are several rules to follow:

- **The length of each one of the primers** should range between 18 and 24 bases because it has been proved that longer (30-35 bases) primers do not enhance efficiency and short primers lack enough specificity.
- Both primers should have a similar **T_m** (the difference between both temperatures should be 5°C/41°F at the most).
- Puric base vs. pyrimidinic base ratio should be 1:1 (or 40-60% at the most).
- **Primer sequence** should start and end with 1-2 puric bases.
- In order to prevent **primer dimer** formation, it must be ensured that primers do not have complementary sequences.

Primer dimers consist of double-stranded fragments whose lengths are very close to the sum of the primers and they are produced when one primer is extended after another. The exact dimer formation mechanism has not been fully determined. It has been observed that their formation is facilitated by primers with complementary 3'

ends, and this fact suggests that the initial step is produced by transitory interactions that bring complementary ends near to each other. Some polymerases, including Taq, have shown a weak non-template polymerizing activity, which is able to bind additional nucleotides to the annealed double end. If this activity could be performed on a single oligonucleotide strand, it would be a good opportunity for the extension to form a short overlapping on the 3' end with the other primer, which is enough to foster dimer formation.

2.2.3 DEOXYNUCLEOTIDE TRIPHOSPHATES

Typically used dNTP concentrations are around 200 μM each. In a 25 μl reaction volume with this dNTP concentration, 6-6,5 μg of DNA could be synthesized. dNTP and MgCl_2 concentrations are related because Mg binds to the dNTPs, and high dNTP concentrations would inhibit the reaction due to the lack of enough Taq Polymerase to incorporate dNTPs. For a 200 μM concentration of each dNTP, MgCl_2 is usually added at a 1.5 mM concentration.

2.2.4 TAQ-POLYMERASE

Optimum Taq Polymerase quantities required for DNA synthesis are around 2 units in a 25 μl final reaction volume. This enzyme activity is influenced by the concentration of dNTPs, Mg^{2+} , and some monovalent ions, and high concentrations of these constituents inhibit enzyme activity.

On the other hand, small KCl concentrations enhance Taq synthetic activity by 50-60%, with an apparent maximum when its concentration is 50 mM. There are some data related with the influence of certain reagents used before amplification which alter Taq activity. For instance, 1M urea concentrations enhance Taq activity; whereas low SDS concentrations and concentrations of ethanol above 10% inhibit it.

2.2.5 TEMPLATE DNA

It is the DNA from which a particular fragment is copied. Therefore, it is the template DNA used by the Taq polymerase to synthesize new polynucleotide strands. Necessary DNA quantity for the PCR depends on diverse factors:

The **amplification kit**: some kits have more specific and optimized amplification primers. Consequently, a certain DNA quantity (when using minimal quantities mainly) may be amplified by some markers but not by others. Thus, when using a new kit in a laboratory, a validation study with a sensitivity examination must be carried out. The minimal DNA quantity for a standard amplification can be inferred from that sensitivity study. Generally, for almost all of the STRs used in Forensic Genetics, the optimum DNA quantity that guarantees an adequate efficiency is around 5 ng.

DNA quality: Usually, there are no amplification problems when working with high quality DNA, and even quantities over or under 5 ng can give good results. The problem occurs when the quality of DNA obtained is not optimal either because it is degraded or because it is associated with a series of pollutants that can inhibit

polymerase activity. If DNA is degraded by restriction enzymes, whether results are obtained in the amplification depends on the conservation of the amplification fragment. In the case of a non-degraded DNA containing some pollutants, the sample would have to be diluted to the utmost in order to reduce those pollutants, but it has to be done within a DNA range no lower than the PCR sensitivity limit. The problem of minimum DNA quantities and the presence of pollutants or Taq inhibitors is a usual fact in Criminalistics, and it requires an exhaustive study of the sample before amplification.

PCR ADJUVANTS

These are elements that improve PCR efficiency and specificity. Though some authors have recommended the use of **DMSO** and **glycerol**, the most common adjuvant is **BSA**. Concentrations per enzyme of over 0.8 $\mu\text{g}/\mu\text{l}$ of BSA increase PCR efficiency, since it acts as protein that captures ions that can be Taq Polymerase inhibitors.

When compared with previously used genetic analysis techniques, PCR has many advantages, such as:

- The capacity to get results in cases when DNA quantity is **minimal** or the DNA is partially **degraded**.
- In a short period of time, it generates a **high copy number of target DNA**, enabling the use of simpler and faster visualization techniques than those of probes marked with radioactivity.

- It permits allelic determination and grouping into **discrete classes**, providing an easier database development, since standardization is made immediately; and
- It makes possible the application of biostatistic methods and elaborated programs.

When DNA is degraded, the use of low molecular weight microsatellite markers increases the probability of obtaining positive amplification results because those fragments may not have been digested yet. This advantage is of great importance in Criminalistics because it is typical for collected biological evidence to have been undergoing diverse factors (heat and moisture) that encourage bacterial growth.

However, one of the best advantages of PCR -its high sensitivity- can become a great problem in some cases, since a foreign DNA could be co-amplified with the target one. Nevertheless, precautionary measures to prevent pollution or manipulation problems in Forensic Genetics laboratories are extreme.

Once the DNA has been amplified, the resulting fragments are separated according to their size by means of a process called **electrophoresis**. Currently, two types of electrophoresis are used in Forensic Genetics laboratories:

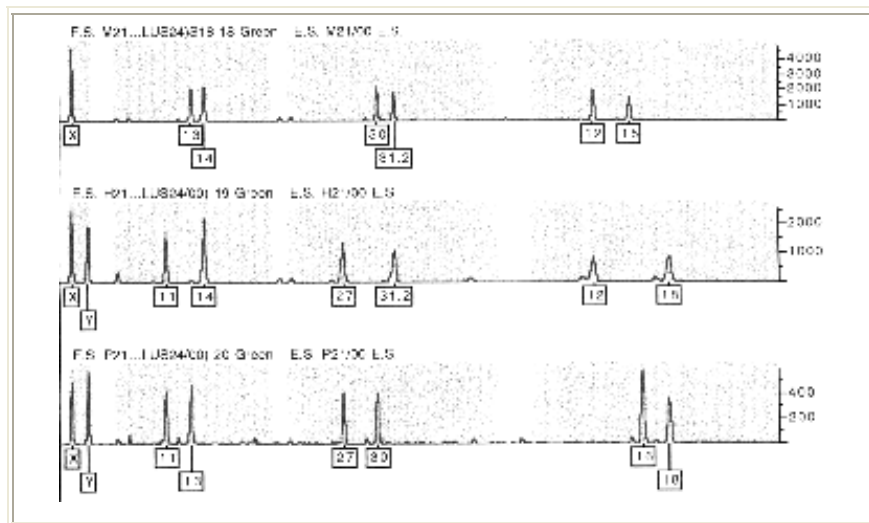
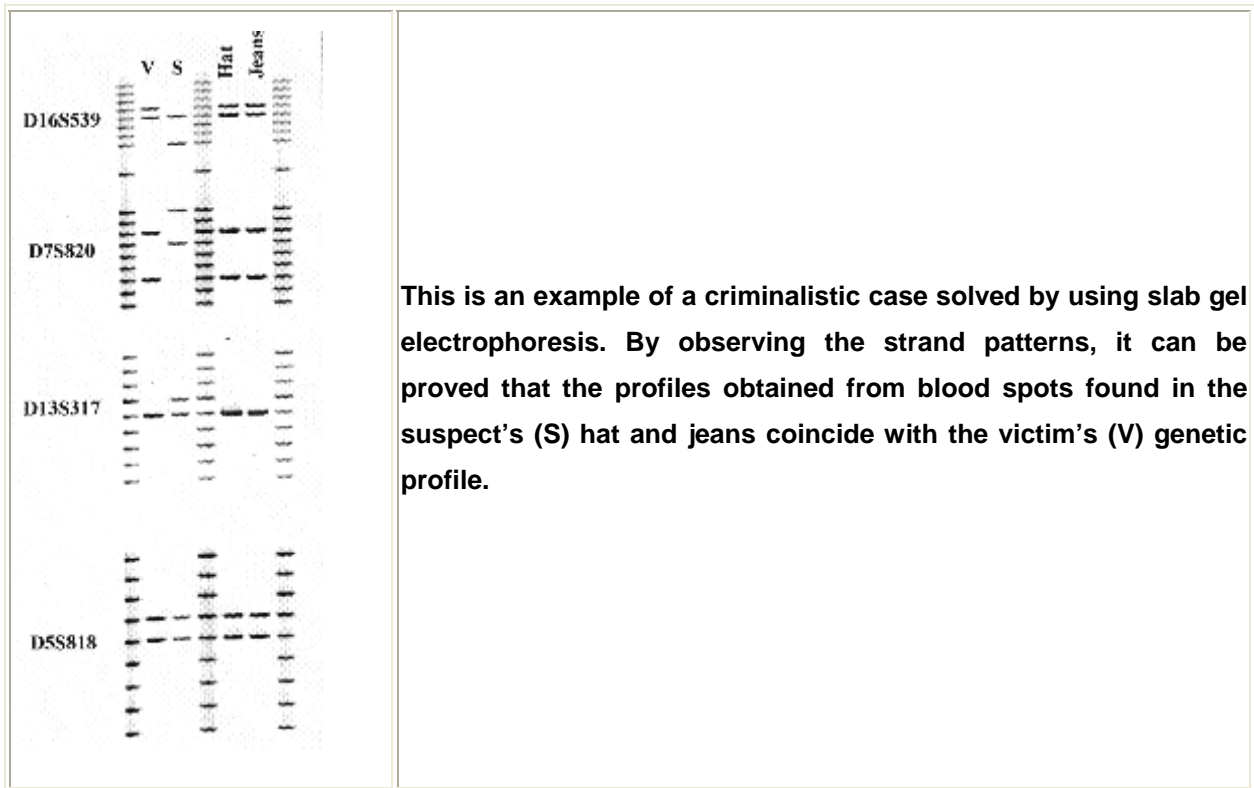
- **Slab gel electrophoresis**
- **Capillary electrophoresis**

Capillary electrophoresis is a relatively new technique in Forensic Genetics, and it is gradually replacing slab gel electrophoresis systems. In this case, the electrophoresis process is carried out in a ~50 μ m-diameter silicon capillary, lowering

the heat amount generated; as a consequence, higher voltages can be applied. In order to analyze primers or dideoxynucleotides (when sequencing) by capillary electrophoresis, they should be fluorescently labeled with some molecules called fluorochromes, which emit fluorescence at a determined wavelength when being excited with laser. The equipment used for this process is attached to a computer in charge of translating fluorescent emission data from sequences or fragments with their correspondingly assigned alleles. Capillary electrophoresis has several advantages over slab gel electrophoresis, such as:

- **Speed:** it enables simultaneous analysis of several loci, though some might have alleles with overlapping sizes.
- **Sensitivity:** it makes possible the detection of tiny amounts of amplified DNA.

Results are obtained by computer, thus preventing interpretation problems and simplifying their analysis by means of software.



3. Future Technologies: biochips

Today, DNA analysis techniques are continually developing and evolving. There is a need for implementing techniques that can enable isolation and analysis of the almost 100,000 genes that constitute the human genome. Therefore, there are a number of research projects addressing the discovery of new methods, which make possible parallel monitoring of high volumes of genetic information as well as analysis time and cost reduction.

Forensic Genetics has significantly evolved since the first DNA polymorphisms were analyzed for identification. Experts in this field have witnessed the revolution that occurred in genetic identification techniques as a result of the discovery of PCR. It is possible that the new revolution will be caused by the **biochips or microarrays**.

Biochips emerged from a combination of microelectronic techniques used for computer microprocessor and biologic material fabrication. Generally, it can be said that the main characteristic of chips is their ability to generate information in a small space, since they produce simultaneous processing of multiple assays. This property will probably make biochips the future technology in biomedical research.

Biochip production is similar to that of computer chips: by means of the so-called photolithography technique, microscopic circuits are placed on silicon sheets. In the particular case of biochips, those sheets are made of glass, and DNA strands are placed on them instead. Glass sheets have several advantages, such as:

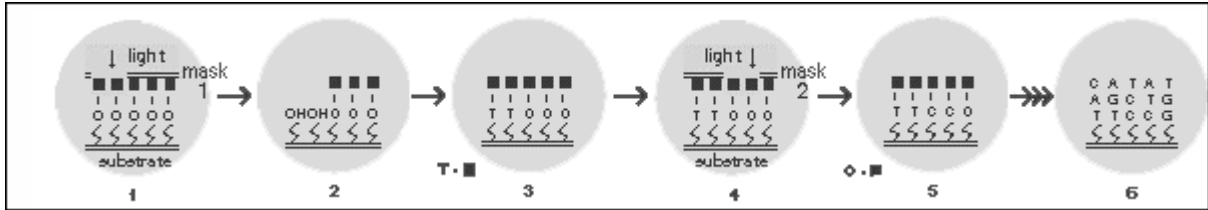
- The possibility of binding DNA strands to the appropriately treated glass surface by means of covalent bonds.

- Resistance to high temperatures and washings at a high ionic force.
- Since the material is not porous, hybridization volume can be reduced at the minimum.
- Low fluorescence prevents background noise.

Other companies have developed different biochip fabrication strategies. Even so, the photolithographic-based biochips are currently the most used and have more applications.

As mentioned before, these chips consist of a small glass sheet with some reagent groups to which nucleotides are later bound, and protected with a film prepared from a photodegradable chemical agent. By way of the mask placed on the chip, a light beam is focused towards some specific regions and positions where the protective chemical agent gets degraded, while regions that are protected by the mask are left intact. Then, the chip is added to a medium containing one of the four nucleotides, which through a covalent bond, will be bound to the glass with reagent groups that have remained unprotected. Each group added carries a photodegradable receptor molecule.

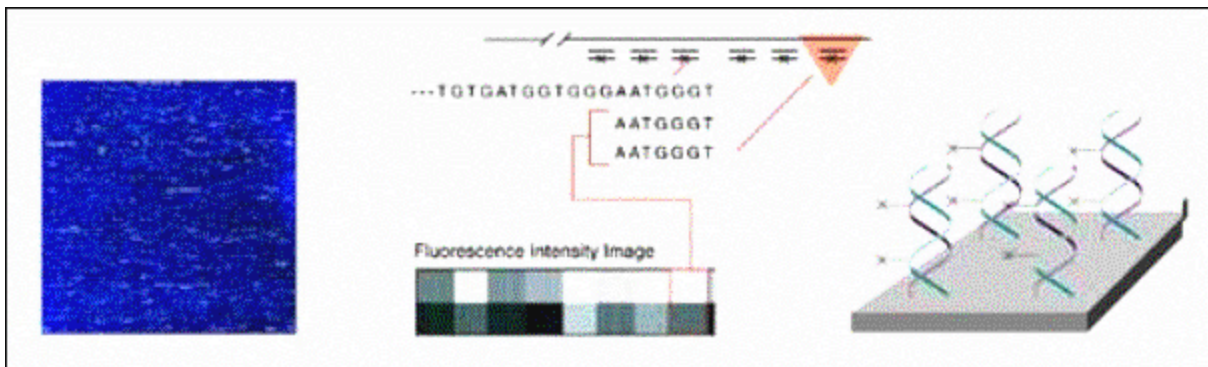
The overall process is repeated with different nucleotides and masks until generating oligonucleotides with the target sequences. The following diagram shows the complete process explained above:



Each chip box has an oligonucleotide strand. Thus, the DNA fragment that can perfectly hybridize with this strand will be the only one to remain bound after several washing procedures.

Before hybridization, the DNA from the target sample should be amplified and fluorescently labeled in one of the ends. Once labeled, it has to be incubated in the chip container, and several washes should be made in order to remove fragments that could not reach hybridization. Then, the chip is placed in a scanner for hybridization pattern detections. This detection is based on the fluorescence emitted by the light-excited fluorochromes from the sample. Fluorescence will be higher at vessels in which binding has been completed than that of vessels containing an unannealed base.

Probe sequence identification according to their chip position is performed by a computer in connection with a scanner.



3.1 Biochip Applications

Though it is a very recent technology that is still under experimentation, biochips are currently being applied in:

- **Genetic expression monitoring:** It enables the determination of the genetic expression pattern and it quantifies the expression level of a high number of genes simultaneously. Those aspects are useful for comparative studies of particular gene activation in healthy and diseased tissue, and thus for the determination of gene function.
- **Mutation and polymorphism detection:** It contributes to the study of all of the possible polymorphisms and it favors the detection of complex gene mutation.
- **Sequencing:** While some biochips have been designed for short-DNA fragment sequencing, there is still no a state-of-the-art biochip that allows long-DNA *de novo* sequencing.
- **Clinical diagnosis and microorganism detection:** It enables fast identification by using genetic markers for pathogens.
- **Screening and drug toxicology:** Biochips make it possible to analyze genetic expression changes that occur during fast drug administrations. They also facilitate the localization of new possible therapeutic trends and their related toxicological effects.
- **Therapy follow-up:** With biochips it is possible to assess genetic features that may influence therapy response.

- **Preventive medicine:** Knowledge and possible diagnosis of certain genetic features associated with particular pathologies are useful to prevent diseases before symptoms arise.

Copyright © 2000 by Carmen Entrala, Doctor in Biology.

Documents for the doctorate course

Informe de investigación

**LOS PROCESOS SINTÁCTICOS DE INTERFERENCIA
LINGÜÍSTICA EN LA TRADUCCIÓN INVERSA
(ESPAÑOL-INGLÉS) DEL SINTAGMA NOMINAL
EN EL GÉNERO CIENTÍFICO**

INTRODUCCIÓN

Presentación general

En este proyecto se traducen y analizan tres artículos científicos, dos en inglés y uno en español, cuyos respectivos campos temáticos pertenecen a la genética y biología molecular, aplicadas al área de investigación forense. El título original de cada artículo es: 1. “An Overview of Human DNA Analysis in Forensics” 2. “A Duplex Real-Time qPCR Assay for the Quantification of Human Nuclear and Mitochondrial DNA in Forensic Samples: Implications for Quantifying DNA in Degraded Samples” 3. “Técnicas de análisis del ADN en genética forense”. El primer artículo se publicó en noviembre del 2006 en la *Revista de Ciencias Forenses* y se presentó oficialmente en Costa Rica por medio de la Asociación Costarricense de Ciencias Forenses (ACCF) del Poder Judicial de Costa Rica. El segundo se publicó en septiembre del 2005 en los Estados Unidos por medio de la revista: *Journal of Forensic Sciences*. Finalmente, el artículo en español se publicó como un documento del curso de doctorado en biología forense de la Universidad de Granada, España, en el 2000.

El primer artículo consiste en una reseña científica sobre la historia y desarrollo de los distintos tipos de análisis de ADN dentro del área forense. Se espera que los lectores sean tanto especialistas como estudiantes universitarios de grado avanzado, o de posgrado de biología interesados en informarse sobre estas técnicas. La función textual consiste en presentar, a manera de resumen, la evolución y los avances científicos sobre este campo temático, y así presentar una

visión en conjunto sobre las investigaciones señaladas. La macroestructura textual es típica de los artículos de reseña científica: resumen, introducción, desarrollo de subtemas, conclusión y referencias. Además, las tablas de contenido y las imágenes son elementos no verbales que forman parte del texto.

El segundo artículo es una publicación científica primaria. Se espera que los lectores meta sean expertos en genética y biología molecular, y que trabajen en el área forense. Este tipo de texto supone un conocimiento más especializado, propio de expertos de un nivel de postgrado, conocedores de un léxico muy específico. La función textual es informar a la comunidad científica sobre los avances en el análisis de ADN en muestras degradadas. Para ello se basa en la presuposición de que los lectores conocen todos los conceptos fundamentales sobre el análisis de la PCR o reacción en cadena de la polimerasa, y están interesados en conocer más sobre el nuevo método denominado “reacción en cadena de la polimerasa cuantitativa dúplex en tiempo real”, lo cual define la importancia y el aporte del artículo. La macroestructura textual presenta los movimientos típicos de los artículos científicos (resumen, descriptors, introducción, materiales y métodos, resultados y discusión, conclusiones y referencias). Asimismo, el artículo presenta elementos no verbales como gráficos, imágenes y cuadros.

Por último, el tercer artículo es un compendio de información científica especializada sobre las técnicas de análisis de ADN. El léxico es especializado, y se presupone que los lectores meta sean estudiantes de doctorado en biología de la Universidad de Granada, España, que desean conocer los fundamentos temáticos de esta área de estudio. Dado que este artículo se publicó en la Internet, se espera que

sea consultado también por estudiantes equivalentes de universidades latinoamericanas, entre otros. La macroestructura textual es similar a la del primer artículo de este proyecto, y se asemeja a un artículo de reseña científica porque presenta la evolución de las técnicas de análisis, describe algunas de ellas, especialmente la PCR, y se refiere a las tecnologías previstas para el futuro, partiendo del año 2000.

Justificación de la escogencia y el análisis de la traducción

Respecto a la justificación temática, a nivel mundial los artículos científicos por lo general se escriben en inglés, o bien, se traducen al inglés. El motivo de este tipo de encargo de traducción se debe a que este idioma se ha convertido en la lengua internacional para la publicación científica. Según algunas entrevistas realizadas con especialistas, para que un investigador alcance lectores y audiencias reconocidas, necesariamente debe publicar sus artículos en inglés y quien sólo publica en español se limita a una divulgación de nivel latinoamericano o peninsular, sin poder alcanzar a las esferas científicas de África, Asia, Europa, Oceanía y parte de Norteamérica. Por lo tanto, la labor del traductor latinoamericano respecto a este género de texto en gran parte se centra en la traducción inversa, y más si se tiene también en cuenta que la mayoría de los científicos leen y comprenden los artículos en inglés. Además, este proyecto de traducción se presenta con el propósito de penetrar en el campo de la traducción especializada, la cual es un desafío para el traductor costarricense.

Respecto a la justificación del análisis, en este trabajo se establece que la traducción inversa (español-inglés) del sintagma nominal supone un esfuerzo

sintáctico importante que implica el desapego de las estructuras de la lengua materna, e implica una adquisición o asimilación, inconsciente o consciente de algunas estructuras de la segunda lengua. Esta afirmación se fundamenta en la experiencia de traducción de este proyecto, la exploración de artículos en inglés traducidos por hispanohablantes nativos, y la práctica de la retrotraducción al inglés de artículos traducidos al español, como técnica sintáctica comparativa. Como consecuencia, en este proyecto se considera la necesidad de profundizar sobre este tema, y se esbozan algunos puntos de debilidad en los que el traductor hispanohablante es más susceptible a incurrir en interferencias sintácticas sutiles, específicamente en la traducción del sintagma nominal.

Hipótesis

El proceso de traducción inversa (español-inglés) de los sintagmas nominales del inglés en el género científico suele implicar un grado de interferencia sintáctica sutil por parte de los traductores hispanohablantes nativos, la cual se manifiesta mediante la reproducción de construcciones sintácticas del español, o por el contrario, mediante la hipercorrección y la sobregeneralización.

Antecedentes

La traducción científica y técnica ha sido estudiada en algunos manuales y libros sobre traducción, tales como los de Jiménez (2002), Maillot (1997), entre otros. Esta bibliografía se centra mayormente en estudios sobre terminología, convenciones del género textual como el uso de abreviaturas, siglas, etc; así como algunos rasgos macroestructurales y microestructurales generales. El papel de estas

fuentes en este trabajo se ha limitado a proveer un panorama general sobre la caracterización del género científico. Dado que este proyecto se ocupa de los sintagmas nominales, la bibliografía más importante en el campo de la traducción es tomada primeramente de Alcaraz (31-32, 62-64), que enumera las preferencias sintácticas de estos textos, y entre ellas menciona la importancia de la complejidad de los sintagmas nominales largos como un factor que contribuye a la precisión expresiva y a la economía comunicativa. Sin embargo, Alcaraz (62) sólo menciona este rasgo lingüístico y se refiere a su efecto estilístico en el texto, pero no profundiza más sobre su descripción lingüística, ni se centra en su traducción.

En este proyecto, el fundamento teórico para analizar la sintaxis del sintagma nominal es la lingüística descriptiva de Bloomfield, desarrollada por Gleason (138-298), Francis (291-325, 355-366) y Hockett (177-197), principalmente. También se consulta bibliografía detallada sobre la gramática inglesa, como el caso de Quirk y otros (1985), con el fin de identificar algunos rasgos estructurales del inglés compilados por otros autores. Respecto a la descripción de los procesos de interferencia sintáctica, este proyecto se plantea con base en el análisis del corpus y el seguimiento de la metodología de investigación. Esto da al proyecto un carácter innovador porque trata la interferencia sintáctica desde el punto de vista de la traducción inversa.

A continuación, se resumen los objetivos planteados para este trabajo.

Objetivo general

Identificar algunos factores sutiles de interferencia sintáctica que supone la traducción inversa (español-inglés) del sintagma nominal dentro del género científico, mediante el estudio lingüístico descriptivo y comparado de este tipo de construcciones.

Objetivos específicos

1. Describir y analizar distintos tipos de construcciones sintácticas de las frases nominales presentes en los artículos científicos traducidos, tanto en inglés como en español, mediante el método de la gramática descriptiva.
2. A partir de la exploración descriptiva y comparada anterior, confirmar y describir algunos procesos de interferencia sintáctica que pueden experimentar los traductores hispanohablantes durante la traducción inversa de los sintagmas nominales al inglés.

Para llevar a cabo el análisis, se procede de acuerdo con los siguientes pasos:

Metodología

- 1) Analizar los tipos de sintagmas nominales de los artículos científicos traducidos en este proyecto, así como los encontrados en algunos textos paralelos, y los retrotraducidos al inglés, comenzando por la descripción y

comparación lingüística de los tipos básicos en ambos idiomas, para luego profundizar con la caracterización de los más complejos.

- 2) Describir los procesos de interferencia sintáctica a partir de la investigación teórica anterior, con el fin de proponer posibles causas estructurales, las formas resultantes y los mecanismos de este fenómeno.

CAPÍTULO 1

MARCO TEÓRICO

Características sintácticas de los textos científicos

Este proyecto de investigación parte del análisis de la traducción inversa al inglés de artículos científicos por parte de hispanohablantes. La hipótesis surge mediante la observación del comportamiento de los traductores en el manejo de rasgos lingüísticos sintácticos específicos, y el análisis de su efecto estilístico en el género textual científico. La característica estructural en cuestión es el sintagma nominal complejo o múltiple, el cual se ha catalogado como una construcción muy frecuente y casi exclusiva del discurso escrito en el campo de la ciencia. De acuerdo con Quirk (1985), el sintagma nominal es un indicador de estilo cuyo efecto es sensible dentro de los textos científicos, y su modalidad de uso depende del propósito y el campo de estudio. Por lo tanto, el uso de esta construcción contribuye en gran parte a la aceptabilidad de la traducción inversa en este género textual.

Alcaraz (2004) menciona que entre los rasgos sintácticos del inglés académico, y en este caso científico, se pueden citar: transitividad, pasivización, nominalización, verbos ergativos, personificación, cosificación, además de algunas preferencias sintácticas (62) como:

- 1) la sencillez de la estructura oracional para lograr precisión conceptual y cortesía comunicativa,

- 2) las relaciones paratácticas (yuxtaposición y coordinación) que promueven la objetividad expresiva y comprensiva de los datos, así como el uso de algunas expresiones hipotéticas,
- 3) la complejidad de los sintagmas nominales largos para lograr precisión expresiva y economía comunicativa.

Por otro lado, en cuanto al género científico en español, Jiménez (116-117) cita a Criado Pérez (19), y establece que en este género hay preferencia por los complementos circunstanciales, uso de oraciones subordinadas que expresan condiciones límites, construcciones personales con *se*, imperativos que evitan apelación a una persona determinada, así como evasión de referencia directa y concreta del agente de la acción. Además, Jiménez enumera otras particularidades morfosintácticas expuestas por otros autores, entre ellas: escasez de artículos, abundancia de sustantivos (sobre todo abstractos), la adjetivación eminentemente especificativa, el predominio de verbos descriptivos en general, uso del tiempo presente en las definiciones, uso de conectores como elementos de primer orden, y el carácter enunciativo y objetivo de la modalidad (117-118).

Duque clasifica a los artículos científico-técnicos como publicaciones primarias o inéditas y publicaciones secundarias. Las primeras corresponden a artículos originales de investigación, donde se describen nuevos experimentos o enfoques teóricos, se informan los resultados y se argumentan las implicaciones de estas investigaciones (44). Además, van dirigidos a una audiencia muy experta en un campo determinado (45). Por ejemplo, el artículo "*A Duplex Real-Time qPCR Assay*

for the Quantification of Human Nuclear and Mitochondrial DNA in Forensic Samples: Implications for Quantifying DNA in Degraded Samples” traducido en este proyecto, corresponde a una publicación primaria o inédita. Por otro lado, Duque (45) incluye dentro de las publicaciones secundarias aquellos artículos cuyo fin consiste en recopilar las investigaciones y avances de un campo específico para mantener a los expertos informados y actualizados. Los artículos “*An Overview of Human DNA Analysis in Forensics*” y “*Técnicas de análisis del ADN en genética forense*” traducidos en este proyecto son ejemplos de este tipo. A diferencia de la variabilidad léxica especializada entre un tipo de publicación y otra, la estructura del sintagma nominal es comparable en ambas, y en este proyecto de análisis no se establecen diferencias sintácticas entre los artículos.

Esta investigación traductológica se desarrolla desde una perspectiva lingüística, dado que los mecanismos de ajuste o desviación de la norma estructural por parte de los traductores hispanohablantes, en este caso, corresponden a fenómenos del bilingüismo. Este fundamento lingüístico se justifica con el fin de alcanzar una perspectiva científica. Los estudios de traducción en sus orígenes se caracterizaron por la subjetividad, la falta de métodos de investigación, y la especulación. Con el surgimiento del concepto de la traducción como una ciencia, las teorías adquirieron mayor objetividad y se comenzaron a formular sobre un fundamento más investigativo.

La visión de la traducción como disciplina o arte puede abordarse desde una perspectiva científica sin afectar los aspectos estéticos, culturales o técnicos de la temática. Con la formulación de hipótesis, el desarrollo de experimentos, la

interpretación objetiva de datos y la comprobación de las teorías, la traducción deja de ser un estudio basado en opiniones, y se despoja de su estereotipo de oficio o técnica.

Según Catford (1), “any theory of translation must draw upon a theory of language.” Si bien el fundamento del proceso de traducción es de naturaleza lingüística, no se puede reducir este estudio a ser tan sólo una rama de la lingüística aplicada o de la lingüística comparada. Los fenómenos artísticos, socio-culturales y filosóficos llevan esta ciencia más allá de un simple análisis de diferencias lingüísticas entre una lengua y otra. Sin embargo, lo anterior no significa que este tipo de estudios se deban soslayar en beneficio de otros nuevos campos de investigación. El análisis descriptivo y comparado de las lenguas permite examinar con detalle las causas de algunos fenómenos traductológicos tales como la interferencia lingüística en el desempeño del traductor, la hipercorrección al traducir a una segunda lengua y el desarrollo de la competencia estilística estructural según el género y la tipología textual.

Lingüística descriptiva y constituyentes inmediatos

El origen de la lingüística moderna se caracteriza por el surgimiento de la gramática comparada como una disciplina de la lingüística histórica. La comparación lingüística detallada permitió que los investigadores descubrieran las ramificaciones de las lenguas indoeuropeas y sus troncos comunes. A partir de allí, se desarrolló mucha actividad en torno a la descripción lingüística durante el siglo XIX. Sin embargo, antes de que la lingüística descriptiva se convirtiera en una rama de la

ciencia, primero tuvo que formar una base conceptual de acuerdo con el avance de otros campos de estudio (Gleason 39).

Aunque la lingüística descriptiva moderna se fundamentó en el estudio del fonema, investigadores estadounidenses como Edward Sapir (1921) desarrollaron la idea del morfema, reconocieron la individualidad estructural de los idiomas y negaron la relevancia del significado dentro del análisis lingüístico (Gleason 44). Los trabajos de Sapir fueron retomados y enriquecidos por Leonard Bloomfield, considerado como el gran sintetizador de la lingüística en América (Gleason 46). Entre sus grandes aportes, el más importante para efectos de este proyecto es la caracterización sintáctica publicada en el libro *Language*.

En esta obra, Bloomfield divide las oraciones en dos partes, y repite el proceso sucesivamente hasta separar todos los constituyentes de la oración. El análisis no se basa en los constituyentes aislados, sino en los pares estructurales de constituyentes obtenidos a partir de la división sucesiva de la oración. Estos forman niveles estructurales, en los cuales se basa Bloomfield para formular su enfoque descriptivo. Según este autor, “any complex form (ranging from a long sentence to a multielement word) could be described in terms of its constituents and the grammatical features by which these are joined” (citado en Gleason 52). Con esta perspectiva, Bloomfield introduce la noción de nivel estructural, y a los constituyentes de un mismo nivel los denomina “constituyentes inmediatos”.

El enfoque de Bloomfield se utilizó en el análisis de este trabajo, dado que no es prescriptivo, y se ajusta al estudio sincrónico de las lenguas. Además, no se basa en reglas de transformación ni en el establecimiento de procesos o pasos para

generar las estructuras, sino que se limita a describir los patrones de formación estructural del idioma. Después de Bloomfield, los estudios estructuralistas en América se amplían con el aporte de algunos profesores estadounidenses como Charles Carpenter Fries (1952), quien introduce el concepto de “modificadores”, y relaciona el estudio de los constituyentes inmediatos con la descripción de la estructura y clase de las palabras, dejando de lado el significado de estas (Gleason 79-80).

La descripción del sintagma nominal del inglés ha sido desarrollada por autores como Gleason, Hockett, Francis y Roberts, entre otros, a partir del enfoque estructural propuesto por Bloomfield y el estructuralismo. El sintagma nominal forma parte de un tipo de construcciones denominadas por Hockett como endocéntricas (184), las que define como aquellas en las que la forma o clase del sintagma resultante es similar a la forma y clase de al menos uno de los constituyentes inmediatos de dicho sintagma. Un sintagma como “*young cat*”, que es un sustantivo singular, está compuesto de un adjetivo descriptivo y un sustantivo singular. El constituyente resultante (sustantivo singular) es similar a la forma y clase de uno de los constituyentes, representado por la palabra “cat”.

El constituyente cuya forma y clase calzan con los del constituyente resultante se denomina *núcleo*, y el otro constituyente es el *atributo*. Las construcciones endocéntricas que contienen un atributo son *atributivas* o *subordinadas*, y aquellas sin atributos se denominan *coordinadas* (Hockett 184-185). Este autor divide las construcciones endocéntricas coordinadas en tipos cuyos equivalentes en español son: copulativas, disyuntivas, apositivas, entre otras. En cuanto a las construcciones

atributivas del inglés, Hockett las clasifica en cuatro grupos principales (186), pero para el estudio del sintagma nominal sólo se retomarán tres:

- 1) Atributo antes del núcleo: *big tree*; *stone wall*
- 2) Núcleo al inicio: *number three*; (the) *book on the shelf*
- 3) Atributo discontinuo: (the) latest *volume to come out*; better *plan* than yours.

Las construcciones atributivas anidadas son expresiones complejas formadas por series de construcciones atributivas, insertas o anidadas una dentro de la otra (Hockett 188), y las lenguas difieren en cuanto a la complejidad de este tipo de frases. El anidamiento permite que se agreguen atributos en niveles más inclusivos respecto al núcleo. La siguiente figura muestra la forma en que Hockett esquematiza este tipo de construcciones para un sintagma nominal:

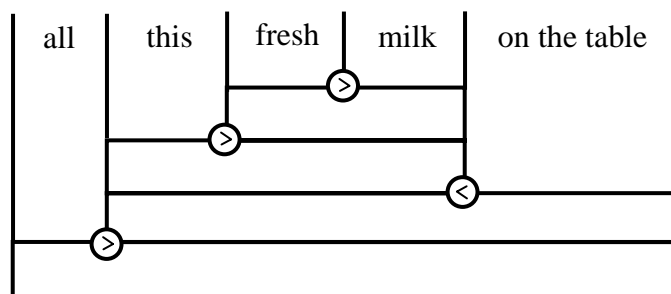


FIGURA 21.1 de Hockett (188)
Las marcas “>” y “<” se ubican en los puntos de unión de los
constituyentes inmediatos, e indican la dirección del atributo hacia la
posición del núcleo

En este diagrama de constituyentes, se observa que la sintaxis del inglés presenta una organización en varias capas formadas por pocos tipos de unidades básicas, tal como lo describe también Francis (293), quien desarrolló diagramas de constituyentes con la misma idea, pero distinta representación gráfica. Según

Hockett, las construcciones atributivas anidadas son cerradas: “*the process must stop: the nesting of constructions is closed. The whole form can then participate in still larger forms only by coordinate or exocentric constructions*” (189).

Roberts denomina a los grupos nominales como estructuras de modificación (189), formadas por el núcleo del sintagma más los modificadores. Roberts señala que los únicos modificadores que aparecen delante de los determinantes en inglés son “all” y “both” (193), tal como se observa en el ejemplo anterior de Hockett, e indica que el orden de las palabras en inglés es rígido en comparación con otros idiomas (205).

Un sintagma nominal se puede plantear, según Francis (311), de la forma:

determinante + adjetivo + sustantivo + **núcleo** + grupo preposicional + adverbio

en el que los elementos a la izquierda o a la derecha del núcleo se denominan modificadores. Como se verá en los siguientes capítulos, algunos autores introducen ligeras variantes a este orden propuesto, como en el caso de Gleason, que antepone el adverbio a las frases preposicionales.

El estudio de los atributos o modificadores del inglés ha sido retomado con detalle por profesores de lenguas y gramáticos como Quirk (1985) y Flores (2002), entre otros. Por ejemplo, entre los premodificadores, Quirk incluye a los predeterminantes (*all, both, double*), los determinantes centrales (artículos, posesivos), los postdeterminantes (adjetivos numerales), y la premodificación (adjetivos o frases adjetivas, los sustantivos). Entre los postmodificadores, Quirk incluye las frases

preposicionales, las cláusulas no finitas, las cláusulas relativas y la complementación. Por otro lado, Flores (2002) retoma otros estudios más detallados y dentro de los adjetivos incluye a los modificadores compuestos, y en el caso de los postmodificadores, incluye a un grupo de adjetivos denominados “true adjectives”, los cuales se postponen a los pronombres terminados en “body”, “one” y “thing”. Mientras tanto, un análisis exhaustivo de los modificadores del sintagma nominal en español se encuentra en la gramática descriptiva de Ignacio Bosque (1999), y a raíz de una comparación de ambos sistemas lingüísticos surge el análisis que provee una base teórica para confirmar la hipótesis de las interferencias lingüísticas estructurales presentadas en los siguientes capítulos.

Interferencia y bilingüismo

En términos generales, el lingüista Weinreich define a la interferencia como:

“Those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language, i.e. as a result of language contact” (1).

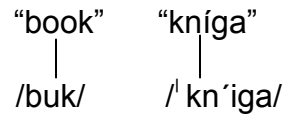
Según Weinreich, cuanto más difieran los sistemas lingüísticos entre sí, mayor será la cantidad de formas y patrones que se excluyan mutuamente en cada idioma, y mayor también será la dificultad de aprendizaje y el potencial para que ocurran las interferencias (1). Weinreich defiende a la lingüística descriptiva como la base teórica a partir de la cual se enuncian las formas de interferencia mutua de las lenguas en contacto, y señala que en la mayoría de los casos, las causas de los fenómenos específicos de interferencia se pueden determinar mediante métodos lingüísticos (3).

Este lingüista señala que una interferencia demasiado común entre los bilingües es aplicar el orden de las palabras de un idioma a la sintaxis de otra lengua, lo cual ocurre en un plano gramatical (37).

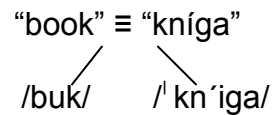
Weinreich incluye dos fuerzas opuestas dentro del fenómeno de la interferencia: el estímulo y la resistencia. Este lingüista señala que ambas pueden ser estructurales o no estructurales, y entre los criterios estructurales se encuentra la economía y la inteligibilidad lingüística (66). Algunos factores de estímulo son el valor social y el prestigio de la lengua, la diferencia o la congruencia entre los sistemas lingüísticos, el bilingüismo de los interlocutores, etc. Entre los factores de resistencia se incluye las actitudes puristas hacia una de las lenguas, lealtad hacia la lengua materna, entre otros (64-65).

En sus estudios sobre la interferencia lingüística, Weinreich se basa en la noción sassureana de “significante” y “significado”, y formula tres tipos de bilingüismo: el coordinado, el compuesto y el subordinado (9-10). El primero consiste en evocar el significado y los significantes de una lengua y otra en forma separada durante la comunicación interlingüística. El bilingüismo compuesto consiste en asociar el significado del signo en ambos idiomas, y luego evocar el significante en cada idioma por separado. Finalmente, el sistema subordinado implica evocar el significante de un idioma a partir del significante y el significado de un signo en otra lengua. Esto lo representa con un diagrama donde incluye el semantema “book” del inglés y su equivalente “kníga” en ruso:

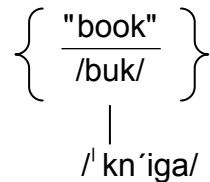
1) Bilingüismo coordinado:



2) Bilingüismo compuesto:



3) Bilingüismo subordinado:

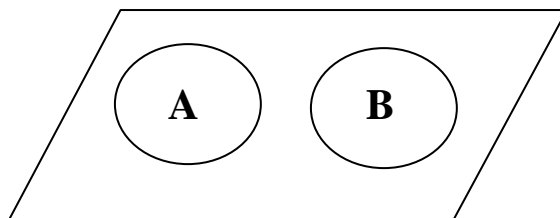


Aunque estos diagramas se aplican a semantemas, se podrían hacer diagramas para representar la codificación y descifrado de estructuras.

Osgood y Sebeok también dividen a los sistemas lingüísticos en compuestos y coordinados (196). Con base en esta división, este autor plantea dos modelos de traducción basados en la secuencia de estímulo, respuesta, codificación y decodificación, desde la perspectiva conductual. Indica que las situaciones de conducta son diferentes para un traductor si (a) mantiene lenguas compuestas o coordinadas, o si (b) la traducción es directa o inversa (200).

De Greve y Van Passel toman esta idea y representan mediante diagramas y esquemas de descifrado y codificación (137-142), una dicotomía en el campo del bilingüismo. Verifican, a partir de Osgood y Sebeok (1965), por un lado, la existencia de dos sistemas lingüísticos coordinados pero independientes y sin compenetración

recíproca, al que llaman “bilingüismo perfecto”, y lo representan de la siguiente manera:

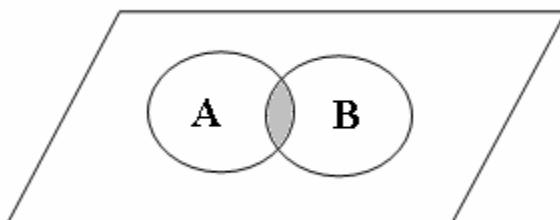


Con un esquema complementario de codificación:

Lengua A \longrightarrow Lengua B

donde A y B representan dos sistemas lingüísticos diferentes.

Por otro lado, estos lingüistas describen también una intersección de dos sistemas lingüísticos que provocan interferencias, de la siguiente manera:



y agregan el esquema de descifrado y codificación siguiente:

Lengua B \longrightarrow [Lengua A] \longrightarrow Lengua B

De Greve y Van Passel indican que:

“...casi en todos los casos existe interferencia entre los dos sistemas, o mejor aún, uno de ellos influye sobre el otro. Las circunstancias ideales que pueden eventualmente hacer posible la existencia de sistemas coordinados pero independientes se producen cuando se viven dos

campos lingüísticos unidos por analogías, o por los menos por afinidades «escenciales» (139).

Estos lingüistas proponen que las analogías o las afinidades entre los sistemas lingüísticos son condiciones ideales para favorecer la independencia y coordinación de los idiomas en el individuo bilingüe. A partir de este enunciado, se deduce que De Greve y Van Passel consideran que las semejanzas entre idiomas contribuyen a la ausencia de interferencias. Más adelante señalan que hay circunstancias particulares, como la presión de un sistema privilegiado sobre otro, que destruyen esa denominada “perfecta unidad analógica” y las interferencias se pueden hacer más manifiestas cuando esa presión es ejercida por una lengua de la misma familia lingüística que el idioma materno del individuo bilingüe (140). Estas ideas son tomadas de Osgood, sin embargo, este psicolingüista afirma claramente en su libro “Psicolingüística” que “incluso dentro de un sistema coordinado puede que se produzca la interferencia entre los dos conjuntos de procesos” (198). Además, él no atribuye la presión lingüística a la presencia de un sistema privilegiado sobre otro, sino a la presión mental que experimenta el bilingüe al interpretar los signos cuando se expone a una situación interlingüística, aunque el hablante tenga un sistema mental coordinado. En este caso Osgood y Sebeok afirman que:

“Cuanto más semejantes sean los signos... y cuanto más similares sean los mediadores –*significados*--, más fuerte será esta presión. Es más probable que se produzca la interferencia cuando las lenguas estén estrechamente relacionadas...” (198).

Osgood y Sebeok razonan a partir de las interpretaciones léxicas, sin embargo, esta base teórica abre el camino para discutir y plantear los mecanismos de la equivalencia estructural entre las lenguas, caso específico de este trabajo.

Otros lingüistas como Lado (1961), Odlin (1989) y Cook (1993) señalan que las diferencias entre los sistemas lingüísticos son los factores de mayor dificultad para el aprendizaje de una segunda lengua y contribuyen a la transferencia negativa. Lado afirma que “when the structural pattern is not paralleled in the native language, the student will have trouble learning it because of interference from the native language habit patterns” (146). Por otro lado, estos investigadores consideran a las semejanzas útiles para la transferencia positiva. Por ejemplo, Odlin señala que “*much of the influence of the native language (or some other previously learned language) can be very helpful, especially when the differences between two languages are relatively few*” (26), y considera que en este caso no se estaría hablando de interferencia, pues el término interferencia solamente incluye a la transferencia negativa: “*interference implies no more than what another term, negative transfer, does...*” (26). El lingüista Lado también respalda que las semejanzas entre los idiomas facilitan el aprendizaje, y señala que éste ocurre mediante transferencia. Sin embargo, incluye a la función, forma y distribución como parámetros para medir la semejanza.

Dentro del bilingüismo y el campo de la interferencia también ocurren fenómenos sociolingüísticos como es el caso de la hipercorrección y la sobregeneralización. Anteriormente se mencionó que el prestigio es un factor de estímulo para la interferencia. Aquellos individuos que evitan los errores en pro del

prestigio y estatus, podrían caer en otro grupo de interferencias. Por ejemplo, Odlin menciona a la hipercorrección como parte del fenómeno de transferencia negativa. Dentro de las consecuencias de la transferencia negativa menciona la baja producción, la sobreproducción, los errores de producción y la mala interpretación. La baja producción se refiere a una desviación de la norma en el idioma meta caracterizada por una baja frecuencia en el uso de una estructura, y funciona en forma contraria a la sobreproducción. Por otro lado, Odlin divide a los errores de producción en: sustituciones, calcos, y alteraciones de estructuras. Las sustituciones implican el uso exacto de la estructura de una lengua en otra, mientras que los calcos son aproximaciones muy cercanas del mismo caso, dentro de los cuales se pueden incluir algunos errores en el orden de las palabras. En cuanto a las alteraciones estructurales, Odlin menciona a la hipercorrección y la define aproximadamente como reacciones exageradas hacia una influencia particular del idioma nativo. Aclara también que otras alteraciones similares a la hipercorrección no reflejan ninguna influencia directa de la lengua materna (38). Otro fenómeno asociado es el de la sobregeneralización de las reglas del segundo idioma. Con el afán de evitar errores, algunos hablantes podrían aplicar determinadas reglas en forma generalizada, y no contemplar las excepciones o casos particulares. Estos aspectos sociolingüísticos también son parte del fenómeno de interferencia y se contemplan en el desarrollo de este proyecto.

CAPÍTULO 2

DESCRIPCIÓN Y COMPARACIÓN LINGÜÍSTICA DE LOS SINTAGMAS NOMINALES DEL INGLÉS Y EL ESPAÑOL

Este capítulo se centra en la descripción y comparación lingüística (inglés-español) de los sintagmas nominales típicos de los textos científicos; con base en ese estudio, en el capítulo siguiente de este proyecto se establecen factores de interferencia lingüística aplicados a la traducción inversa de estos sintagmas, según su comparación sintáctica en ambos idiomas, si se toma en cuenta que los traductores sean hablantes nativos del español.

El sintagma nominal del género científico y técnico se diferencia de otros por su estructura anidada que forma largas cadenas de atributos, que generalmente constituyen estructuras de modificación complejas con escasa coordinación y yuxtaposición. Según Alcaraz (63), en el inglés profesional y académico, el sintagma nominal cumple una función tanto expositiva como descriptiva. Se utiliza para procesos, ordenaciones lógicas, explicación razonada de un fenómeno, evaluación objetiva de datos, entre otros. Por ejemplo:

(1) qPCR assay for the simultaneous quantification of human and mitochondrial DNA in forensic samples (Timken y otros 1)

En (1), el sintagma nominal cumple una función explicativa porque expone la aplicación o utilidad del ensayo qPCR. Por otro lado, la función descriptiva o especificativa a menudo se manifiesta mediante una cadena de atributos que anteceden al núcleo nominal:

(2) cross-species sequence homology investigations (Timken y otros 2)

En este caso, el núcleo nominal de todo el sintagma es el sustantivo plural “investigations;” el resto de elementos a su izquierda son atributos estructurados en niveles sintácticos denominados constituyentes inmediatos (CI) que se ilustrarán posteriormente.

En inglés, los sintagmas nominales complejos se caracterizan por una larga premodificación. Por otro lado, en español se caracterizan por la ausencia de premodificación compleja y larga postmodificación. Este tipo de construcciones permiten que los expertos establezcan de una manera condensada y concisa, descripciones y explicaciones que implican enumeración exhaustiva de condiciones o características. Los traductores deben tener en cuenta que este tipo de sintagmas representa un rasgo estilístico propio de los textos científicos en inglés, y por lo tanto podrían tratar de incorporarlo estilísticamente a sus traducciones, como parte de un proceso para lograr la aceptabilidad textual. A continuación se describen algunos tipos comunes de sintagmas nominales del inglés y del español, y se comparan para observar estructuras sintácticas inversas, semejantes o diferentes entre ambas lenguas.

Tipo de estructura del sintagma nominal:

Determinante+premodificador(es)+(núcleo)+postmodificador(es)

Los modificadores del sintagma pueden ser sustantivos, adjetivos, adverbios, grupos o frases preposicionales, grupos subordinados, entre otros.

Ejemplos¹:

INGLÉS:

A continuación se plantean los tipos básicos de construcciones endocéntricas atributivas del inglés científico, las cuales se comparan luego con las del español. Se aclara que este análisis no es exhaustivo, sino más bien trata de representar las construcciones más comunes, según el corpus de este proyecto. Estos tipos básicos de construcciones se describen lingüísticamente, para luego proceder a los sintagmas más complejos que se basan en estas construcciones. Este estudio le servirá al traductor para comprender los fundamentos sobre los constituyentes

¹ Los ejemplos a continuación se ajustan a la siguiente nomenclatura:

A) El núcleo del sintagma se indica en **negrita**.

B) Los puntos suspensivos (...) indican la posible continuación del sintagma.

C) El término “frase” incluye las frases nominales, adjetivas, adverbiales, preposicionales, de participio, de gerundio y de infinitivo. Aunque son menos comunes, el término “frase” también puede aplicarse a casos de subordinación. Sin embargo, este análisis no se centra en este aspecto.

D) Los participios, aunque funcionen como adjetivos, se incluyen por su nombre en los diagramas.

Abreviaturas: **sust:** sustantivo; **adj:** adjetivo o adjetival; **adv:** adverbio o adverbial; **verb:** verbo o verbal; **prep:** preposición o preposicional; **particip:** participio; **infinit:** infinitivo(a); **conj:** conjunción; **cons. coord:** construcción coordinada; **determ:** determinante; **modif. comp:** modificador compuesto; **frase nom:** frase nominal.

inmediatos y para explorar las combinaciones atributivas posibles y aceptadas en ambas lenguas.

Construcciones atributivas de premodificación

Constituyentes inmediatos básicos:

A) Modificadores simples:

(4) ...sust+(núcleo): DNA **extracts** (Timken y otros 2)

En este caso, la construcción del sintagma no presenta dificultad porque el núcleo sólo cuenta con un modificador o atributo.

(5) ...sust+sust+(núcleo): TH01* target **sequence** (Timken y otros 2)

*TH01: Designa a un tipo específico de gen.

Dada la presencia de más de un premodificador, es necesario establecer niveles de CI para averiguar si TH01 modifica sólo al sustantivo “target”, o si modifica más bien al sintagma “target sequence”. El análisis muestra que los constituyentes “target sequence” encierran un solo concepto y por lo tanto son constituyentes inmediatos. Si se realiza un diagrama similar al de Gleason (197), se puede representar los niveles de construcción atributiva con el objetivo de interpretar su significado.

TH01 target **sequence**


```
graph TD; A[TH01 target sequence] --- B[TH01]; A --- C[target sequence]; C --- D[target]; C --- E[sequence];
```

Con este ejemplo sencillo se muestra que el modificador más inclusivo se coloca más cercano al núcleo del sintagma, mientras que el más global ocupa un nicho más hacia la derecha según el número de constituyentes que modifique.


(6) ...adj+sust+(núcleo): low room-**temperature** (Liu y otros 1).

Si no hubiera un guión entre “room” y “temperature”, y el sintagma apareciera asilado, el ejemplo podría utilizarse para dos significados distintos: la temperatura corresponde a un cuarto ubicado en una planta baja, o se refiere a una temperatura ambiental baja.

Para el primer caso: low room **temperature**



Para el segundo caso: low room-**temperature**



Una posible solución para resolver esta ambigüedad sería utilizar un guión para unir “low-room” en el primer caso.

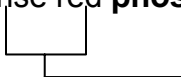
(7) ...adj+(núcleo): red **phosphorescence** (Marriot y otros 4)

(8) ...adj+adj+(núcleo): intense red **phosphorescence** (Peter y otros 3)

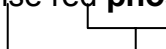
Mientras (7) es un tipo básico de construcción atributiva, (8) requiere determinar los niveles de constituyentes para evitar cambios de significado. Se debe

discriminar si la fosforescencia es de un rojo intenso, o si la fosforescencia roja es intensa:

Para el primer caso: intense red **phosphorescence**



Para el segundo caso: intense red **phosphorescence**



El traductor al inglés debe ser conciente de estos niveles sintácticos. De lo contrario, cuando los sintagmas sean más complejos, se podría alterar el orden esperado de los atributos y como consecuencia, se generaría un significado completamente distinto.

(9) ...adv+adj+(núcleo): highly degraded **samples** (Timken y otros 9)

En este tipo de construcciones, el adverbio y el adjetivo forman un solo constituyente inmediato que luego modifica al núcleo. El adverbio nunca puede aparecer como modificador de un núcleo nominal.

(10) ...particip+(núcleo): flanking **regions** (Timken y otros 4);

delayed **CT** (Timken y otros 6)

Como se verá más adelante, también los participios se pueden posponer según el caso.

B) Construcciones coordinadas

Estas construcciones a menudo se presentan como atributos que modifican al núcleo. Las coordinaciones copulativas a menudo se sintetizan para formar modificadores compuestos que se explican en el punto C. Este tipo de construcciones se combinan con otros atributos para formar sintagmas complejos.

Ejemplos:

(11) **Copulativas:** “__and__”: the HVI and HVII **portions** (Timken y otros 3)

En muchos casos, los dos atributos no necesitan un orden preferencial y se pueden intercambiar de posición. Sin embargo, en otras situaciones los atributos presentan un orden lógico.

(12) **Disyuntivas:** “__or__”: known or deduced **primers** (Timken y otros 3)

(13) **Apositivas** “__ __”: TH01 **gene** (Timken y otros 3)

C) Modificadores compuestos léxicos

Tipos comunes:

adj-adj; adv-adj; sust-adj; adj-sust; sust-sust; frase nominal; frase adj.; frase adv.; frase prep.; frase verb.; frase coord.; frase conj.;(adj, sust) + adj deverbal.

Algunos ejemplos: (subrayados)

(14) ~170-190 bp **sequence** (Timken y otros 1)

(15) non-root hair **shafts** (Timken y otros 2)

- (16) DOS-based **program** (Timken y otros 4)
- (17) real-time qPCR **assay** (Timken y otros 1)
- (18) end-point PCR **assays** (Timken y otros 1)
- (19) state-of-the-art **chemistry** (Wait 5)

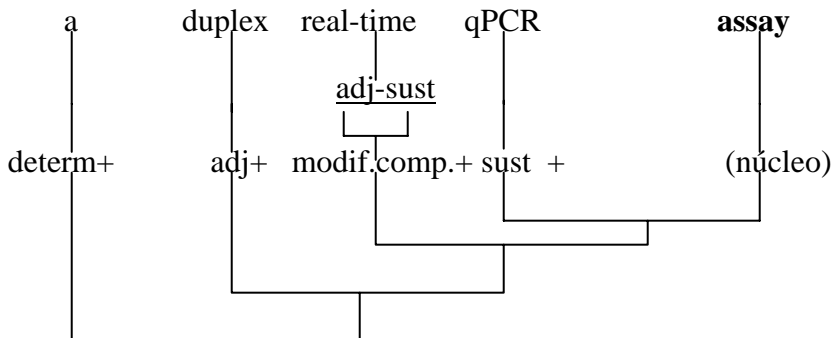
Estructuras complejas de premodificación:

Estas estructuras presentan una combinación de varios constituyentes inmediatos superpuestos en varios niveles estructurales. Para el traductor hispanohablante, la construcción de estas estructuras debe hacerse con base en estas nociones lógicas de constituyentes inmediatos, con el objetivo de evitar ambigüedad y sin sentido.

Ejemplos:

Se utiliza un diagrama de constituyentes inmediatos similar (en parte) a los propuestos por Gleason para la lingüística descriptiva.

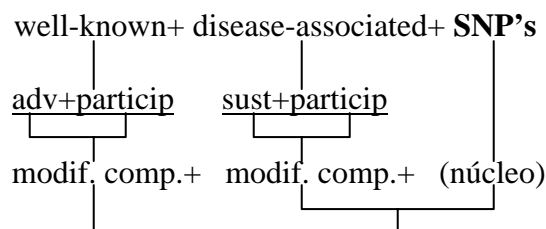
- (20) a duplex real-time qPCR* **assay** (Timken y otros 1)



Sigla del término fraseológico: qPCR: “quantitative Polymerase Chain Reaction.”

En este ejemplo se muestran los distintos niveles de constituyentes inmediatos que forman parte del sintagma. Los constituyentes o atributos que presentan una relación más cercana al núcleo, es decir, los más inclusivos, se colocan en los niveles más bajos, mientras que los atributos más generales se van anidando hacia la izquierda en niveles más altos.

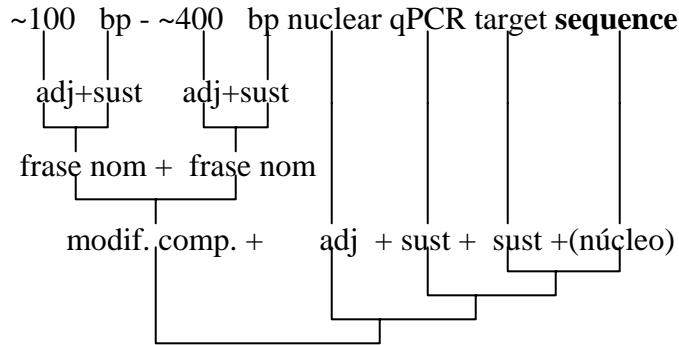
(21) well-known disease-associated **SNP's** (Timken y otros 5)



*Sigla del término fraseológico: SNP: Single nucleotide polymorphism.

El ejemplo complejo anterior presenta dos grupos adyacentes de modificadores compuestos, formados a su vez por un par de constituyentes inmediatos separados por un guión, el cual indica asociación inmediata. Este tipo de modificadores constituyen una forma efectiva del inglés científico para economizar las expresiones y simplificar la estructura de los atributos en una sola cadena premodificadora, en vez de recurrir a la subordinación.

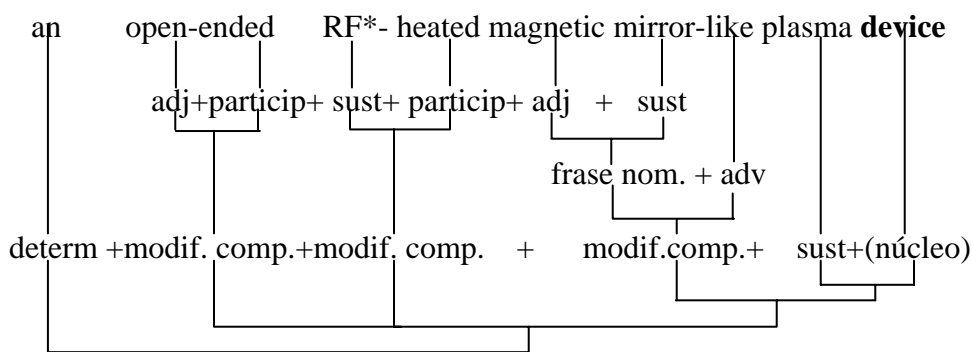
(22) ~100 bp* - ~400 bp nuclear qPCR target **sequence** (Timken y otros 12)



*Sigla del término fraseológico: bp: base pair.

En este ejemplo también se muestra que los atributos premodificadores pueden estar formados por otros constituyentes inmediatos que es necesario construir e interpretar. Tal como ha sido descrito por los lingüistas mencionados en el marco teórico de este proyecto, la estructura de atributos anidados de estos ejemplos de sintagmas nominales complejos no se presentan como una cadena lineal de modificadores, sino como niveles sintácticos superpuestos. Como se mostró en los ejemplos anteriores sobre los tipos básicos de construcciones, los niveles sintácticos pueden complicarse cuando se incluyen atributos que se prestan para ambigüedad. A continuación se muestra un ejemplo de esta situación en un sintagma nominal complejo, tomado de un artículo técnico sobre cohetes de plasma:

(23) an open-ended RF-heated magnetic mirror-like plasma **device** (Chang y otros 1)



*Término fraseológico: RF: Radio Frequency.

Este sintagma nominal complejo presenta atributos que modifican a otros atributos en niveles distintos. Es posible que un lector no experto realice asociaciones de constituyentes equivocadas y malinterprete el sintagma. Para un traductor hispanohablante nativo no experto en el tema, esta construcción tan sintética podría ser difícil de alcanzar debido a que supone una organización de premodificadores compuestos anidados y separados por un guión, los cuales representan una estructura sintáctica muy distinta a la del español.

En ese ejemplo también se observa la formación de una frase nominal, que junto a un adverbio, conforma un modificador compuesto: “magnetic mirror-like”. Nótese que un lector no experto podría asociar el modificador “mirror-like” con el sustantivo “plasma”, y pensar que el adjetivo “magnetic” engloba al resto del

sintagma de izquierda a derecha. De ser así, se interpretaría que el plasma es de tipo espejo, y que el dispositivo es magnético. En realidad, es el espejo el que es magnético, y más bien, el dispositivo funciona con el plasma. Este ejemplo demuestra la necesidad de identificar correctamente los niveles sintácticos estructurales.

Construcciones atributivas de postmodificadores

A continuación se describirá brevemente la parte postmodificadora compleja del sintagma nominal del inglés. No sólo se incluyen las estructuras de postmodificación posibles en la lengua, sino también aquellas que no pertenecen a la gramática de este idioma, sino al español. Estas estructuras se señalan con el símbolo \emptyset , que indica ausencia, y se incluyen con el objetivo de utilizarlas más adelante en un cuadro comparativo de las estructuras de ambas lenguas.

Constituyentes inmediatos básicos:

A) Modificadores simples:

Ejemplos:

- (24) (núcleo)+sust= \emptyset
- (25) (núcleo)+adj...: **concentrations** identified (Timken y otros 5)
- (26) (núcleo)+adj+adj= \emptyset
- (27) (núcleo)+adj+adv...: chemical **intermediates** used widely in the manufacture of dyes... (IARC 1)

- (28) (núcleo)+adv+adj+frase...: **method** widely used in biomedical research and molecular diagnostics (Timken y otros 1)
- (29) (núcleo)+adj+frase...: **studies** performed on an instrument (Timken y otros 1)
- (30) (núcleo)+frase...: **DNA** to be quantified (Timken y otros 12)
- (31) (núcleo)+frase +frase...: **proportion** of undetected alleles at STR loci (Timken y otros 1)
- (32) (núcleo)+cláusula relativa...: ...**amplification** that spans the TH01 STR-locus repeat region (Timken y otros 1)

B) Construcciones coordinadas

- (33) (núcleo)+adj+conjunción+adj...: **Information** obtained and reviewed (Minnesota Planning 10)
- (34) (núcleo)+apositivo: sequence **number** 96 (Timken y otros 5)

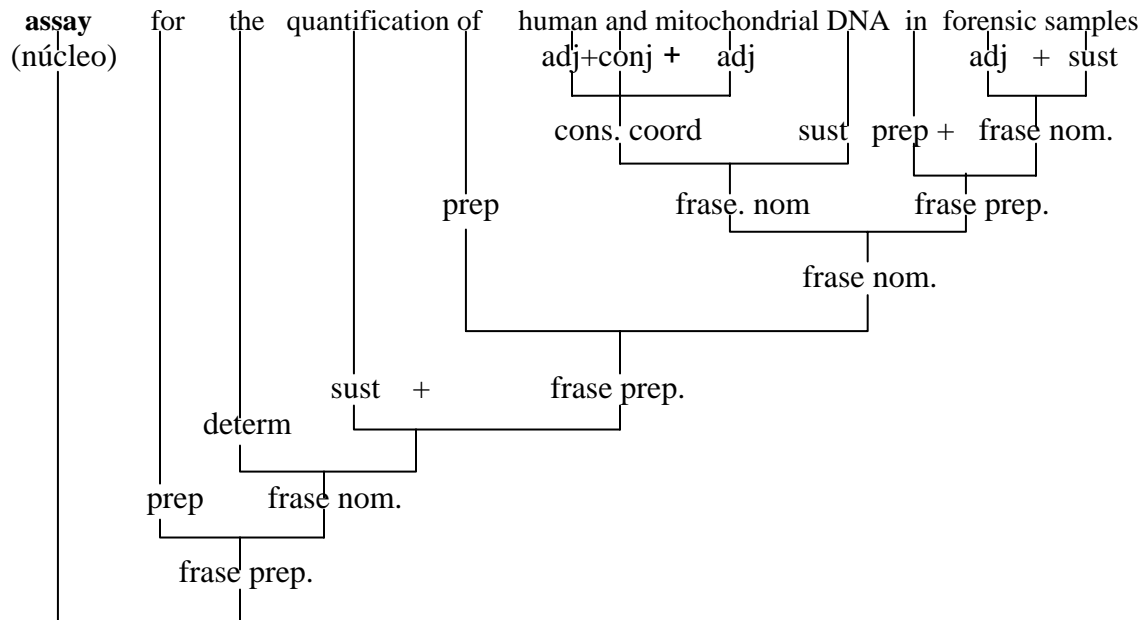
C) Modificadores compuestos léxicos (separados por guión)=∅

Este tipo de modificadores no se suelen posponer al núcleo nominal.

Estructuras complejas de postmodificación:

A continuación se cita solamente un ejemplo de muestra para representar una estructura de larga postmodificación:

(35) ... **assay** for the quantification of human and mitochondrial DNA in forensic samples (Timken y otros 1)



El ejemplo anterior es una muestra de que las construcciones postmodificadoras del inglés tienden a estar formadas por frases preposicionales que se anidan de derecha a izquierda, desde la menos relacionada con el núcleo hasta la más inclusiva.

ESPAÑOL

Construcciones atributivas de premodificadores

En español, las construcciones atributivas premodificadoras no suelen ser complejas. En este idioma no se forman cadenas de atributos anidados delante del

núcleo, tal como sucede con preferencia en el inglés técnico-científico. A continuación se presenta un examen breve sobre algunas construcciones del inglés que están ausentes en español, dentro del género textual discutido.

Constituyentes inmediatos básicos

A) Atributos simples

(36) sust+(núcleo)= ∅

(37) adj+(núcleo): No es común en el género científico, excepto cuando el adjetivo es numeral. Ejemplo:

(35.1) tres **fases** (Entrala 7).

En el género literario suele aparecer con frecuencia como epíteto:

(35.2) ...con la leve **sonrisa**, con la luz que ha borrado, como un río de desechos, aquellos viejos **signos**. (Monge 120).

Ocasionalmente se construye algún sintagma como este:

(35.3) altas **temperaturas** (Entrala 17).

Nótese que este último ejemplo corresponde a un subtipo textual típico de la macroestructura de los movimientos de introducción o discusión de los textos científicos, y no se forman con expresiones de lenguaje especializado. Por ejemplo, no se acostumbra decir:

(35.4) Los descartados **ensayos**, o los repetitivos **datos**, etc.

(36) adj+adj+(núcleo): Similar a (35), tienen frecuencia escasa en el género científico, podría ocurrir en presencia de adjetivos numerales, junto con algún adjetivo que se debe resaltar (de función especificativa). Ejemplo:

(36.1) dos únicas **bandas** (Penacino 1)

(37) adv+adj+(núcleo): (Poco común)

Ejemplo: poco controlables **diferencias** (Penacino 1)

B) Construcciones coordinadas

(38) ...adj+conj. coordinante+adj+(núcleo): cuarto y último **paso** (Díaz y otros 7).

Este tipo de estructura aparece de forma ocasional en los textos científicos y no se utiliza con frecuencia en las descripciones o explicaciones que incluyen lenguaje especializado.

C) Modificadores compuestos léxicos (separados por guión)=∅. En español no se utilizan estos modificadores antes del núcleo.

Estructuras complejas de premodificación:

No son comunes en el género científico, y suelen ocurrir en el género literario en forma coordinada y yuxtapuesta. Ejemplo:

(39) La dolorosa y desesperante fría sala de hospital.

Construcciones atributivas de postmodificadores

En los textos de español científico, este tipo de construcciones son ideales para describir y explicar procesos. En contraste con el inglés, la mayoría de los sintagmas nominales utilizados en el lenguaje técnico y científico están formados por un núcleo seguido por postmodificadores, que pueden ser frases preposicionales, adjetivas, de participio, entre otros. Los grupos postmodificadores de subordinación no se estudiarán en este proyecto. A continuación se describirán en forma general algunos constituyentes inmediatos postmodificadores básicos del español con el objetivo de contrastarlos con el inglés y comparar luego las construcciones complejas de sintagmas nominales en ambos idiomas.

Constituyentes inmediatos básicos:

Ejemplos:

A) Modificadores simples:

(40) (núcleo)+sust...: En textos científicos escritos por autores que presentan interferencia del inglés, a menudo se encuentran este tipo de

construcciones, las cuales se incorporan al lenguaje técnico y científico como anglicismos de estructura. Incluso, algunos de estos anglicismos están normalizados internacionalmente y se utilizan, aunque no coincidan con la lengua receptora. Ejemplo:

(40.1) **ADN** molde (Entrala 7).

En este caso, molde no es un tipo específico de ADN, y por lo tanto, no es un apositivo. El sintagma que se ajusta a la lengua sería: **ADN** de molde.

(41) (núcleo)+adj... : **bases** pirimidínicas (Entrala 10).

(42) (núcleo)+adj+adj...: **ADN** nuclear inespecífico (Entrala 4).

(43) (núcleo)+adj+adv...: **gel** aplicado posteriormente (Timken y otros 31)

(Traducción)

(44) (núcleo)+adv+adj...: **muestras** moderadamente degradadas (Timken y otros 30) (Traducción).

La construcción (44) no es muy común en español porque más bien se utiliza el adverbio pospuesto al adjetivo. Sin embargo, el lenguaje científico la ha adoptado como un anglicismo de estructura.

(45) (núcleo)+adj+frase...: **ADN** mitocondrial en muestras forenses (Timken y otros 7) (Traducción)

(46) (núcleo)+adj+frase+frase...: **ADN** genómico de alto peso molecular para la cuantificación (Timken y otros 16) (Traducción).

(47) (núcleo)+frase...: **Hibridación** de la sonda marcada (Entrala 4).

(48) (núcleo)+frase +frase...: **Métodos** para cuantificar el ADN en la genotipificación (Timken y otros 32) (Traducción).

(49) (núcleo)+cláusula relativa...: el **ADN** que se utiliza (Entrala 1).

B) Construcciones coordinadas

Ejemplos:

(50) **Copulativas:** “__y__”; “__ni__”: **Sonda** marcada y desnaturalizada (Entrala 4).

(51) **Disyuntivas:** “__o__” : **ensayos** de qPCR uniplex o dúplex (Timken y otros 21) (Traducción).

(52) **Apositivas** “__ __”: Los **marcadores** ABI Quantifiler™ (Timken y otros 1) (Traducción).

C) Modificadores compuestos (separados por guión): Estos modificadores se utilizan frecuentemente en el género científico. Ejemplo:

(53) **Análisis** clínico-genético (Saria 1)

Estructuras complejas de postmodificación:

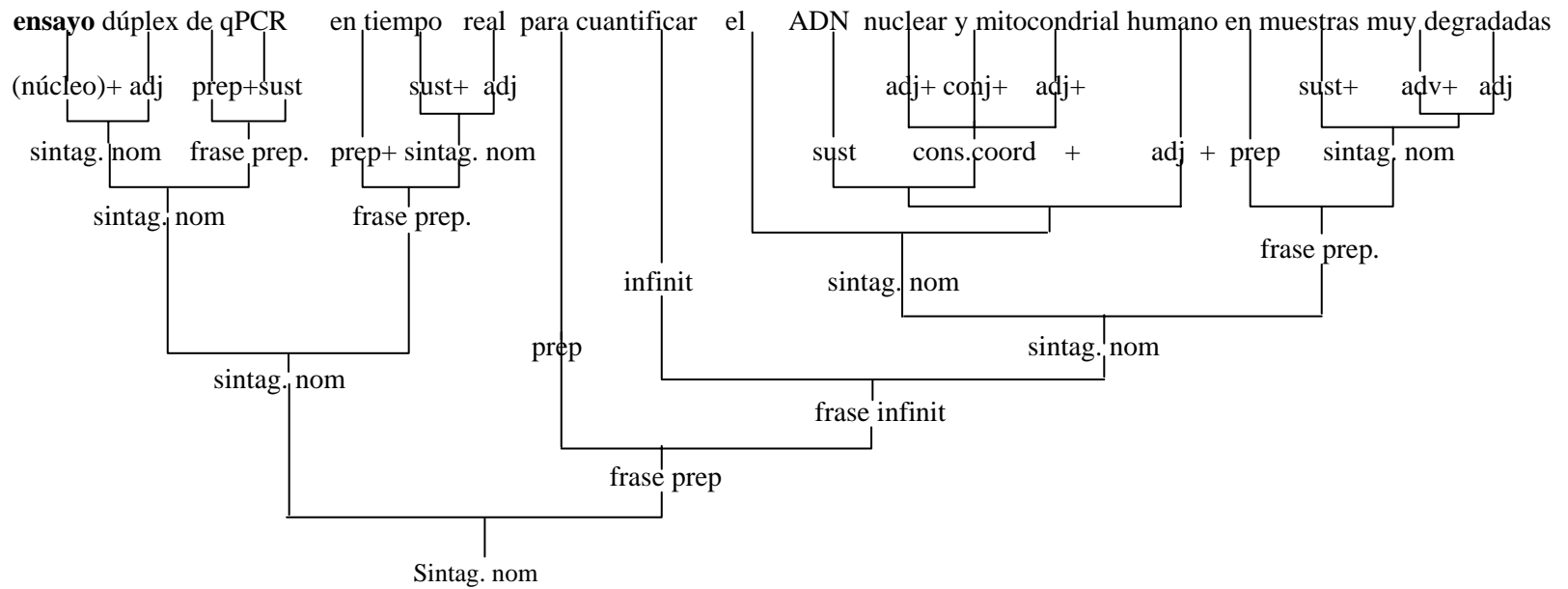
Como se indicó anteriormente, el español presenta preferencia por las construcciones de postmodificación, y este análisis las describe detalladamente para compararlas con las construcciones del inglés.

Ejemplo:

A continuación se muestra que el núcleo de un sintagma es seguido por una serie de tres niveles sintácticos que crecen progresivamente de izquierda a derecha según su menor especificidad respecto al núcleo, de manera que el nivel más alto es el que engloba aspectos más generales. Si se compara esta estructura compleja con los ejemplos en inglés, se observa que los sintagmas postmodificadores presentan formas similares, pero la información que se incluye no es igual. Es decir, al traducir este sintagma al inglés, una parte se debería construir en forma de atributos premodificadores. Las diferencias de las estructuras postmodificadoras en ambos idiomas causan interferencia lingüística sintáctica, la cual se debe también a la ausencia de premodificación compleja en español, y al orden variable de los atributos según su categoría en el sintagma de ambos idiomas.

Diagrama de construcción atributiva de postmodificadores en español

(54) ensayo dúplex de qPCR en tiempo real para cuantificar el ADN nuclear y mitocondrial humano en muestras muy degradadas (Timken y otros 1) (Traducción)



A continuación se muestran dos cuadros comparativos de las estructuras del sintagma nominal del inglés y el español.

Cuadro 1. Estructuras de premodificación²

Tipo de estructura	Lengua		Correspondencia funcional	Observaciones
	Inglés	Español		
sust+(núcleo)	X	∅	NA	
adj+(núcleo)	X	X*	≠	*Adj numeral, cuantificador, en función explicativa o epíteto.
adj+adj+(núcleo)	X	X*	≠	*Adj numeral, cuantificador, en función explicativa o epíteto.
adv+adj+(núcleo)	X	X*	≠	*En función explicativa o epíteto
particip+(núcleo)	X	X*	≠	*En función explicativa o epíteto.
construcciones coordinadas +(núcleo)	X	X*	≠	*En función explicativa o epíteto.
modificadores compuestos léxicos	X	∅	NA	
estructuras complejas de premodificación	X	X*	≠	*Acompañados de yuxtaposición y coordinación, en textos literarios.

² **Simbología:**

X: Estructura presente en la lengua

X*: Estructura presente pero con uso restringido

≠: La función de la misma estructura en ambos idiomas es distinta.

≈: La función de la misma estructura en ambos idiomas es similar porque coincide en algunos casos.

=: La función es equivalente.

∅: Estructura ausente en la lengua

NA: No aplica

Cuadro 2. Estructuras de postmodificación

Tipo de estructura	Lengua		Correspondencia funcional	Observaciones
	Inglés	Español		
(núcleo)+sust	X*	X*	≈	*Apositivos
(núcleo) +adj	X*	X	≠	*Incluye adjetivo en función de participio o en función explicativa.
(núcleo) +adj+adj	∅*	X	NA	*Excepto cuando se utiliza coordinación o yuxtaposición
(núcleo) + adj +adv	X*	X	≠	*Participio en función de adjetivo, suele continuar con frases prep.
(núcleo) + adv + adj	X	X*	≠	En inglés suele continuar con frases prep. *Poco común, adoptado como anglicismo de estructura.
(núcleo)+adj+frase	X*	X	≠	*Adjetivo en función de participio o en función explicativa.
(núcleo)+frase	X	X	≈	Pueden diferir en la función comunicativa
(núcleo)+frase+ frase	X	X	≈	Pueden diferir en la función comunicativa
(núcleo)+cláusula	X	X		Presentan distinta frecuencia. En inglés son cuatro veces menos frecuentes respecto a la forma: (núcleo)+frase (Quirk 1274).
(núcleo) + construcciones coordinadas	X	X	=	
(núcleo) +modificadores compuestos léxicos	∅	X	NA	
estructuras complejas de postmodificación	X	X	≈*	* La postmodificación en inglés tiende a buscar la explicitación (Quirk et.al 1243), mientras que en español es de uso generalizado.

En el Cuadro 1 se observa que las estructuras premodificadoras del español ocurren con poca frecuencia en el género científico, y sobre todo se reservan para funciones explicativas, o epítetos típicos de los textos literarios. A diferencia de los resultados del Cuadro 1, el Cuadro 2 muestra que todas las construcciones postmodificadoras del inglés son compartidas por el español. A partir de estos datos se deduce que las semejanzas estructurales entre ambos idiomas no necesariamente implican una correspondencia funcional. Algunas de esas funciones son la simplificación de construcciones pasivas, y la función explicativa relacionada con el enfoque y el valor comunicativo del sintagma nominal, su uso restrictivo y no restrictivo (Quirk et.al 1281).

La traducción inversa del sintagma nominal del inglés depende no sólo de la construcción de estructuras válidas en la lengua, sino también del estilo del género textual, y la función comunicativa que cumple el sintagma según el contexto. Tal como se muestra en los ejemplos del análisis, en el género científico el sintagma nominal del inglés suele construirse con premodificaciones largas y complejas, y postmodificaciones complejas de frases preposicionales. Esta característica no aplica a otros textos como los literarios, o los textos convencionales utilizados para lectores no especializados. En el próximo capítulo se analizarán las interferencias lingüísticas desde el punto de vista causal. Se describirá y ejemplificará los tipos de interferencias, y se explicarán los rasgos estructurales y funcionales que el traductor hispanohablante podría adoptar como resultado de este fenómeno lingüístico.

CAPÍTULO 3

FACTORES DE INTERFERENCIA LINGÜÍSTICA EN LA TRADUCCIÓN INVERSA (ESPAÑOL-INGLÉS) DEL SINTAGMA NOMINAL

El grado de interferencia lingüística de quien traduce a una segunda lengua depende de la competencia y el desempeño lingüístico, entre otros factores. El presente trabajo no pretende evaluar los distintos grados de interferencia lingüística en las traducciones, sino revelar algunas áreas donde se pueden producir interferencias sintácticas sutiles durante el proceso de traducción inversa. Estas interferencias se identifican como sutiles porque no corresponden a los errores sintácticos comunes que cometen los hispanohablantes nativos durante el proceso de aprendizaje de una segunda lengua, los cuales se suelen corregir conforme el estudiante avanza de nivel y logra asimilar o interiorizar algunas reglas. Las interferencias sintácticas que se tratan en este análisis se observan en hispanohablantes cuyo nivel de competencia lingüística en el idioma inglés es avanzado y traducen aceptablemente a su segunda lengua. Estas interferencias no necesariamente corresponden a errores sintácticos, sino a errores de uso y función que causan debilidad estilística sintáctica que podría afectar la aceptabilidad del texto meta.

Los datos que sustentan este análisis provienen, por un lado, de la observación continua de trabajos de traducción inversa realizados por estudiantes de posgrado en traducción, y por otro, de los resultados de un proceso de retrotraducción así como de traducción inversa efectuados en este proyecto. Para

comprender mejor la importancia de la metodología, se debe tener en cuenta que la lengua materna de la traductora de este proyecto es el español, y el inglés es su segunda lengua.

Para realizar la retrotraducción se tomaron al azar varias muestras de sintagmas nominales complejos del corpus de los artículos científicos traducidos al español que se adjuntan en este proyecto. Estos se tradujeron de nuevo al inglés para comparar sus estructuras sintácticas con la de los sintagmas originales. Los ejemplos que se comentan a continuación son tomados del proceso de retrotraducción al inglés, así como del artículo que se tradujo al inglés como parte de la investigación. Además, se incluyen ejemplos de algunos textos paralelos o de informes científicos con temática distinta.

Como es de esperar, la traducción inversa que forma parte de este trabajo presenta sintagmas nominales traducidos de la manera más parecida a las construcciones sintácticas correspondientes al estilo del inglés científico estándar, y se evitó el uso de estructuras sintácticas que reflejaran las interferencias. Sin embargo, este esfuerzo fue conciente, y muchas veces, el proceso de traducción implicó replanteamiento sintáctico de los sintagmas para lograr un estilo más condensado, es decir, un estilo con precisión expresiva y economía comunicativa tal como lo señala Alcaraz en su descripción del inglés académico y profesional (62).

Como se expuso en el capítulo anterior, el sintagma nominal del inglés y el español difieren en cuanto a la posición de los modificadores respecto al núcleo. Mientras se observa que atributos como los adjetivos y sustantivos aparecen con mayor frecuencia en función premodificadora en inglés, en español más bien ocurren

en función postmodificadora, y con un orden inverso en la mayor parte de los casos. Por otro lado, el inglés también presenta estructuras postmodificadoras, las cuales suelen aparecer construidas por un adjetivo, una forma coordinada en función modificadora, y principalmente por frases preposicionales y cadenas anidadas de ellas.

En este proyecto se analizan dos causas de interferencia lingüística. La primera es que el tipo de construcción sintáctica del sintagma nominal del español y del inglés presenta orden inverso, de modo que las estructuras del idioma materno compiten con las del segundo idioma. La segunda razón es que el tipo de construcción sintáctica del inglés está ausente o presenta menor frecuencia en la lengua materna del traductor.

Los procesos de las interferencias sintácticas en la traducción son de tres tipos según el mecanismo que las origina. Primero, las interferencias ocurren al trasladar e incorporar una estructura de la lengua materna al sistema gramatical de la segunda lengua. Esto implica formar una gramática “nueva” que incluye algunos calcos sutiles del español, y a partir de aquí producir el texto meta. Este proceso suele ocurrir en presencia de estructuras de orden distinto o inverso entre los dos sistemas lingüísticos. El segundo tipo corresponde a la búsqueda de estructuras semejantes en ambos idiomas, mientras que el tercer tipo es la hipercorrección. Los dos últimos mecanismos están relacionados con la ausencia o menor frecuencia de estructuras del español respecto al inglés. Estas interferencias pueden resultar en construcciones que se ajustan a las normas sintácticas de la lengua, pero no necesariamente a la pragmática del texto.

El tipo de construcciones basadas en las semejanzas de ambos idiomas producen interferencias sutiles porque, como se ha visto, el hecho de que dos estructuras sintácticas existan en dos idiomas no significa que se utilizan con la misma función y efecto comunicativo. Por otro lado, el proceso hipercorrectivo no busca semejanzas sino que persigue lograr una construcción muy aproximada a la utilizada en el estándar de la segunda lengua, y como resultado se produce un uso sobregeneralizado de las reglas sintácticas. A continuación se describen las interferencias según las posibles causas que las originan.

INTERFERENCIAS LINGÜÍSTICAS CAUSADAS POR CONSTRUCCIONES SINTÁCTICAS DE ORDEN DISTINTO O INVERSO EN AMBOS IDIOMAS

Orden de los atributos de premodificación

Según la descripción sintáctica del capítulo 2, los atributos premodificadores del inglés por lo general (con excepción de adjetivos numerales, epítetos y otros casos específicos), se incluyen como atributos postmodificadores en español. Según el orden de ocupación de nichos “Slot-and-filler” propuesto por Gleason (140) para los determinantes y atributos, el inglés establece el siguiente orden:

Determinante+numeral+adjetivo+sustantivo+ **(núcleo)**+adverbio+frase prep.

Ejemplo modificado de Gleason (140):

(55) the three old boxer **dogs** downstairs with a mean temper.
 | | | | | | | |
 determ + num+ adj + sust +(núcleo) + adv + frase prep.

Si la estructura (55) se compara con la de español, parecería que son similares, excepto por la ausencia de sustantivos en premodificación, la ausencia de adverbios como modificadores de sustantivos y la preferencia del español por colocar el adjetivo de forma postpuesta:

Determinante+numeral+adjetivo*+(**núcleo**) +adverbio + adjetivo+frase prep.

Ejemplo:

(56) las tres **familias** más ricas de Santo Domingo
 | | | | | | |
 determ + num + (núcleo) +adv +adj + frase prep

* Este nicho sólo se ocupa en casos especiales, en los que el adjetivo (no numeral) cumple una función explicativa o epíteto.

A diferencia del inglés, si se examina la posición del adverbio en español respecto a sus constituyentes inmediatos, este no modifica directamente al núcleo, sino al adjetivo que le sucede.

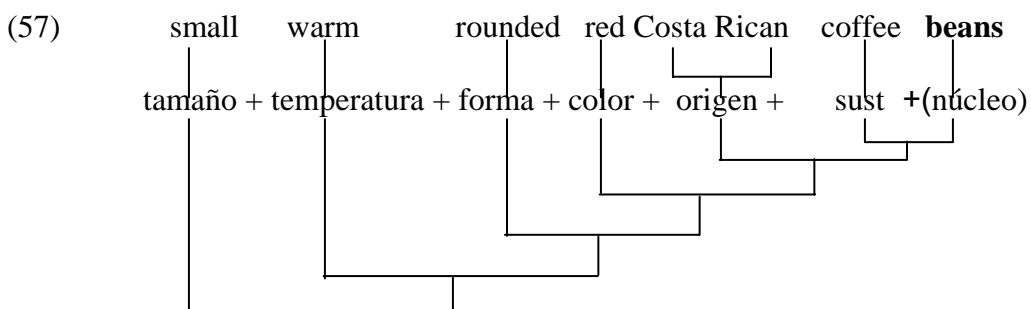
El estudio comparado de estas estructuras no es suficiente para detectar áreas posibles donde se produzca alguna interferencia sintáctica sutil, especialmente cuando se tratan de sintagmas complejos. Es necesario examinar el orden anidado en que los adjetivos descriptivos suelen presentarse en ambos idiomas. A continuación se describen algunos patrones de ordenamiento sintáctico que provocan interferencia en caso de construcciones sintácticas contrarias en la otra lengua:

INGLÉS

Premodificadores:

Valor + tamaño + temperatura o edad + forma + textura + color + origen + material
+ (**núcleo**) (Flores y otros 53).

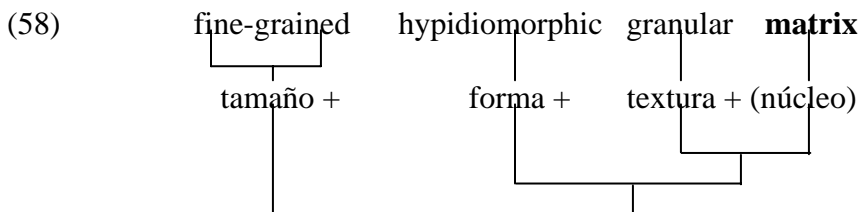
Ejemplo:



Sintagma tomado de: (Flores y otros 53).

Aunque *coffee* ya no forma parte de la cadena de adjetivos descriptivos, es un sustantivo modificador del núcleo del sintagma, y según Gleason (140) debe siempre ir adyacente al núcleo.

Los informes petrográficos muestran el orden de los atributos:



Sintagma tomado de: (Shannon y otros 6)

En este caso se describen las muestras de cristales de una roca. Aunque *fine* se encuentra en un nivel sintáctico más interno respecto al resto de los atributos de la frase, es decir, es un constituyente inmediato de *grained*, sirve para ejemplificar el orden del adjetivo de tamaño en el sintagma. Luego se incluye la forma (*hypidiomorphic*), y por último la textura (*granular*).

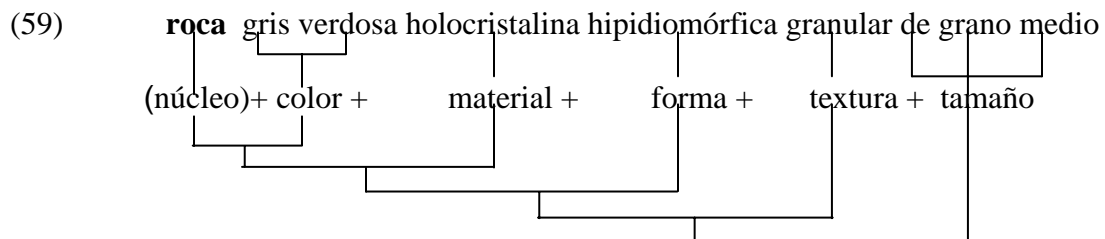
ESPAÑOL

A partir de los casos analizados en el capítulo anterior, se obtiene que el sistema lingüístico del español por lo general sólo permite un adjetivo antes del núcleo nominal excepto si se aplica coordinación o yuxtaposición, o si un segundo adjetivo es numeral o cuantificador. En presencia de varios adjetivos descriptivos, la prioridad para ubicar uno pospuesto al núcleo nominal en español sería la siguiente:

(núcleo) + color + material + forma+ textura+ edad o temperatura + tamaño + valor + origen

Nota: El orden podría variar en algunos casos.

Ejemplo típico de un informe petrográfico:



(Kussmaul 39-41)

Si se compara este sintagma con el del ejemplo (58) “fine-grained hypidiomorphitic-granular **matrix**,” se puede notar que el orden de postmodificación del primero es distinto al orden de premodificación del segundo. Esta caracterización es esencial para que el traductor sea consciente de que este orden inverso o distinto podría causar interferencias en las traducciones. Por ejemplo,:

(60) **basalto** verde con bordes irregulares (Geo. Schofield 3) (Traducción)

Angloparlante nativo: ragged-edged green **basalt** (Geo. Schofield Co 3)

Traductor hispanohablante nativo: green ragged-edged **basalt**

En el caso del sintagma producido por el hablante nativo del español, existe una interferencia porque se incluye primero el adjetivo de color. Este tipo de error puede producir un discurso poco natural y desacostumbrado para el lector hablante nativo del inglés. El ejemplo anterior sirve para ilustrar cómo un tipo de construcción sintáctica de orden inverso en ambos idiomas podría causar interferencias.

Un ejemplo de interferencia tomado de las retrotraducciones realizadas en este trabajo es:

(61) **ensayo** dúplex en tiempo real (Timken y otros 1) (Traducción)

Hablante nativo del inglés: duplex real-time qPCR **assay** (Timken y otros 1)

Traductor hablante nativo del español: real-time duplex qPCR assay

En este caso, el adjetivo *duplex* es numeral, y según el orden propuesto por Gleason (140), debería anteponerse a cualquier otro adjetivo. La tendencia del hispanohablante fue la de conservar la adyacencia de “dúplex ” al núcleo en español y trasladarla al inglés.

Otro caso en el que los traductores hispanohablantes podrían variar el orden de los modificadores es el caso de la estructura del inglés: adj + sust + (núcleo) descrita en (6), dado que en español es inversa: (núcleo) + sust + adj como corresponde a los apositivos. Por ejemplo:

(62) **regiones** minisatélites hipervariables (Entrala 2)

Angloparlante nativo: hipervariable minisatellite **regions** (Entrala 3) (Traducción)

Traductor hispanohablante nativo: minisatellite hipervariable **regions**

En (62) se observa que el traductor coloca el sustantivo antes del adjetivo en premodificación, lo cual rompe con la estructura del inglés descrita en (6).

Orden de modificación según función especificativa y explicativa

Los elementos del sintagma nominal que cumplen funciones especificativas o explicativas se ordenan en forma inversa en inglés y español.

Las funciones especificativas y explicativas son importantes porque están ligadas a estructuras sintácticas en idiomas como el español y el inglés. La función especificativa da mayor énfasis al núcleo, mientras que la explicativa se enfoca más en los atributos. En español, la forma especificativa del sintagma nominal se expresa mediante estructuras de postmodificación, mientras que la explicativa, se caracteriza por la anteposición del adjetivo (ver Cuadro 1). Por el contrario, la función especificativa del sintagma nominal del inglés se expresa mediante adjetivos premodificadores, y la explicativa suele asociarse a las estructuras de postmodificación, que incluye posposición del adjetivo en algunos casos, y la utilización de frases preposicionales.

INGLÉS

Especificativo: premodificador(es)+(núcleo)

Explicativo: (núcleo) + postmodificador(es)

Ejemplos:

(63) forensic DNA **quantifications** (Timken y otros 1)

(64) **quantification of DNA** in forensic samples (Nicklas y otros 1)

Los adjetivos antepuestos en (63) cumplen una función especificativa, es decir, delimitan las características del núcleo nominal y son restrictivos, aunque en casos específicos también pueden ser no restrictivos. Por otro lado, los atributos pospuestos en (64) más bien cumplen una función explicativa, y aparte de ello, este

tipo de estructuras también se utilizan en casos particulares. Por ejemplo, los atributos poseídos por objetos inanimados suelen colocarse preferentemente como frases preposicionales de postmodificación en inglés (Stockwell y otros 94):

(65) the **top** of the mountain (en vez de) the mountain **top**

En español: la **cima** de la montaña

(66) the **surface** of the water (en vez de) the water **surface**

En español: la **superficie** del agua

En casos particulares como (65) y (66), la posmodificación del sintagma nominal del inglés coincide tanto en estructura como en función comunicativa con el sintagma del español.

ESPAÑOL

Especificativo: (núcleo) + postmodificadores(s)

Explicativo (epíteto): premodificadores(s)+(núcleo)

Si el adjetivo es **especificativo** se coloca después del núcleo.

Si el adjetivo es **explicativo** se coloca antes del núcleo.

Ejemplo:

- (67) **transiciones** de fusión falsas (Timken y otros 24) (Traducción)
- (68) falsas **transiciones** de fusión

Los adjetivos del ejemplo (67) cumplen una función especificativa. Nótese que dicho ordenamiento es inverso al que se utiliza en inglés para la función especificativa. En (68), el adjetivo “falsas” es explicativo y no se utiliza en las fuentes estudiadas.

Este tipo de estructuras son fuente de interferencias en las traducciones, porque el traductor puede confundir las funciones comunicativas y lograr efectos contrarios. Por ejemplo:

- (69) **amplificación** de una secuencia de análisis de ~170-190 pb que abarca la región de repetición STR del locus TH01 (Timken y otros 1) (Traducción)

Angloparlante nativo: a ~170-190 bp target sequence **amplification** that spans the TH01 STR-locus repeat region (Timken y otros 1)

(Cumple una función especificativa).

Traductor hispanohablante nativo: an **amplification** of a ~170-190 bp target sequence that spans the repeat region of the TH01 STR locus

(Cumple una función explicativa).

El discernimiento entre las funciones comunicativas depende de la situación y de cuán específico sea el núcleo. Dado el alto grado de especificidad de los sintagmas en el género científico-técnico, el traductor puede caer en el otro extremo de sobregeneralizar la función especificativa. Por ejemplo:

(70) la **probabilidad** de que dos individuos sean genéticamente iguales

(Entrala 2)

Angloparlante nativo: the **probability** of genetic matching between two individuals

Traductor hispanohablante nativo: the genetic matching **probability** between two individuals

En este ejemplo el traductor hispanohablante sobregeneraliza la función especificativa. Esta “probabilidad” no es específica porque no se está siendo diferenciada de otros tipos de probabilidades en el texto.

INTERFERENCIAS LINGÜÍSTICAS CAUSADAS POR LA AUSENCIA O BAJA FRECUENCIA DE ESTRUCTURAS DEL SEGUNDO IDIOMA EN LA LENGUA MATERNA DEL TRADUCTOR

A. Traducción basada en la semejanza de estructuras alternativas presentes en ambos idiomas

Esta tendencia sintáctica se basa en el estudio descriptivo del capítulo anterior, en el que se comparan algunos tipos de construcciones del idioma inglés que se encuentran ausentes en español. En este punto se establece que una de las formas en la que los traductores hispanohablantes nativos resuelven esta ausencia o baja frecuencia de construcciones es basándose en las semejanzas sintácticas de estructuras alternativas en ambos idiomas. Por ejemplo, si en español no se utiliza premodificación compleja, cuando el traductor traduce desde su lengua materna podría recurrir a otro tipo de construcciones del segundo idioma, las cuales son más semejantes a la de su lengua materna. Este tipo de solución puede ser ventajosa porque se encuentra una manera eficiente de traducir los sintagmas. Sin embargo, los efectos estilísticos de estas construcciones no siempre son los más apropiados según el género textual que se trate. A continuación se ejemplifica este fenómeno:

1. Baja producción de modificadores compuestos léxicos

Las traducciones inversas al inglés pueden presentar preferencia por los adjetivos premodificadores coordinados en vez de premodificadores compuestos léxicos. En español no se utiliza premodificadores compuestos, y esta ausencia

estructural podría conducir al traductor a buscar alguna otra estructura del segundo idioma que sea más semejante a su idioma materno.

Por ejemplo:

(71) **ensayo** dúplex de qPCR nuclear y mitocondrial (Timken y otros 1)

(Traducción)

Angloparlante nativo: nuclear-mitochondrial duplex qPCR **assay**

(Timken y otros 1)

Traductor hispanohablante nativo: nuclear and mitochondrial duplex qPCR **assay**

Nótese que la construcción formada por el hispanohablante utiliza una forma que se asemeja a una construcción del idioma materno del traductor, y coincide con (38). La opción del traductor en (71) no es estructuralmente incorrecta, sino que es menos sintética. La opción de formar modificadores compuestos en vez de coordinados contribuiría al efecto estilístico condensado del inglés del género científico, el cual persigue utilizar relaciones sintácticas más concisas en la modificación. Otro ejemplo que muestra la baja producción de estos modificadores compuestos léxicos es el siguiente:

(72) cualquier **SNP** conocido por su relación con alguna enfermedad
(Timken y otros 19) (Traducción)

Angloparlante nativo: any well-known disease-associated **SNPs**

(Timken y otros 5)

Traductor hispanohablante nativo: any **SNP** known due to its relation to some
disease

En este caso, el traductor hispanohablante prefiere la postmodificación y utiliza un participio junto con formas coordinadas. En otros casos el traductor también puede recurrir a las cláusulas relativas.

Utilizar frases preposicionales postmodificadoras adyacentes al núcleo o a un participio postpuesto al núcleo, en vez de premodificadores de compuestos léxicos, es otra de las tendencias sintácticas donde los traductores hispanohablantes podrían diferir en estilo respecto a los hablantes nativos del inglés en los textos científicos. Dada la ausencia de estructuras formadas por modificadores complejos cuyo segundo componente se deriva de una forma verbal, adjetiva o adverbial en español, el traductor recurre a una estructura del idioma meta que sea más semejante a su idioma materno, tal como la que se menciona. Un ejemplo de este caso es:

(73) **sondas** de detección de sustancias específicas

(Timken y otros 1) (Traducción)

Angloparlante nativo: target-specific detection **chemistries** (Timken y otros 1)

Traductor hispanohablante nativo: **chemistries** for the detection of specific targets

Este ejemplo muestra que el traductor hispanohablante recurre a la posmodificación de su idioma materno, y construye un sintagma más largo en inglés, formado por frases preposicionales postmodificadoras. Sin embargo, el hablante nativo del inglés prefiere utilizar un modificador compuesto de uso generalizado en textos científicos, cuyo segundo componente es una forma adjetiva. El siguiente ejemplo muestra la alternativa “núcleo+participio+frase” del mismo fenómeno:

(74) un **gel** que ilustra el grado de fragmentación del ADN para las muestras tratadas con DNasa (Timken y otros 29) (Traducción)

Angloparlante nativo: a **gel** illustrating the degree of DNA fragmentation for the DNase-treated samples (Timken y otros 9)

Traductor hispanohablante nativo: a **gel** indicating the degree of DNA fragmentation for the samples treated with DNasa

En este ejemplo, el traductor soluciona la ausencia de modificadores compuestos premodificadores en su idioma materno, y forma una estructura que se comparte en ambas lenguas.

2. Baja producción de participios (-ing) postpuestos al núcleo:

(núcleo) + participio

A partir del análisis del corpus de este trabajo, se observa que la frecuencia de participios postpuestos al núcleo es menor en español respecto al inglés (ver ejemplo (43)), y la estructura más utilizada es la cláusula relativa. Dadas estas condiciones, el traductor hispanohablante nativo recurre a esta última forma estructural que se comparte en ambas lenguas pero con distinta distribución. Por ejemplo:

(75) un **medio** que contiene uno de los cuatro nucleótidos (Entrala 13)

Angloparlante nativo: a **medium** containing one of the four nucleotides

Traductor hispanohablante nativo: a **medium** that contains one of the four nucleotides

En este caso, el participio presente (-ing) del inglés se utiliza como modificador. En español no ocurre así con el gerundio, especialmente cuando el núcleo es un nombre inanimado. Se observa que el traductor prefiere introducir el pronombre relativo “that” y utiliza una cláusula subordinada. La preferencia por estas cláusulas también se refleja en el uso del participio pasado (-ed).

3. Sobreproducción de sintagmas nominales con determinante

Dado que en inglés los sintagmas nominales pueden o no incluir un artículo definido según la función restrictiva o no restrictiva del núcleo, mientras que en español el artículo se conserva para ambas situaciones, los traductores hispanohablantes suelen incluir mayor número de estos determinantes en los textos.

Por ejemplo:

(76) la **actividad** de la *Taq polimerasa* (en un contexto general: núcleo sustantivo inespecífico o no restrictivo) (Entrala 7)

Angloparlante nativo: Taq polymerase **activity**

Traductor hispanohablante nativo: the Taq polymerase **activity**

El traductor incorpora el artículo “the” al sintagma, y lo torna restrictivo. Para un contexto general, este artículo se omite en inglés. Otro extremo consistiría en evitar siempre el artículo, lo que causaría una sobregeneralización de la regla.

4. Baja producción de premodificación compleja

Como se mostró en el capítulo anterior, el español carece de construcciones anidadas de premodificadores, y una de las alternativas de los traductores hispanohablantes para resolver este faltante es el uso de frases postmodificadoras que existen tanto en inglés como en español, pero con frecuencia y funciones no completamente correspondientes (Ver Cuadro 2). Por ejemplo:

(77) los **métodos** proteicos y séricos utilizados con anterioridad en el área forense (Miller y otros 4) (Traducción)

Angloparlante nativo: previously used forensic protein and serological **methods**

(Miller y otros 4)

Traductor hispanohablante nativo: **methods** based on proteins and serum, which were previously used in forensics

Aquí el traductor sigue la ramificación derecha de modificadores de su idioma materno, y alarga más el sintagma al introducirle subordinación.

B. Hipercorrección

La hipercorrección es un fenómeno que se produce cuando el hablante “abusa” de determinada estructura lingüística y la emplea en contextos inapropiados. En este estudio se detecta como parte del fenómeno de interferencia lingüística. De acuerdo con Gleason (382), la hipercorrección sintáctica se produce cuando el hablante no nativo de una lengua realiza un esfuerzo por corregir sus estructuras, pero a pesar de esta actitud, sus producciones sintácticas reflejan que no comprende completamente el patrón normativo de la lengua meta. A continuación se ejemplifican dos tipos de situaciones sintácticas en las que se manifiesta el uso hipercorrectivo de los patrones del segundo idioma.

1. Abuso en la longitud de construcciones atributivas anidadas

El traductor hispanohablante podría ser consciente de la necesidad de incluir construcciones premodificadoras anidadas en los textos de inglés académico. Como consecuencia, es probable que incurra en complicar de manera extrema dicho tipo de construcciones. Por ejemplo:

(78) **cuantificación** simultánea de ADN nuclear y mitocondrial en muestras forenses (Timken y otros 1) (Traducción)

Angloparlante nativo: simultaneous **quantification** of human nuclear and mitochondrial DNA in forensic samples (Timken y otros 1)

Traductor hispanohablante nativo: human nuclear and mitochondrial DNA simultaneous forensic sample **quantification**

En (78) la estructura formada por el traductor resulta muy compleja porque incluye una construcción coordinada antepuesta a una estructura anidada de atributos. Además, utiliza un sintagma nominal complejo para modificar a otro igual. Como se observa en la descripción lingüística del capítulo 2, en inglés no suelen ocurrir tales tipos de anidamientos. La utilización de sintagmas complejos modificadores de otros sólo ocurre en posición posmodificadora como en (31). Duque señala que el número de premodificadores raramente debe exceder a tres por

motivos de comprensión de las ideas, y que se pueden usar frases preposicionales para expresar el resto de la información (146).

Sin embargo, Quirk y otros (1298) indica que la postmodificación múltiple en inglés podría resultar en ambigüedad y en problemas estilísticos como la inaceptabilidad estructural, y afirma que un orden cuidadoso de los constituyentes en el sintagma nominal es fundamental para comunicar la intención. En relación con lo mismo, Duque señala que más de tres frases preposicionales seguidas producen un estilo pomposo y artificioso (170). Por lo tanto, un uso balanceado de las estructuras es clave para mantener la claridad textual. Otro ejemplo de hipercorrección es el abuso en el número de sustantivos premodificadores:

(79) **evolución** de las técnicas de estudio de los polimorfismos de ADN

Angloparlante nativo: **evolution** of techniques in the study of DNA polymorphism

Traductor hispanohablante nativo: **evolution** of DNA polymorphism study techniques

El sintagma del traductor es poco natural, y en el manual de estilo para el género científico-técnico en inglés, Duque recomienda no exceder dos sustantivos antes del núcleo, y advierte que es preferible que haya un sólo sustantivo en función premodificadora (146).

2. Sobregeneralización del participio con terminación (-ed) en posición de premodificador

Otro tipo de hipercorrección sintáctica asociada con la sobregeneralización de la segunda lengua es el uso extendido del participio con terminación (-ed) del inglés en posición de premodificador. Algunos hispanohablantes generalizan la función adjetiva antepuesta del participio y obvian la función propia del participio derivada de construcciones pasivas, frases fijas, la condición de estado o proceso, y de permanencia o temporalidad, etc. Respecto a estos últimos factores, Quirk describe detalladamente los casos en los que un participio, en este caso el terminado en *-ed*, se utiliza de forma antepuesta o postpuesta en el sintagma nominal del inglés (1327-1330). Además del esfuerzo correctivo, este tipo de interferencia conlleva el factor de ausencia o rareza de los participios antepuestos al núcleo en español. Algunos ejemplos de esta sobregeneralización son:

(80) **Información** provista (en un título)

Angloparlante nativo: **Information** Provided (Entrala 5) (Traducción)

Traductor hispanohablante nativo: Provided **Information**

(81) Personas entrevistadas (en un título).

Angloparlante nativo: **People** Surveyed (Salleh 1)

Traductor hispanohablante nativo: Surveyed **People**

Nótese que en estos casos, el hablante nativo del inglés opta por posponer el participio de función adjetiva dada la condición fija de la frase en el título, y además el participio indica una condición temporal, que implica postposición al núcleo.

En (80) se observa una forma simplificada de la cláusula:

(82) Information that was provided

En (81) también se simplifica la siguiente cláusula:

(83) People that were surveyed.

En resumen, las interferencias lingüísticas que suelen aparecer en la traducción inversa (español-inglés) del sintagma nominal son producto del orden inverso de algunas construcciones sintácticas en ambos idiomas, así como de la ausencia de construcciones equivalentes de la segunda lengua respecto al idioma materno del traductor. Aunque un traductor hispanohablante presente un nivel de competencia lingüística muy amplio en su segundo idioma, es importante que retome un estudio como éste, en el cual podría visualizar de forma descriptiva algunos de los procesos sintácticos que causan desventajas estilísticas y comunicativas en los textos traducidos, y en este caso especial, en los textos científicos o de inglés profesional y académico.

Sería ideal para el traductor interiorizar el orden usual de los modificadores en inglés, dado que este es uno de los principales factores de interferencia. Además, quienes traducen deben tener en cuenta que las semejanzas o coincidencias

estructurales entre un idioma y otro no implican correspondencia pragmática. También, es importante considerar factores de hipercorrección como los analizados en este proyecto porque el afán por producir estructuras “correctas” podría implicar una sobregeneralización de algunas.

El siguiente cuadro resume el tipo de interferencia, las causas estructurales y la forma simplificada del sintagma nominal en ambos idiomas para explorar los puntos de divergencia y equivalencia. Aquí se muestra el aporte de esta investigación en cuanto a la clasificación de las interferencias encontradas, y la descripción general de la forma resultante.

Cuadro 3. Resumen de resultados³

Tipo de interferencia	Causa estructural	Forma en L1⁴	Forma en L2	Forma producida por interferencia en L2
Desubicación de premodificadores	Orden inverso entre modificadores de L1 y L2	(núcleo)+color+material+ forma+ textura+ edad o temperatura+tamaño+valor + origen (núcleo)+sust+adj	valor + tamaño + temperatura o edad + forma + textura + color + origen + material + (núcleo) adj+sust+(núcleo)	Una posible: color+ forma+(núcleo) sust+adj+(núcleo)
Desubicación de modificadores	Orden inverso de modificación explicativa o especificativa	Especificativo: (núcleo)+postmodificador(es) Explicativo (epíteto): premodificador(es)+(núcleo)	Especificativo: premodificador(es)+(núcleo) Explicativo: (núcleo)+postmodificador(es)	Especificativo: (núcleo)+postmodificador(es) Explicativo: premodificador(es)+(núcleo)
Baja producción de modificadores compuestos léxicos	Ausencia en L1 de modificadores compuestos léxicos en premodificación	∅ Estructuras alternativas: ver la forma producida por la interferencia en L2	palabra-palabra+(núcleo)	modif+cons.coor+modif+ (núcleo) núcleo+cláusula (núcleo)+frase núcleo+participio+frase
Baja producción de participios (-ing) postpuestos	Mayor frecuencia de (núcleo)+cláusula en L1	<< (núcleo)+gerundio/particip Estructura alternativa: ver la forma producida por la interferencia en L2	(núcleo)+participio (-ing) (-ed)	(núcleo)+cláusula

⁴ **Simbología:** L1: Lengua materna (español); L2: segunda lengua (inglés); <<: mucho menos frecuente; >>: mucho más frecuente; ∅: estructura o elemento ausente.

Cuadro 3 (continuación). Resumen de resultados

Tipo de interferencia	Causa estructural	Forma en L1	Forma en L2	Forma producida por interferencia en L2
Sobreproducción de sintag. nom. con determinante	Muy baja frecuencia de sintag. nom. sin determinante en L1	artículo+(núcleo) << Ø+(núcleo)	artículo+(núcleo) Ø+(núcleo)	>>artículo+(núcleo)
Baja producción de premodificación compleja	Ausencia en L1 de premodificación anidada compleja sin yuxtaposición	Ø Estructuras alternativas: ver la forma producida por la interferencia en L2	distintos niveles de premodificadores+(núcleo)	(núcleo) +distintos niveles de postmodificadores Oraciones en vez de sintagmas
Sobreproducción de premodificación compleja	Ausencia en L1 de premodificación anidada compleja sin yuxtaposición	Ø	algunos niveles de premodificadores+(núcleo)	exceso de niveles de premodificadores+(núcleo)
Sobregeneralización de atributos antepuestos al núcleo	Baja frecuencia de premodificadores adjetivos en L1	<< adj+(núcleo)	>>adj+(núcleo)	Sobregeneralización de adj+(núcleo)

CONCLUSIONES

Los estudios lingüísticos aplicados a la traducción reivindican el empleo de los métodos analíticos científicos al análisis de los fenómenos traductológicos. Una teoría del lenguaje suministra una base objetiva y confiable para derivar conclusiones, con lo que se evitan los planteamientos subjetivos o especulativos. El análisis lingüístico descriptivo y contrastivo de este proyecto permite detallar las implicaciones estructurales y estilísticas de las interferencias en la traducción inversa, y explora sus causas según el contacto entre la lengua materna y la segunda lengua.

Hay diversidad de opiniones entre los lingüistas respecto a la ventaja o desventaja de las diferencias y similitudes entre los sistemas lingüísticos del individuo bilingüe, que en este caso sería el traductor hispanohablante. A partir de este proyecto se establece que el orden inverso o contrastante entre los modificadores del sintagma nominal del inglés y el español son una fuente de interferencias basadas en las disimilitudes entre los sistemas lingüísticos. Este factor incluye el ordenamiento de los premodificadores según el tipo de adjetivo, y el orden según la función especificativa o explicativa del sintagma. Otro aspecto que contribuye al desarrollo de las interferencias y que está relacionado con la diferencia tipológica es la ausencia en español de premodificadores compuestos léxicos y con afijos, y la ausencia de premodificación anidada compleja.

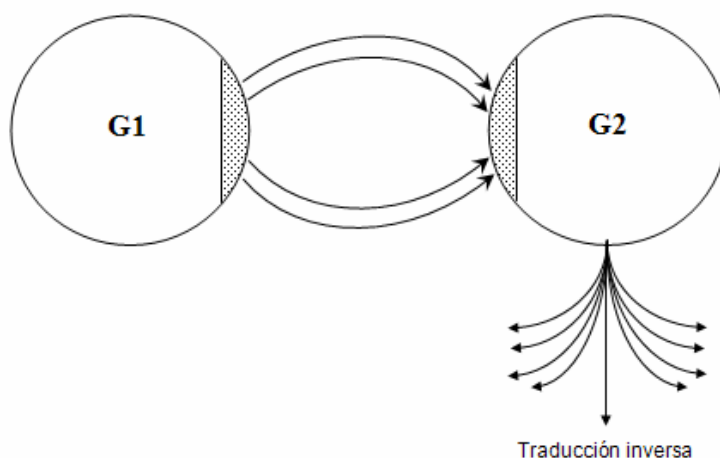
A partir de los datos analizados en el corpus se infiere que un factor de interferencia en la traducción inversa del sintagma nominal es la búsqueda de

equivalencia formal para facilitar la labor traductológica, la cual no es exacta, sino de naturaleza aproximada, tal como lo señala Catford (33). Este factor implica que a veces se eluda la aplicación situacional y textual de esas estructuras. A nivel general se vislumbran dos reacciones del traductor hispanohablante ante las diferencias de tipología estructural entre las lenguas que utiliza. La primera consiste en refugiarse en la lengua materna o dominante, y la segunda implica refugiarse en la aplicación de reglas del L2. Estas dos actitudes se convierten en factores potenciales para las desviaciones de la norma, no sólo dentro del sistema lingüístico en general, sino en lo relativo a los géneros discursivos.

Cuando hay diferencias tipológicas entre las lenguas, lo ideal es que el traductor reaccione con un sistema de decodificación y codificación compuesto o coordinado en el que se genere un canal de respuesta independiente para la lengua meta, y la traducción se base en la equivalencia estructural dinámica. Uno de los mecanismos de interferencia confirmados por este proyecto es el uso recursivo de estructuras alternativas en ambos idiomas que sean semejantes entre sí en cuanto a su forma, pero no en cuanto a sus características funcionales, comunicativas, o discursivas. A partir de este punto se producen calcos estructurales de la lengua materna, fundamentados en aproximaciones formales que no necesariamente producen errores gramaticales en el texto meta. Si la estructura formal semejante que se plantea para resolver la diferencia tipológica no tiene en cuenta los factores de función y distribución, propuestos por Lado (146), el traductor se desvía de la norma lingüística y disminuye la aceptabilidad de su texto.

El mecanismo de interferencia basado en el calco estructural, que ocurre cuando los premodificadores en el sintagma presentan rasgos de orden inverso, se puede representar así:

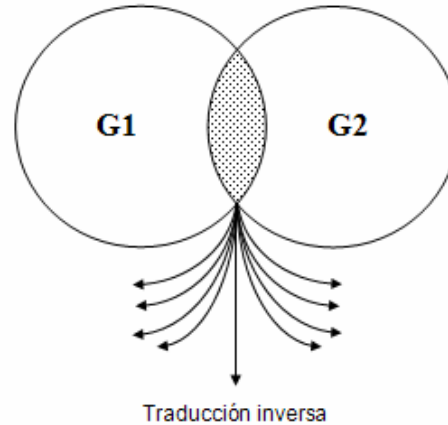
Esquema 1:



En este esquema G1 es la sigla para la gramática del español, y G2 la del inglés. La zona sombreada en G1 indica las construcciones de ordenamiento de los premodificadores en el sintagma nominal. Esta zona sombreada se traslada a G2 durante la traducción inversa. Como consecuencia, el traductor construye una gramática nueva, no totalmente compatible con G2, y el texto meta se compone de una parte gramatical G2 con equivalencia dinámica respecto a la G1, y otra parte correspondiente a calcos estructurales sutiles adoptados durante el proceso.

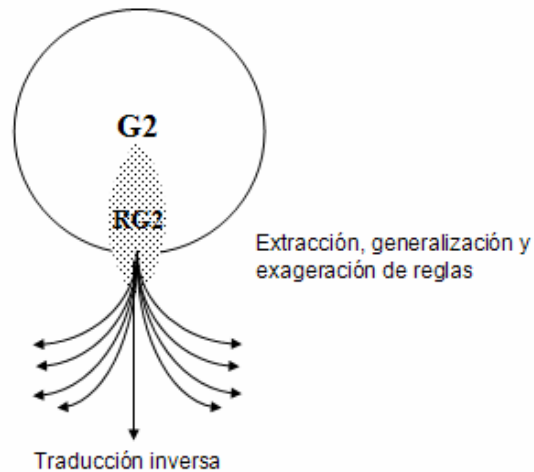
Por otro lado, el mecanismo de interferencia basado en las semejanzas de forma sintáctica se presenta en el esquema 2. Aquí, el proceso parte de una zona común entre los dos sistemas, que se indica mediante la parte sombreada, y se relaciona con la forma de las estructuras. Las interferencias que surgen conllevan rasgos estilísticos de tipología textual y función.

Esquema 2:



Finalmente, el mecanismo de traducción basado en la hipercorrección o sobregeneralización, relacionado con la aplicación de reglas gramaticales en la traducción, se representa en el siguiente esquema:

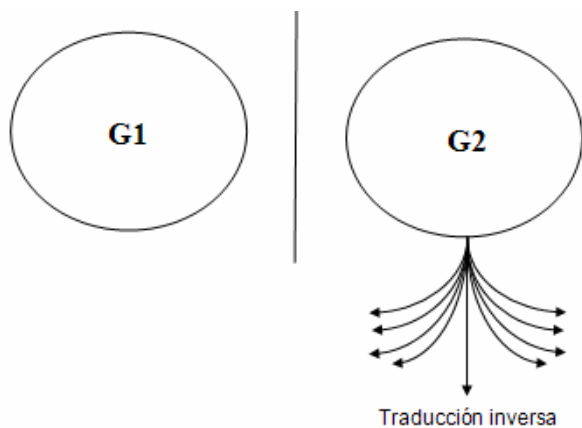
Esquema 3:



En el esquema 3. el traductor recurre sólo a la gramática de su segunda lengua, y la traducción de una estructura parte de un conjunto de reglas gramaticales extraídas (RG2) del sistema gramatical global del L2. Como consecuencia, ocurre la hipercorrección y la sobregeneralización en la traducción inversa, porque no se parte del conjunto total de posibilidades gramaticales de L2.

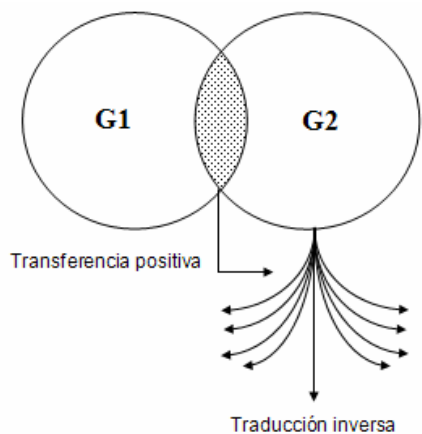
El modelo ideal de la traducción inversa se puede representar de dos maneras en los esquemas siguientes:

Esquema 4:



En este esquema, el sistema de la lengua materna y la segunda lengua son totalmente independientes, y el traductor no tendría interferencias. Sin embargo, este modelo no sólo se aleja de la realidad, sino que las interferencias pueden ser positivas en algunos casos, tal como Odlin señala al referirse a la “transferencia positiva”. Dado que los sistemas lingüísticos pueden compartir estructuras, especialmente cuando tienen parentesco cercano, un modelo más realista respecto al uso y combinación de gramáticas en la traducción, se representa en el siguiente esquema:

Esquema 5:



Como se ha señalado en este proyecto, las semejanzas no siempre son ventajosas, y el esquema 5 se aplicaría más a casos donde el traductor trabaja con lenguas muy afines entre sí, por lo que no dejaría de ser un modelo idealizado.

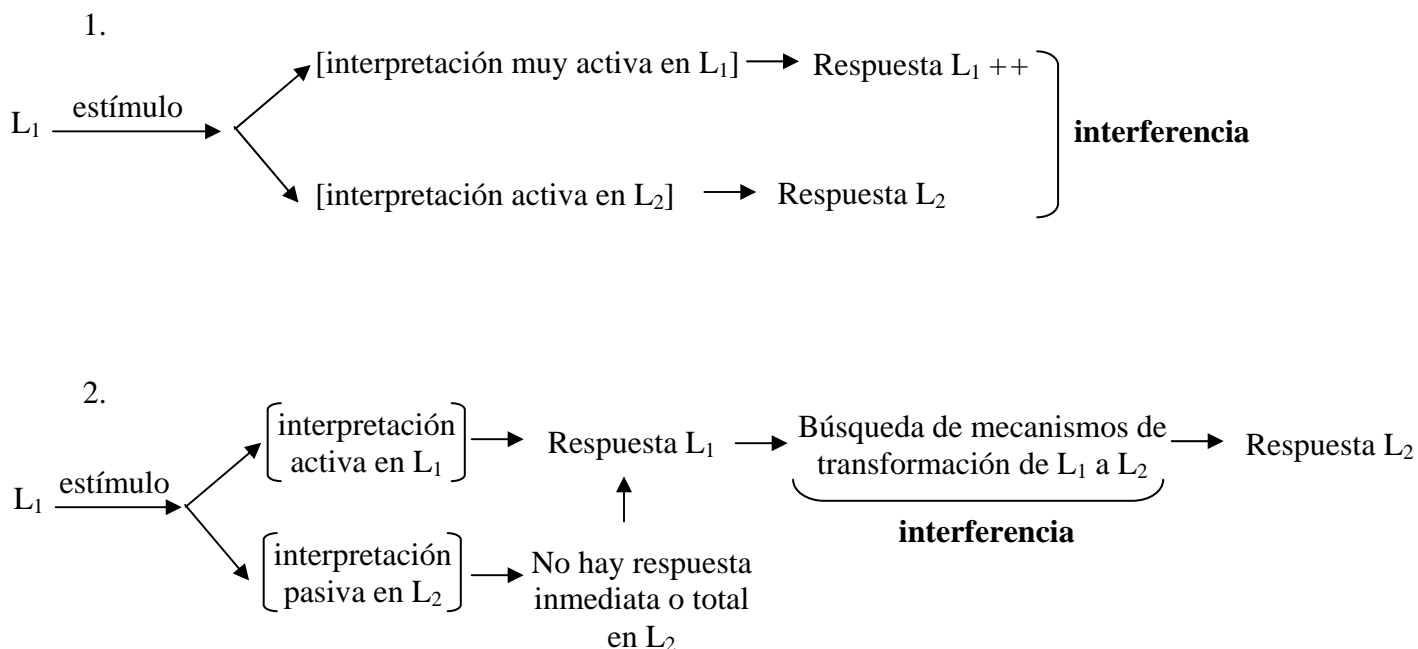
Una parte del proceso de las interferencias se explica mediante el análisis inductivo del comportamiento del traductor, a partir de dos premisas que se desprenden del modelo bilingüe de Weinreich, pero con una aplicación completamente traductológica. La primera es que el traductor interpreta los estímulos simultáneamente en ambas lenguas, si se trata de una situación interlingüística. La dirección de la traducción no es un factor a considerar dentro de esta premisa porque independientemente de si el estímulo lingüístico inicial se produce en la lengua materna o en la segunda lengua, el sistema de codificación y decodificación actuaría en forma simultánea para ambas lenguas. Durante este proceso, la lengua materna puede producir la interferencia si esta se considera como la más influyente en el individuo, ya sea por la exposición a un medio hispanohablante, como es el caso de los traductores que residen en Costa Rica, o por el orden de aprendizaje de las lenguas.

Por otro lado, la segunda premisa es que el traductor subordina la respuesta en la segunda lengua a los estímulos recibidos en la lengua materna. Partiendo de ahí, la estructura de la lengua materna condicionaría en gran parte a la traducción sintáctica inversa. De ser así, el traductor calca estructuras de su lengua materna, o bien, realiza un esfuerzo por aplicar las reglas del inglés de forma rígida con el fin de evitar las desviaciones de la norma. Este proceso de interferencia no se consideraría

común en un traductor competente, sin embargo podría surgir de forma sutil en casos específicos.

Dentro de un modelo bilingüe compuesto, y teniendo en cuenta los conceptos de estímulo y respuesta utilizados por Osgood, los procesos de interferencia de la lengua materna en la traducción inversa, se pueden plantear así:

Diagrama psicotraductológico de la interferencia de la lengua materna en la traducción inversa



En 1. se observa que el estímulo inicial producido por la lengua materna del traductor, que en este caso sería un texto T1, pero que se podría proyectar al ámbito oral, se interpreta simultáneamente en español e inglés durante una situación interlingüística. Sin embargo, la interpretación en L1 es más activa que en L2. Como consecuencia, la respuesta en L1 hacia ese estímulo es mucho más pronunciada que en L2. La interferencia surge al activar ambos sistemas simultáneamente, dado

que puede haber intercambio inconsciente según las reacciones mentales del traductor. Este modelo de traducción es variable y depende de la situación. Según el medio interlingüístico al que el traductor se exponga, varía el grado de interferencia que experimente, y esto tiene que ver con su desempeño, y no con su competencia. Este aspecto abre las puertas para nuevas investigaciones que también abarquen el desempeño oral, como en el caso de la interpretación, dado que durante el proceso de producción oral los intérpretes podrían experimentar distintos grados de presión de una lengua sobre otra al tener que combinar ambos sistemas lingüísticos simultáneamente.

En 2., la interpretación del estímulo en L1 causa una interpretación activa en L1, pero una pasiva en L2. Dadas estas condiciones, no se produce una respuesta inmediata o total en L2, y el traductor recurre a la respuesta que surgió en L1. A partir de aquí se produce una búsqueda de mecanismos de transformación de las estructuras de L1 a L2, lo que sería un proceso de traducción a partir de la lengua materna, y es en esta fase donde se producen las interferencias que se transfieren a L2. Este proceso implica no sólo la interferencia de la lengua materna, sino también la aplicación de reglas de conversión gramaticales, que podrían resultar en hipercorrección y sobregeneralización. Las interferencias generadas por este modelo tienden a ser más permanentes y no tan situacionales como las del punto anterior. Aquí se pueden incluir aspectos como el tipo de asimilación o los métodos de aprendizaje de la segunda lengua que experimentó el traductor, entre otros. En este aspecto entra en juego la competencia lingüística estructural en casos sutiles.

Al incurrir en las interferencias, el traductor puede trabajar con ambos mecanismos psicolingüísticos de interpretación de estímulo, combinándolos según la situación o característica sintáctica en particular. Esto significa que no todas las áreas de asimilación y producción lingüística trabajan igual en la mente del traductor. La influencia de la estructura de la lengua materna puede ser más fuerte en algunas áreas que en otras. Este tipo de modelos son útiles para plantear nuevas hipótesis y metodologías de investigación.

La observación es una herramienta importante para el planteamiento de hipótesis, tal como ha sido el caso de este proyecto de investigación. La observación comparativa de características estructurales del sintagma nominal en textos científicos del español y el inglés, indica que la organización de los constituyentes inmediatos en el idioma materno, así como su estructura en distintos niveles puede influir de manera sutil en la traducción inversa. Una exploración a la luz de la lingüística descriptiva y comparada permite concluir que las diferencias o semejanzas entre los dos sistemas lingüísticos pueden traicionar la mente del traductor y conducir el proceso de reacción lingüística hacia la lengua materna. La semejanza en la forma puede competir con la semejanza en la función comunicativa de las estructuras. La competencia textual y de género en ambos idiomas es un factor indispensable para reconocer el empleo inusual de alguna norma estructural, dada su distribución distinta según la tipología textual. Se puede afirmar que los juegos de semejanzas estructurales sutiles podrían incluso denominarse “cognados estructurales”, en analogía con los léxicos.

El planteamiento de problemas sintácticos en la traducción científico-técnica ha pretendido trascender los problemas de equivalencia terminológica. Como afirma Maillot, es un error “no ver en la traducción científica más problemas que los de vocabulario” (32), y recomienda la práctica de la estilística comparada. Una traducción científica técnica planteada sólo desde el punto de vista terminológico especializado podría verse reducida al esfuerzo de buscar o elaborar glosarios y diccionarios para lograr una equivalencia léxica. Esta investigación, en cambio, mediante una relación estrecha con la lingüística, ha intentado cuestionar ciertos comportamientos lingüísticos y analizarlos a la luz del estudio teórico-descriptivo y comparado de un corpus. Estas investigaciones contribuyen a que los traductores evalúen su desempeño y puedan establecer pautas de calidad según el estado sincrónico de las lenguas. Además, esto permite el desarrollo de una visión introspectiva sobre los mecanismos de traducción y sobre el grado de interferencia del traductor.

Lamentablemente, ha habido reticencia entre algunos en cuanto a analizar los cuestionamientos traductológicos desde una perspectiva lingüística. Según señala Maillot, entre los traductores “hay muchos que, no sólo no son lingüistas, sino que muestran incluso, con respecto a la lingüística, cierta repugnancia, la cual se explica en gran parte por la aversión que sienten hacia la anarquía reinante en el vocabulario de la lingüística” (170). Ante esta situación, se espera que este trabajo anime a los traductores a fundamentar sus hipótesis desde el punto de vista científico de la lingüística.

Para continuar con más estudios estructurales, se puede analizar por ejemplo la sobregeneralización del genitivo posesivo en la traducción inversa, la cual es común en otro tipo de textos. En futuras investigaciones afines, se recomienda el campo de las interferencias estructurales transferidas del inglés al español. En el caso del sintagma nominal, se puede estudiar el artículo contracto “del” frente al uso de la preposición “de”, y la forma de las frases adjetivas.

BIBLIOGRAFÍA

- Alcaraz Varó, Enrique. *El inglés profesional y académico*. Madrid: Alianza, 2000.
- Bosque, Ignacio y Violeta Demonte. *Gramática descriptiva de la lengua española: sintaxis básica de las clases de palabras*. Primer tomo. Madrid: Espasa Calpe, 1999.
- Cal Teba, María Luisa. *Análisis de polimorfismos de ADN microsatélite de cromosoma Y. Estudio de la población de Galicia y aplicaciones forenses*. Tesis de doctorado. 2001. <<http://www.cervantesvirtual.com/FichaObra.html?Ref=6669>> 24 de setiembre del 2006.
- Catford, J.C. *A Linguistic Theory of Translation: an essay in applied linguistics*. Oxford: Oxford University Press, 1965.
- Chang Díaz, Franklin., J.P. Squire., A.V. Ilin., G.E. McCaskill., T.X. Nguyen., D.S. Winter., A.J. Petro., G.W. Goebel., L.D. Cassady., K.A. Stokke., C.E. Dexter., T.P. Graves., L. Amador Jr., J.A. George., M.D. Carter., F.W. Baity Jr., G.C. Barber., R.H. Goulding., D.O. Sparks., S.W. Schwenterly., R.D. Bengtson., B.N. Breizman., V.T. Jacobson., A.V. Arefiev., R.Z. Sagdeev., K. Karavasilis., S.V. Novakovski., A.A. Chan y T.W. Glover. *National Aeronautics and Space Administration*. "The Development of the VASIMR Engine". 1999. <<http://spaceflight.nasa.gov/shuttle/support/researching/aspl/reference/develop.pdf>>. Octubre del 2006.
- Claros, Gonzalo M., Saladrigas, Verónica M., & Diego González Halphen. *Vocabulario inglés-español de bioquímica y biología molecular*. 2006. <<http://www.biorom.uma.es/contenido/Glosario/>>. 6 de setiembre del 2006.
- Cook, Vivian. *Linguistics and Second Language Acquisition*. Londres: The Macmillan Press, 1993.
- De Grève, Marcel y Frans Van Passel. *Lingüística y enseñanza de lenguas extranjeras*. Madrid: Fragua, 1971.
- Del Valle, Christian., Anayanci Rodríguez y Marta Espinoza. *Comparación de tres métodos de extracción de ADN a partir de restos óseos*. 2004. <http://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S0034-77442004000300032&lng=e&nrm=iso> 25 de setiembre del 2006.

- Díaz, José Andrés., Luis A. Abarca y Eduardo Macías. *Revista de Ciencias Forenses* I. "Probabilidad de ganar en un juego de azar cuestionado por fraude en Costa Rica". Asociación Costarricense de Ciencias Forenses (ACCF), 2006.
- Domínguez Vásquez, María José. *En torno a la interferencia lingüística*. 2001. <<http://www.ucm.es/info/circulo/no5/dominguez.htm>>. Octubre, 2006.
- Duque García, Ma. Mar. *Manual de estilo: el arte de escribir en inglés científico-técnico*. Madrid: Paraninfo, 2000.
- Entrala, Carmen. *Técnicas de análisis del ADN en genética forense*. 2000. <<http://www.ugr.es/~eianez/Biotecnologia/forensetec.htm>>. 4 de agosto del 2006.
- Francis, W. Nelson. *The Structure of American English*. Nueva York: Ronald Press Company, 1958.
- García Yebra, V. *Teoría y práctica de la traducción*, 2 tomos, Madrid: Gredos, 1989.
- Genome Glossary*. 2004. <http://www.ornl.gov/sci/techresources/Human_Genome/glossary/>. 6 de setiembre del 2006.
- Geo. Schofield Co., Inc.* "Natural Stone Products". <<http://www.geoschofield.com/boulders.htm>>. 1 de noviembre del 2007.
- Gleason, H.A Jr. *Linguistics and English Grammar*. Nueva York: Holt, Rinehart and Winston, Inc., 1965.
- Glick, David M. *Glossary of Biochemistry and Molecular Biology*. 1992. <<http://www.portlandpress.com/pp/books/online/glick/default.htm>>. 6 de setiembre del 2006.
- Hockett, Charles F. *A Course in Modern Linguistics*. Nueva York: The Macmillan Company, 1958.
- Ingman, Max. *El ADN mitocondrial esclarece la evolución humana*. 2001. <<http://www.actionbioscience.org/esp/evolution/ingman.html>>. 6 de setiembre del 2006.
- International Agency for Research on Cancer (IARC)*. "Dichlorobenzenes". 1999. <<http://www.inchem.org/documents/iarc/vol73/73-08.html>>. 2 de noviembre del 2007.

Jiménez Serrano, Oscar. *La traducción técnica inglés-español didáctica y mundo profesional*. Granada: Comares, 2002.

Kusssmaul, Siegfried. *Petrografía de rocas ígneas y metamórficas*. San José: Universidad de Costa Rica, 2003.

Laboratorio de Genética: Glosario de términos. 2000.

<<http://www.geocities.com/maorera/hqlaes2n.htm>>. 9 de setiembre del 2006.

Lado, Robert. *Language Testing*. Londres: Longman Group Limited, 1961.

Lemon, Stanley M. & Alan G. Barbour. *A Glossary of Terms Commonly Used in Molecular Biology*. 1993.

<www.med.unc.edu/wrkunits/3ctrpgrm/pmbb/mbt/GLOS.htm>. 6 de setiembre del 2006.

Liu, G; Stintz, A.; Li, H.; Malloy, K.J.; Lester, L.F. *IEEE*. "Extremely low room-temperature threshold current density diodelasers using InAs dots in $\text{In}_{0.15}\text{Ga}_{0.85}\text{As}$ quantumwell".

<<http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/2220/17077/00788943.pdf?isnumber=&arnumber=788943>>. Octubre del 2007.

Maillot, Jean. *La traducción científica y técnica*. Madrid: Editorial Gredos, 1997.

Mari Mutt, José A. Caribbean Journal of Science: *Manual de redacción científica*.

<<http://www.caribjsci.org/>>. Setiembre, 2006.

Marriott, G; Clegg, R; Arndt-Jovin, D y Jovin, T. *Biophys. Journal*. "Time resolved imaging microscopy". 1991.

<www.biophysj.org/cgi/reprint/60/6/1374.pdf>. 2 de noviembre del 2007.

Métodos de transferencia de proteínas a filtros. 2003.

<www.ub.es/biocel/wbc/tecnicas/westernblot.htm>. 4 de octubre del 2006.

Miller, Sasha., King, Sheria y Kalai Mathee. "An Overview of Human DNA Analysis in Forensics". *Revista de Ciencias Forenses*. Asociación Costarricense de Ciencias Forenses (ACCF), noviembre del 2006, vol. 1, No. 1.

Minnesota Planning. *Technical Work Paper for Human Health Issues*. 2001.

<http://www.eqb.state.mn.us/geis/TWP_HumanHealth.pdf>. 2 de noviembre del 2007.

Monge, Carlos Francisco. *La tinta extinta*. "Como el farero". Heredia: Universidad Nacional, 2007.

- Nicklas, Janice A., Eric Buel. *Analytical and Bioanalytical Chemistry*. "Quantification of DNA in forensic samples". Berlin: Springer Berlin / Heidelberg, agosto del 2003; vol 376, número 8.
- Odlin, Terence. *Language Transfer: cross-linguistic influence in language learning*. Nueva York: Cambridge University Press, 1989.
- Osgood, Charles E., Thomas A. Sebeok y A. Richard Diebold. *Psicolingüística: problemas teóricos y de investigación*. Barcelona: Planeta, 1974.
- Penacino, Gustavo Adolfo. *Investigación e implementación de sistemas de identificación de individuos por técnicas de biología molecular con especial referencia a los estudios post-mortem*. Tesis doctoral. 1997. <<http://www.biotech.bioetica.org/docta62.htm>>. 19 de setiembre del 2006.
- . *Actualización en estudios de paternidad por análisis de ADN*. 2002. <www.slagf.org/trabajopat.pdf>. 4 de agosto del 2006.
- Peter, Roland., Robert Gschwentner., Wolfgang Schürmann., Reinhard M. Rieger y Peter Ladurner. *Journal of Applied Biomedicine*. "The significance of stem cells in free-living flatworms: one common source for all cells in the adult". <<http://homepage.uibk.ac.at/~c719135/pdf/Peter-et-al-2003.pdf>>. Octubre del 2007.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech y Jan Svartvik. *A Comprehensive Grammar of the English Language*. Nueva York: Longman, 1985.
- Roberts, Paul. *Understanding English*. Nueva York: Harper & Brothers Publishers, 1958.
- Rodríguez Tarduchy, Gemma y María Concepción Santiago Martínez. *Secuenciación automática de ADN*. <<http://www.iib.uam.es/servicios/seq/otros/SecuenciaADN/SecuencaciaADN.html>>. 6 de setiembre del 2006.
- Sancho Rodríguez, Alfonso. "Definición de sintagma. Tipos de sintagma". <<http://personal.telefonica.terra.es/web/apuntesasr/SintaxSN.htm>>. Octubre del 2006.
- Salleh, Anna. *News in Science*. "Genetic discrimination cases investigated". <<http://www.abc.net.au/science/news/stories/s1496914.htm>>. Octubre del 2007.

- Saria, Tesi. *Basque Research*. “Análisis clínico-genético de las diastrofias musculares LGMD2A y desarrollo de un prototipo DNAarray para la rápida detección de mutaciones”. 2007. <http://www.basqueresearch.com/tesia_irakurri.asp?Kodea=6&lehiaketa_urtea=2003&hizk=E>. 2 de noviembre del 2007.
- Shannon, William M., Calvin G. Barnes y M.E Bickford. *Oxford Journals*. Grenville Magmatism in West Texas: Petrology and Geochemistry of the Red Bluff Granitic Suite. <http://www.oxfordjournals.org/our_journals/petroj/online/Volume_38/Issue_10/pdf/ega057_gml.pdf>. 1 de noviembre del 2007
- Stockwell, Robert P., J. Donald Bowen., John W. Martin. *The Grammatical Structures of English and Spanish*. Chicago: The University of Chicago Press, 1965.
- Timken, Mark D., Swango, Katie L., Orrego, Cristián y Martin R. Buoncristiani. “A Duplex Real-Time qPCR Assay for the Quantification of Human Nuclear and Mitochondrial DNA in Forensic Samples: Implications for Quantifying DNA in Degraded Samples.” *Journal of Forensic Sciences*. L., setiembre del 2005, vol. 50, No. 5.
- Wait, A. Dallas. *Environmental Forensics*. “Evolution of Organic Analytical Methods in Environmental Forensic Chemistry”. 2000. <<http://www.idealibrary.com>>. 22 de abril del 2007.
- Weinreich, Uriel. *Languages in Contact: findings and problems*. Hague: Mouton & Co. N.V, 1970.

Apéndice:

Textos originales

AN OVERVIEW OF HUMAN DNA ANALYSIS IN FORENSICS

Sasha Miller^{1,2}, Sheria King^{1,3} and Kalai Mathee^{1*},

¹Department of Biological Sciences, Florida International University,
Miami, FL 33199, USA

Current address:

²Consultant DNA Analyst, Forensic Science Center, Culloden Road, St. Michael,
Barbados

³Royal Bahamas Police Force, East Street, Nassau, Bahamas.

Running Title: Human DNA Forensics

Key Words: VNTR, STR, Y-STR, SNP, MtDNA, PCR

* Correspondence:

Dr. Kalai Mathee, Ph.D.

Profesora adjunta

Department of Biological Sciences

Florida International University

University Park

Miami, FL 33199

Tel. No: 305 348 1261

FAX: 305 348 2913

E.mail: matheek@fiu.edu

AN OVERVIEW OF HUMAN DNA ANALYSIS IN FORENSICS

Abstract: Genetic variations among individuals of a particular species are a fundamental fact of life. However, the commonalities between genomes, even across related species, are very striking. The work of Alec Jeffreys, which recognized the commonalities and differences found in human DNA that provide a unique fingerprint of an individual, led to the introduction of DNA analysis into the legal arena. Since then, DNA analysis techniques have continued to evolve and become more discriminatory. This article will give an overview of the sequence features that are commonly used by forensic analysts for human DNA identification. These include variable number of tandem repeats (VNTR), short tandem repeats (STR), mitochondrial DNA, Y-chromosomes (or Y-STRs) and single nucleotide polymorphisms (SNPs). Thus, DNA forensics can be defined as the use of measured molecular variations to identify and distinguish between individual samples. The legal admissibility of DNA profiles as evidence has been expedited by (i) the establishment of the national DNA databases, such as Combined DNA Information System (CODIS) and human SNP reference database (dbSNP), and (ii) the commercial development of robust and reproducible multiplex kits. These progresses have enabled the DNA profiling technique to emerge as a potent and priceless tool for criminal investigations, either for incriminating or for exonerating individuals.

INTRODUCTION

Forensic DNA analysis has largely been directed at human identification [1]. As a result of human genome sequencing, it has been determined that the human genome comprises about 3×10^9 nucleotide base pairs, and the vast majority of human DNA (99.9%) is shared between individuals [2, 3]. However, there are regions of DNA, in the case of humans 0.1% (one base pair out of every 1000), that have sufficient polymorphism (or makes an individual unique) to be used as a tool for individual identification [4]. These polymorphisms serve as the basis for DNA profiling. Hence, DNA forensics can be defined as the use of measured molecular variations to identify and distinguish individual samples from one another.

Forensic DNA evidence made its first appearance in the US court system in 1986 in *Florida v Andrews* case and DNA has remained the focus of public attention ever since [5]. During these years, forensic DNA analysis has experienced tremendous progress towards the development of robust, efficient and precise protocols. Consequently, results of forensic DNA profiling are now largely accepted in courts as crucial evidence. In addition, powerful databases of forensic DNA profiles have been implemented in many countries. The databases, such as Combined DNA Index System (CODIS), help to solve crimes and bring criminals to justice, as well as aid in freeing people that have been wrongly accused [6]. DNA analyses are also used in other types of human identification such as paternity testing, missing person's investigations, identification of unknown soldiers and investigations of mass disaster scenes. Here, we review the techniques that are commonly used by forensic analyst for human identification.

Variable number of tandem repeats (VNTRs)

Eukaryotic genomes are full of repeated DNA sequences of varying sizes that are usually designated by the length of the core repeat unit and the number of contiguous repeat units, or the overall length of the repeat region [7, 8]. These regions, usually referred to as satellite DNA, are dispersed throughout the genome and are categorized as either mini- or micro-satellites. Minisatellites, also known as variable number of tandem repeats (VNTR), consist of core repeat units in the range of approximately 8-100 bases in length (bp) [9]. Typically, the VNTR region represents a DNA fragment of a few thousand bases. On the other hand, short tandem repeats (STRs) or

microsatellites, which are smaller versions of the VNTR sequences, are DNA regions with core repeat units that are 2-7 bp in length, with a total length generally between 100 and 450 bp [6, 10].

Alec Jeffreys was the first scientist to discover that certain regions of DNA contained sequences that were repeated in tandem [11, 12]. He also found that the number of repeats present at any site could differ between individuals, and this formed the basis of DNA fingerprinting [11, 12]. Consequently, Jeffreys developed a technique to examine the length variation of these DNA repeat sequences or VNTRs located between two restriction sites. This technique involved the use of appropriate restriction enzymes to cut the regions of DNA surrounding the VNTRs generating fragment lengths that vary between individuals, and was therefore called restriction fragment length polymorphism (RFLP). The VNTR fragments were separated using slab gel electrophoresis and their presence detected with the Southern blot technique. A Southern blot involves transfer of the DNA from a gel to a nylon membrane and then labeled with radioactive or luminescent probes for detection. DNA variation is then exhibited in the form of different fragment lengths or alleles, or a variety of possibilities at a particular locus [13].

VNTR technology offers greater stability than previously used forensic protein and serological methods [6]. Blood typing (ABO group) determination was the first genetic tool used for distinguishing between individuals. Though it was a rapid method, blood typing was not very discriminative since there are only four possible phenotypic groups - A, B, AB, and O [14]. In addition, 40% of the population in the United States is type O [14]. Thus, the usefulness of the ABO blood groups is limited to excluding an individual from being the source of a crime scene sample. The test was not very helpful when an inclusion was made, especially if the sample was type O. Therefore, the combination of several VNTR loci provided a higher power of discrimination between individuals, and as a result, VNTR test results are widely accepted by courts.

In DNA profiling, multiple VNTR loci are examined because the more DNA markers examined and compared, the greater the chance that two unrelated individuals will have different genotypes. Alternatively, each piece of matching information adds to the confidence that the same individual contributed the DNA samples. This was particularly useful for human identity testing since a

combination of several VNTRs allows for powerful discrimination between individuals [6, 1]. However, the VNTR method still has quite a few serious limitations. As VNTR sequences are relatively long, nondegraded genomic DNA, which is seldom encountered in forensic samples, is needed to obtain reliable data [15, 6]. At the same time, the small size differences between adjacent alleles require grouping them into multiple “bins”, which complicates the statistical analysis [6, 16]. In addition, the method is relatively time consuming.

Short tandem repeats (STR)

Over the years VNTR forensic analysis evolved into another method that uses shorter repeating sequences known as STRs or microsatellites. Microsatellite DNA is scattered throughout the human genome, which is composed of approximately three billion bp, and occur on an average of every 10,000 nucleotides [12, 17]. Over the years, a large number of STR markers have been characterized and compiled in a Short Tandem Repeat DNA Internet Database, frequently referred to as STRBase [18]. Since its development, STRBase served as an information resource on commonly used STR markers for the forensic DNA typing community [18].

STR sequences are named systematically by the length of the repeat unit. For instance, dinucleotide and trinucleotide repeats have two and three nucleotides repeated next to each other over and over. Currently, among the various types of STR systems, tetranucleotide repeat sequences are most commonly used for forensic purposes [6]. Tetranucleotide repeats are highly polymorphic, facilitate easier analysis, whereas penta- and hexanucleotide repeats are less common in the human genome [1, 19, 20, 21, 22].

The introduction of polymerase chain reaction (PCR) has revolutionized the field of molecular biology. Since its initial description by Kary Mullis in 1986, forensic and DNA typing laboratories have greatly benefited from the ability to make million of copies of a specific sequence of DNA, in as little as a few hours [23]. The PCR DNA amplification technology is well suited for analysis of forensic DNA samples because it is sensitive, rapid, and not limited by the quality of the DNA as is the RFLP methods [1, 6, 12]. This is very useful since DNA from crime scenes is usually limited in both quantity and quality. Amplified STR fragments were initially detected on a slab gel by a silver

staining method. However, this technique was replaced by more sensitive and convenient fluorescence detection methodology [24, 25]. Fluorescent labeling of the amplified PCR product is accomplished by incorporating a fluorescent dye into the amplicon through a 5'-end labeled primer, after which the various fluorescently labeled STR fragments are analyzed by length based separation. The separation methods used today include slab gel and capillary electrophoresis (CE) [24].

Multiple STR fragments can be amplified simultaneously, or multiplexed, as PCR can be carried out in one tube containing different primers with different fluorescent tags [6, 26, 27]. In addition, multiplexing allows for consecutive simultaneous detection of these fragments in one lane [6]. STR fragments are analyzed either off-line after completing slab gel electrophoresis using fluorescent scanners, or by means of the slab gel-based or capillary electrophoresis-based sequencers [20, 28, 29]. The two primary instrument platforms used in the United States for fluorescence detection of STR alleles are the ABI 310 Genetic Analyzer and the Hitachi FMBIO II [6]. Multiplexed STRs involve sensitive fluorescence detection. This system is valuable because they are amendable to automation, which is an important benefit as the demand for DNA testing increases [20]. When sites on multiple chromosomes are examined, STRs are highly discriminating between unrelated and even closely related individuals [16, 27]. In addition, discrete alleles make analysis of results easier to interpret and compare through the use of computerized DNA databases, than RFLP-based systems where similar DNA sizes are grouped or 'binned' together [1, 6, 16].

For DNA typing to be effective across a wide number of jurisdictions, a common set of standardized markers were characterized. The FBI laboratory established 13 core STR loci for inclusion within the national DNA database known as CODIS (Table 1) [6, 30, 31]. When all 13 CODIS core loci are examined, the average random match probability is rarer than one in a trillion among unrelated individuals [16, 27, 31, 32]. Today, both PE Applied Biosystems and Promega Corporation have STR kits that address the needs of the Forensic DNA typing community, that cover a common set of STR loci (Table 2) [9, 33, 34].

With the invent of PCR technology, which requires small amounts of template DNA (typically, 1 ng), it is possible to analyze partially degraded DNA as the fragments of interest are short and likely

to be intact as compared to VNTRs [6]. In addition, allelic dropout of larger alleles in minisatellite markers, which is caused by preferential amplification of the smaller alleles, is eliminated from STR analysis [1]. This is due to the fact that both STR alleles from a heterozygous individual are similar in size as the repeat size is small. Furthermore, the accuracy of the technology to detect the fragment lengths is constantly improving. These days, fluorescent capillary and slab gel electrophoresis provide one basepair resolution with fragment lengths up to 500 nucleotides [24]. Therefore, for both biology and technology reasons, the smaller STRs are more advantageous than the larger minisatellite VNTRs. To date, STR analysis remains the chief tool in forensic DNA identification, and will probably continue to be used for many years in the future since they are part of the DNA databases that are continuously growing worldwide.

Mitochondrial DNA (mtDNA) analysis

When the conventional STR multiplexes generate inconclusive results with extremely degraded samples, mitochondrial DNA (mtDNA) may be used to obtain helpful information. This is due to the fact that mtDNA is more resistant to complete sample degradation as it occurs in higher copy numbers per cell than nuclear DNA [1, 6]. The human mtDNA sequence was first determined in 1981 in the laboratory of Frederick Sanger [35]. Human mtDNA contains a small circular genome of 16,569 bp, which like the nuclear genome, consists of both coding and non-coding regions [35]. Today the original sequence is commonly referred to as the Anderson sequence or the Cambridge reference sequence, and it serves as the reference to which new sequences are compared [36]. It was not until 1993 that mtDNA was first used as a tool for forensic identification, and since its debut mtDNA analysis has become widely accepted in the courtroom [37, 38].

MtDNA is a useful marker for human identity testing because it has a 1100 bp non-coding “control region”, also known as the displacement loop (D-loop), which exhibits a fair degree of sequence variation between individuals [1, 6]. Two hypervariable regions within the D-loop have been identified and are commonly known as hypervariable region I (HV1) and hypervariable region II (HV2). Since most of the mtDNA polymorphisms are scattered throughout the hypervariable regions of the D-loop, forensic DNA scientist examine both HV1 and HV2 by PCR amplification followed by sequence analysis [1, 6, 39].

MtDNA can be very helpful in forensic cases involving severely degraded samples or to associate maternally related individuals [39, 6]. As mentioned previously, there are more copies of mtDNA per cell, a fact that gives mtDNA an advantage over single copy nuclear DNA [1, 6]. Therefore, the use of mtDNA as an alternative marker provides forensic DNA scientists with a means to extract valuable information from samples that have strikingly low quantities of DNA, as well as biological material that has been subjected to extreme environmental conditions [40]. In addition, mtDNA is maternally inherited [6, 39]. The sperm does not contribute any mitochondria to the zygote during fertilization. Thus, barring mutation, an individual's mtDNA type is not unique. Siblings and maternal relatives share an identical mtDNA sequence. This reduces the significance of a match in forensic cases but very useful in solving missing persons or mass disaster investigations. To date, mtDNA evidence has played a valuable role in linking families and solving historical puzzles such as identifying the Vietnam Unknown Soldiers and the Romanov family, respectively [9, 40, 41].

Due to the increased labor and expense compared to STR typing, mtDNA sequence comparisons are mainly used in forensic cases where nuclear DNA markers fail to obtain results [6, 37, 42, 43]. One of the major limiting factors in widespread implementation of mtDNA in forensic casework is the existence of heteroplasmy, the presence of more than one mtDNA type in an individual [6, 44]. This is further complicated by the choice of tissues and number of samples extracted; for example different hair roots from the same individual exhibit variable levels of heteroplasmy [45, 46].

A database of mtDNA sequences of HVI and HVII, of more than 1000 unrelated individuals, has been compiled from multiple population groups [47]. In addition, a human mtDNA standard reference material (SRM) has been developed in an effort to streamline mtDNA analysis [48, 49]. These databases provide quality control between laboratories when mtDNA sequences are used for forensic identification. The growing size of the databases is significant because the mtDNA sequence is treated as a single locus. As a result, the current practice of conveying the rarity of a mtDNA type among unrelated individuals is to count the number of times a particular sequence is found in a database [50]. Thus, the larger the number of unrelated individuals in the database, the better the statistics will be for a random match probability calculation. It has been shown that both

length and sequence heteroplasmy exists in mtDNA, and this can be used to increase the discriminating power of the analysis, thus strengthen mtDNA evidence [6, 51].

Y-chromosome systems

MtDNA is not the only lineage marker available to forensic DNA analyst. Y-chromosome markers are also passed down from generation to generation without changing, except in the case of a mutational event [1, 6]. For this reason, Y-chromosome markers are useful for discerning information about paternal lineages. It was not until the early-90s that Y-STR polymorphisms were described, and became a popular method for male identification in forensic casework [52, 53]. Due to the fact that the Y chromosome is passed down from father to son unchanged, it is not as variable between individuals as the autosomal STRs. Consequently, multiple Y-STR markers are required to obtain a high degree of discrimination between individuals. Until recently, the use of the Y-chromosome for forensic purposes was restricted by a lack of polymorphic markers [6]. To date, however, a number of Y-chromosome STRs have been identified and are being developed into robust and reproducible multiplex kits [54, 55, 56, 57]. In a recent detailed analysis Hanson and Ballantyne reported the potential use of at least 389 separate Y-STR markers [58]. The Promega Corporation has a Y-STR kits that amplifies 12 Y-STR (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439) loci [9, 33, 34]. Since commercialized kits for Y-STR typing are just making into the market, utilization in routine forensic casework is limited [59]. Y-STRs are especially suitable for investigation of sexual assault cases because they provide a male-specific DNA profile, which eliminates problems of mixed stain interpretation [1, 6]. Thus, Y-STR typing is a promising technique for forensics in cases where paternal lineage is important, such as personal identification for victims of mass disasters [1, 6].

In addition, Y-STR markers are extremely helpful for paternity testing when a male offspring is in question [52, 60]. As a matter of fact, the ability of Y-STR markers to trace paternal lineage was demonstrated in 1998 by linking modern-day descendants of Thomas Jefferson and Eston Hemmings [61]. The prestigious scientific journal *Nature* published a report entitled 'Jefferson fathered slave's

last child', and the study used Y-chromosome DNA markers to trace Jefferson's male line to a descendant of Sally Hemming's youngest son, Eston Hemmings.

Table 3 contains a summary of the 10 Y-STRs that have been most widely studied for forensic genetic purposes [55]. These markers differ in their power of discrimination and their ability to differentiate between two unrelated male individuals. In addition, there is an accessible internet database containing population data on multiple Y-STR markers that is accessible via the web [62].

Single nucleotide polymorphisms (SNP)

Sequence variation between individuals at a single site in the genome is often referred to as a single nucleotide polymorphism, and Venter *et al.* determined that a million or more SNPs exist per individual [3]. The analysis of multiple SNPs creates a unique profile for that person. Hence, the examination of SNPs has the potential to be a vital part of the future of discriminating between individuals. The study of SNPs is of particular interest to the forensic DNA community for several reasons.

First and foremost, according to Li *et al.*, due to easier result interpretation, SNPs have displayed the ability to easily differentiate between mixtures and single source samples than do STRs [63]. Second, a number of technologies are under development to miniaturize and automate the SNP analysis procedure [1]. For example, a microchip-based assay has been described where more than a thousand SNPs can be examined simultaneously [64]. Therefore, SNPs can be potentially multiplexed to a higher level than STRs. Third, the sample processing and data analysis may be more fully automated because a size-based separation is not needed [63, 65]. Finally, due to the short length of the SNP PCR markers, they would be able to withstand degraded DNA samples better than the larger STR amplicons [1]. However, because each SNP locus possesses only two possible alleles, more markers are needed to obtain a high discriminatory power [31]. That is, though SNPs are more common in the human genome, because they are not as polymorphic, many SNPs have to be typed to achieve the equivalent random match probabilities as the 13 core STR loci [1, 31]. A number of Y-chromosome single nucleotide polymorphism (SNP) markers have been identified, and their usefulness in combination with Y-STR markers for forensic purposes is

currently under investigation [66]. Furthermore, SNPs can be population-specific because of their low mutation rate [67, 68, 69]. Therefore, in the future, it may be possible to predict a perpetrator's ethnic origin with the analysis of a few population-specific SNP loci. Table 4 contains a comparison of SNP and STR markers.

The National Center for Biotechnology Information (NCBI) released a human SNP reference database (dbSNP) in November 2001 that is accessible via the web at <http://www.ncbi.nlm.nih.gov/SNP/>. In addition, The Human Gene Mutation Database (HGMD) provides detailed SNP information that is available to the scientific community at <http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html>.

CONCLUSION

There are a number of new techniques on the horizon for human genetic studies that forensic DNA scientist could potentially use for human identity testing. However, according to the Report of National Commission on the Future of DNA Evidence, the use of mtDNA, Y-chromosome and SNP markers (as well as STRs) will increase in the near future [70]. It is believed that electrophoresis-based STR analysis will remain the method of choice in the forensic community for the foreseen future, though the format of the method may eventually change to more compact and potentially portable micro-devices [1, 6]. The future of human forensic DNA analysis is moving towards increased throughput of sample processing since the forensic community is reaching a point where the time to review, interpret and report DNA test results is taking longer than the actual laboratory testing itself [1].

Ever since DNA profiling was introduced to the field of forensics, DNA evidence has continually proven itself as a reliable and invaluable tool for discerning information in criminal investigations. Forensic DNA laboratories are constantly researching new techniques for human identification. However, developing new methods is only half the challenge. After the technique has proven to be scientifically sound and generally accepted by the scientific community, the new methodology needs to be accepted by the courts [38, 71]. Until lately, human DNA evidence has been the main focus. However, in recent years, non-human DNA, obtained from plant, animal and viruses, made its first

appearances in court [72, 73, 74]. It is clear that forensic DNA laboratories are determined to use DNA to its full potential. However, in comparison non-human DNA forensics has been neglected.

Table 1. Summary of the 13 CODIS core STR loci [6, 30, 31].

Locus Name	Repeat Motif	GenBank Accession^a	Allele in GenBank	Allele Range^b	Number of Alleles seen
CSF1PO	TAGA	X14720	12	6-16	15
FGA	CTTT	M64982	21	15-51.2	69
TH01	TCAT	D00269	9	3-14	20
TPOX	GAAT	M68651	11	6-13	10
VWA	[TCTG][TCTA]	M25858	18	10-24	28
D3S1358	[TCTG][TCTA]	-	-	9-20	20
D5S818	AGAT	G08446	11	7-16	10
D7S820	GATA	G08616	12	6-15	22
D8S1179	[TCTA][TCTG]	G08710	12	8-19	13
D13S317	TATC	G09017	13	5-15	14
D16S539	GATA	G07925	11	5-15	10
D18S51	AGAA	L18333	13	7-27	43
D21S11	[TCTA][TCTG]	AP000433	29	24-38	70

^a GeneBank sequence information for a particular STR locus may be accessed online at <http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank> by entering the accession number shown.

^b Numbers in this column refer to the number of repeat units present in the alleles.

Table 2. Commercially available STR multiplexes (fluorescently labeled) [9, 33, 34].

Name	Source	STR Loci Included	Power of Discrimination
TH01, TPOX, CSF1PO monoplexes	Promega	TH01, TPOX, CSF1PO	1:410
AmpFISTR Blue	PE Applied Biosystems	D3S1358, VWA, FGA	1:5000
AmpFISTR Green I	PE Applied Biosystems	Amelogenin, TH01, TPOX, CSF1PO	1:410
CTTv	Promega	TH01, TPOX, CSF1PO, VWA (vWF)	1:6600
FFFL	Promega	F13A1, FES/FPS, F13B, LPL	1:1500
GammaSTR	Promega	D16S539, D13S317, D7S820, D5S818	1:1.8x10 ⁴
PowerPlex (version 1.1 & 1.2)	Promega	TH01, TPOX, CSF1PO, VWA, D16S539, D13S317, D7S820, D5S818	1:1.2x10 ⁸
AmpFISTR Profiler	PE Applied Biosystems	Amelogenin, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820, D3S1358, FGA, VWA	1:3.6x10 ⁹
AmpFISTR Profiler Plus	PE Applied Biosystems	Amelogenin, D5S818, D13S317, D7S820, D3S1358, FGA, VWA, D8S1179, D21S11, D18S51	1:9.6x10 ¹⁰
AmpFISTR COfiler	PE Applied Biosystems	D3S1358, D16S539, Amelogenin, TH01, TPOX, CSF1PO, D7S820	1:8.4x10 ⁵
AmpFISTR SGM Plus	PE Applied Biosystems	D3S1358, VWA, D16S539, D2S1338, Amelogenin, D8S1179, D21S11, D18S51, D19S433, TH01, FGA	1:3.3x10 ¹²
PowerPlex 2.1	Promega	D3S1358, TH01, D21S11, D18S51, VWA, D8S1179, TPOX, FGA, Penta E	1:8.5x10 ¹⁰
PowerPlex 16	Promega	CSF1PO, FGA, TPOX, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, Amelogenin	1:1.8x10 ¹⁷

Table 3. Summary of the 10 Y-STRs that have been most widely studied for forensic purposes [55].

Marker Name	GenBank Accession	Repeat Motif	Allele Range	PCR Product Sizes
DYS19	X77751	TAGA	8-16	178-210 bp
DYS385	Z93950	GAAA	10-22	252-300 bp
DYS388	G09695	ATT	12-17	128-143 bp
DYS389 I	G09600	(TCTG) (TCTA)	I: 7-13	239-263 bp
DYS389 II	G09600	(TCTG) (TCTA) (TCTG) (TCTA)	II: 23-31	353-385 bp
DYS390	G09611	(TCTG) (TCTA) (TCTG) (TCTA)	18-27	191-227 bp
DYS391	G09613	TCTA	8-13	275-295 bp
DYS392	G09867	TAT	7-16	236-263 bp
DYS393	G09601	AGAT	9-15	108-132 bp

Table 4. Comparison of STR and SNP markers

Characteristics	STRs	SNPs
Occurrence in human genome	~1 in every 10 kb	~1 in every 1 kb
General informativeness	High	Low; only 20-30% as informative as STRs
Marker type	Di-, tri-, tetranucleotide repeat markers	Bi-allelic markers
Number of alleles per marker	Usually more than 5	Usually 2
Detection methods	Gel/capillary electrophoresis	Sequence analysis: microchip hybridization
Multiplex capability	>10 markers with multiple fluorescent dyes	Potential of 1000s on a microchip

REFERENCES

- [1.] V.W. Weedn, *Deoxyribonucleic acid: Future analytical techniques.*, in *Encyclopedia of Forensic Sciences*. Academic Press. 2000 pp 490-499.
- [2.] B.S. Shastry, SNP alleles in human disease and evolution. *J. Hum. Genet.* 47 (2002) 561-566.
- [3.] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, et al., The sequence of the human genome. *Science* 291 (2001) 1304-1351.
- [4.] J.G. Taylor, E.H. Choi, C.B. Foster and S.J. Chanock, Using genetic variation to study human disease. *Trends Mol. Med.* 7 (2001) 507-512.
- [5.] D.J. Reeder, Impact of DNA typing on standards and practice in the forensic community. *Arch. Pathol. Lab. Med.* 123 (1999) 1063-1065.
- [6.] L. Carey and L. Mitnik, Trends in DNA forensic analysis. *Electrophoresis* 23 (2002) 1386-1397.
- [7.] N. Hardman, Structure and function of repetitive DNA in eukaryotes. *Biochem. J.* 234 (1986) 1-11.
- [8.] D. Tautz and M. Renz, Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* 12 (1984) 4127-4138.
- [9.] J.M. Butler, *Forensic DNA typing: biology and technology behind STR markers*: Academic Press. 2001.
- [10.] D. Tautz, Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17 (1989) 6463-6471.
- [11.] A.J. Jeffreys, V. Wilson and S.L. Thein, Hypervariable 'minisatellite' regions in human DNA. *Nature* 314 (1985) 67-73.
- [12.] A. Raina and T.D. Dogra, Application of DNA fingerprinting in medicolegal practice. *J. Indian Med. Assoc.* 100 (2002) 688-694.
- [13.] P.M. Schneider, S. Meuser, W. Waiyawuth, Y. Seo and C. Rittner, Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. *Forensic Sci. Int.* 97 (1998) 61-70.
- [14.] AABB, *Facts about blood and blood banking*. 2003, <http://www.aabb.org/>.
- [15.] K.L. Richie, M.D. Goldsborough, M.M. Darfler, E.A. Benzinger, M.L. Lovekamp, D.J. Reeder, et al., Long PCR for VNTR analysis. *J. Forensic Sci.* 44 (1999) 1176-1185.
- [16.] H.A. Hammond, L. Jin, Y. Zhong, C.T. Caskey and R. Chakraborty, Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am. J. Hum. Genet.* 55 (1994) 175-189.
- [17.] T.J. Hudson, Joe Doupe Young Investigators Award. The Human Genome Project: tools for the identification of disease genes. *Clin. Invest. Med.* 21 (1998) 267-276.
- [18.] C.M. Ruitberg, D.J. Reeder and J.M. Butler, STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res.* 29 (2001) 320-322.

- [19.] P. Gill, C.P. Kimpton, A. Urquhart, N. Oldroyd, E.S. Millican, S.K. Watson, et al., Automated short tandem repeat (STR) analysis in forensic casework--a strategy for the future. *Electrophoresis* 16 (1995) 1543-1552.
- [20.] C.P. Kimpton, P. Gill, A. Walton, A. Urquhart, E.S. Millican and M. Adams, Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Meth. Appl.* 3 (1993) 12-22.
- [21.] M. Litt, X. Hauge and V. Sharma, Shadow bands seen when typing polymorphic dinucleotide repeats: some causes and cures. *BioTechniques* 15 (1993) 280-284.
- [22.] A. Urquhart, N.J. Oldroyd, C.P. Kimpton and P. Gill, Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. *BioTechniques* 18 (1995) 116-121.
- [23.] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich, Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51 (1986) 263-273.
- [24.] E. Buel, M.B. Schwartz and M.J. LaFountain, Capillary electrophoresis STR analysis: comparison to gel-based systems. *J Forensic Sci* 43 (1998) 164-170.
- [25.] J. Fregeau and R.M. Fourney, DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification. *BioTechniques* 15 (1993) 100-119.
- [26.] J.F. Anderson, M.J. Greenhalgh, H.R. Butler, S.R. Kilpatrick, R.C. Piercy, K.A. Way, et al., Further validation of a multiplexing STR system for use in routine forensic identity testing. *Forensic Sci. Int.* 78 (1996) 47-64.
- [27.] P. Gill, Role of short tandem repeat DNA in forensic casework in the UK--past, present and future perspectives. *BioTechniques* 32 (2002) 366-385.
- [28.] C.A. Crouse, S. Rogers, E. Amriott, S. Gibson and A. Masibay, Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. *J. Forensic Sci.* 44 (1999) 87-94.
- [29.] A.M. Lins, C.J. Sprecher, C. Puers and J.W. Schumm, Multiplex sets for the amplification of polymorphic short tandem repeat loci--silver stain and fluorescence detection. *Biotechniques* 20 (1996) 882-889.
- [30.] B. Budowle and R.C. Allen, Analysis of amplified fragment-length polymorphisms (VNTR/STR loci) for human identity testing. *Methods Mol. Biol.* 98 (1998) 155-171.
- [31.] R. Chakraborty, D.N. Stivers, B. Su, Y. Zhong and B. Budowle, The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis* 20 (1999) 1682-1696.
- [32.] B. Budowle, B. Shea, S. Niezgoda and R. Chakraborty, CODIS STR loci data from 41 sample populations. *J. Forensic Sci.* 46 (2001) 453-489.
- [33.] <http://www.appliedbiosystems.com/products>.
- [34.] www.promega.com/geneticidentity/.
- [35.] S. Anderson, A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, et al., Sequence and organization of the human mitochondrial genome. *Nature* 290 (1981) 457-465.

- [36.] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowers, D.M. Turnbull and N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23 (1999) 147.
- [37.] M.M. Holland, D.L. Fisher, L.G. Mitchell, W.C. Rodriguez, J.J. Canik, C.R. Merrill, et al., Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War. *J. Forensic Sci.* 38 (1993) 542-553.
- [38.] E. Marchi and R.J. Pasacreta, Capillary electrophoresis in court: the landmark decision of the People of Tennessee versus Ware. *J. Capillary Electrophor* 4 (1997) 145-156.
- [39.] J.M. Butler and B.C. Levin, Forensic applications of mitochondrial DNA. *Trends Biotechnol.* 16 (1998) 158-162.
- [40.] M.M. Holland and T.J. Parsons, Mitochondrial DNA sequence analysis - validation and use for forensic casework. *Forensic Sci. Rev.* 11 (1999) 21-50.
- [41.] P. Gill, P.L. Ivanov, C. Kimpton, R. Piercy, N. Benson, G. Tully, et al., Identification of the remains of the Romanov family by DNA analysis. *Nat. Genet.* 6 (1994) 130-135.
- [42.] M.N. Gabriel, E.F. Huffine, J.H. Ryan, M.M. Holland and T.J. Parsons, Improved mtDNA sequence analysis of forensic remains using a "mini-primer set" amplification strategy. *J. Forensic Sci.* 46 (2001) 247-253.
- [43.] K.M. Sullivan, R. Hopgood and P. Gill, Identification of human remains by amplification and automated sequencing of mitochondrial DNA. *Int. J. Legal Med.* 105 (1992) 83-86.
- [44.] K. Bender, P.M. Schneider and C. Rittner, Application of mtDNA sequence analysis in forensic casework for the identification of human remains. *Forensic Sci. Int.* 113 (2000) 103-107.
- [45.] K.E. Bendall, V.A. Macaulay and B.C. Sykes, Variable levels of a heteroplasmic point mutation in individual hair roots. *Am. J. Hum. Genet.* 61 (1997) 1303-1308.
- [46.] K. Sekiguchi, K. Kasai and B.C. Levin, Inter- and intragenerational transmission of a human mitochondrial DNA heteroplasmy among 13 maternally-related individuals and differences between and within tissues in two family members. *Mitochondrion* 2 (2003) 401-414.
- [47.] B. Budowle, M.R. Wilson, J.A. DiZinno, C. Stauffer, M.A. Fasano, M.M. Holland, et al., Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci. Int.* 103 (1999) 23-35.
- [48.] B.C. Levin, H. Cheng, M.C. Kline, J.W. Redman and K.L. Richie, A review of the DNA standard reference materials developed by the National Institute of Standards and Technology. *Fresenius J. Anal. Chem.* 370 (2001) 213-219.
- [49.] B.C. Levin, H. Cheng and D.J. Reeder, A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics* 55 (1999) 135-146.
- [50.] M.R. Wilson, M. Stoneking, M.M. Holland, J.A. DiZinno and B. Budowle, Forensic analysis of mitochondrial DNA: application of multiplex solid-

- phase - fluorescent minisequencing to high throughput analysis. *Crime Lab. Digest.* 20 (1993) 68-77.
- [51.] T. Grzybowski, Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots. *Electrophoresis* 21 (2000) 548-553.
- [52.] M. Kayser, A. Caglia, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, et al., Evaluation of Y chromosomal STRs: a multicenter study. *Int. J. Legal Med.* 110 (1997) 125-133.
- [53.] A. Perez-Lezaun, F. Calafell, M. Seielstad, E. Mateu, D. Comas, E. Bosch, et al., Population genetics of Y-chromosome short tandem repeats in humans. *J. Mol. Evol.* 45 (1997) 265-270.
- [54.] A.J. Redd, S.L. Clifford and M. Stoneking, Multiplex DNA typing of short-tandem-repeat loci on the Y chromosome. *Biol. Chem.* 378 (1997) 923-927.
- [55.] L. Gusmao, A. Gonzalez-Neira, C. Pestoni, M. Brion, M.V. Lareu and A. Carracedo, Robustness of the Y STRs DYS19, DYS389 I and II, DYS390 and DYS393: optimization of a PCR pentaplex. *Forensic Sci. Int.* 106 (1999) 163-172.
- [56.] M. Prinz, K. Boll, H. Baum and B. Shaler, Multiplexing of Y chromosome specific STRs and performance of mixed samples. *Forensic Sci. Int.* 85 (1997) 209-218.
- [57.] E.K. Hanson and J. Ballantyne, A highly discriminating 21 locus Y-STR "megaplex" system designed to augment the minimal haplotype loci for forensic casework. *J. Forensic Sci.* 49 (2004) 40-51.
- [58.] E.K. Hanson and J. Ballantyne, Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications. *Leg. Med.(Tokyo)* (2005) 1-11.
- [59.] D. Corach, L.F. Risso, M. Marino, G. Penacino and A. Sala, Routine Y-STR typing in forensic casework. *Forensic Sci. Int.* 118 (2000) 131-135.
- [60.] L. Roewer, M. Kayser, P. de Knijff, K. Anslinger, D. Corach, S. Füredi, et al., A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Sci. Int.* 114 (2000) 31-43.
- [61.] E.A. Foster, M.A. Jobling, P.G. Taylor, P. Donnelly, P. de Knijff, R. Mieremet, et al., Jefferson fathered slave's last child. *Nature* 396 (1998) 27-28.
- [62.] S. Willuweit, L. Roewer, M. Krawczak, M. Kayser and P. de Knijff, *Y-STR haplotype reference database*. 2003, <http://www.ystr.org>.
- [63.] J. Li, J.M. Butler, Y. Tan, H. Lin, S. Royer, L. Ohler, et al., Single nucleotide polymorphism determination using primer extension and time-of-flight mass spectrometry. *Electrophoresis* 20 (1999) 1258-1265.
- [64.] D.G. Wang, J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, et al., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280 (1998) 1077-1082.
- [65.] A. Linacre and D. Graham, Role of molecular diagnostics in forensic science. *Expert Rev. Mol. Diagn.* 2 (2002) 346-353.
- [66.] P.A. Underhill, L. Jin, A.A. Lin, S.Q. Mehdi, T. Jenkins, D. Vollrath, et al., Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* 7 (1997) 996-1005.

- [67.] M.J. Daly, J.D. Rioux, S.F. Schaffner, T.J. Hudson and E.S. Lander, High-resolution haplotype structure in the human genome. *Nat. Genet.* 29 (2001) 229-232.
- [68.] L.B. Jorde, W.S. Watkins and M.J. Bamshad, Population genomics: a bridge from evolutionary history to genetic medicine. *Hum. Mol. Genet.* 10 (2001) 2199-2207.
- [69.] R. Lewis, SNPs as windows on evolution. *The Scientist* 16 (2002) 16.
- [70.] Report, *National Institute of Justice Report: The future of forensic DNA testing*. 2000, National Commission on the Future of DNA Evidence: Washington, D.C.
- [71.] J.A. Auxier and H.M. Prichard, The role of the expert witness: an update. *Health Phys.* 81 (2001) 269-271.
- [72.] M. Menott-Haymond, Pet cat hair implicates murder suspect. *Nature* 386 (1997) 774.
- [73.] G. Sensabaugh and D.H. Kaye, Non-human DNA evidence. *Jurimetrics J* 38 (1998) 1-16.
- [74.] C.K. Yoon, Forensic science - Botanical witness for the prosecution. *Science* 260 (1993) 894-895.

Mark D. Timken,¹ Ph.D.; Katie L. Swango,¹ Ph.D.; Cristián Orrego,¹ Ph.D.;
and Martin R. Buoncristiani,¹ M.P.H.

A Duplex Real-Time qPCR Assay for the Quantification of Human Nuclear and Mitochondrial DNA in Forensic Samples: Implications for Quantifying DNA in Degraded Samples*

ABSTRACT: A duplex real-time qPCR assay was developed for quantifying human nuclear and mitochondrial DNA in forensic samples. The nuclear portion of the assay utilized amplification of a ~170–190 bp target sequence that spans the repeat region of the TH01 STR locus, and the mitochondrial portion of the assay utilized amplification of a 69 bp target sequence in the *ND1* region. Validation studies, performed on an ABI 7000 SDS instrument using TaqMan[®] detection, demonstrated that both portions of the duplex assay provide suitable quantification sensitivity and precision down to 10–15 copies of each genome of interest and that neither portion shows cross-reactivity to commonly encountered non-human genomes. As part of the validation studies, a series of DNase-degraded samples were quantified using three different methods: the duplex nuclear-mitochondrial qPCR assay, the ABI Quantifiler[™] Human DNA Quantification Kit qPCR assay, which amplifies and detects a 62 bp nuclear target sequence, and slot blot hybridization. For non-degraded and moderately degraded samples in the series, all three methods were suitably accurate for quantifying nuclear DNA to achieve successful STR amplifications to yield complete profiles using the ABI AmpF/STR[®] Identifier[™] kit. However, for highly degraded samples, the duplex qPCR assay provided better estimates of nuclear template for STR amplification than did either the commercial qPCR assay, which overestimated the quantity of STR-sized DNA fragments, leading to an increased proportion of undetected alleles at the larger STR loci, or slot blot hybridization, which underestimated the quantity of nuclear DNA, leading to an increased proportion of STR amplification artifacts due to amplification of excess template.

KEYWORDS: forensic sciences, DNA quantitation, quantitative polymerase chain reaction, degraded DNA, TH01, ND1, melt transition

The need to increase throughput in forensic DNA genotyping has led to a growing interest in developing new DNA quantification methods that are more efficient than the slot blot hybridization method currently used in many forensic DNA labs (1). The slot blot method, though it offers excellent specificity and good sensitivity (2,3), relies on a protocol that is time-consuming, labor-intensive, and not readily transferable to automation. In response, there have been a number of recent publications that describe alternative approaches for quantifying DNA in forensic samples, including a liquid hybridization assay (4,5) and several end-point PCR assays (6–8). Another approach that is proving to be useful for forensic DNA quantifications is real-time quantitative PCR (qPCR), a method widely used in biomedical research and molecular diagnostics (9–13). Quantitative PCR assays have been developed for various forensic applications, including the quantification of human nuclear DNA (14–16), human mitochondrial DNA (17–19), and human Y-chromosomal DNA (16,19). These assays exploit features of qPCR that make the technique particularly attractive for

forensic applications: (i) qPCR assays can be designed to quantify specific genomes of interest; (ii) the assays can be sensitive enough to detect only a few copies or even a single copy of target DNA; (iii) qPCR dynamic detection ranges readily span the roughly three orders of magnitude (e.g., 30 pg to 30 ng of nuclear DNA) needed for most forensic applications; and (iv) the experimental protocols for real-time qPCR quantitations are straightforward, labor-saving, and amenable to automation. Moreover, the use of target-specific detection chemistries (e.g., TaqMan[®] (20) or Molecular Beacon (21) probes) makes it possible to design multiplex, real-time qPCR assays that can simultaneously quantify more than one target in a sample, offering the possibility for saving time, labor, and extracted DNA.

In this report, we describe a duplex real-time qPCR assay for the simultaneous quantification of human nuclear and mitochondrial DNA in forensic samples. This assay was designed to be of general utility for forensic DNA quantifications, but to be particularly useful for the post-extraction analysis of samples that contain highly degraded DNA. Such samples, though not uncommon in standard casework, are more often encountered in instances of mass disasters, mass graves, and missing persons' cases (22,23). Short tandem repeat (STR) genotyping, due to its high power of discrimination for human identification, is typically the analytical method of first choice. However, the quality and/or quantity of extracted nuclear DNA in these degraded samples often preclude successful STR genotyping, resulting in partial or no STR profiles. Such samples can then be analyzed by less discriminatory typing methods based

¹ California Department of Justice Jan Bashinski DNA Laboratory, 1001 W. Cutting Blvd., Suite 110, Richmond, CA 94804.

* This work was supported, in part, by Grant Number 2002-IJ-CX-K008 awarded by the National Institute of Justice, Office of Justice Programs, US DOJ. Points of view presented are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice.

Received 2 Oct. 2004; and in revised form 30 March 2005; accepted 9 April 2005; published 3 Aug. 2005.

on polymorphisms in hypervariable subregions I and II (HVI and HVII) of the human mitochondrial control region. Presently, the initial decision as to how to proceed with analysis, either by nuclear STR or mitochondrial typing, is commonly based on a slot blot quantification approach that: (i) has been reported to underestimate the quantity of nuclear DNA in degraded samples (24) and in samples that contain high levels of microbial contamination (23); (ii) provides no information about the quality (fragment length) of the quantified nuclear DNA; (iii) provides no direct information about the quantity of human mitochondrial DNA in the sample; and (iv) has no predictive information on the existence of inhibitors that might interfere with the PCR. Due to these quantification deficiencies, the actual forensic analysis of challenging samples often begins by obtaining inadequate STR typing results, and then proceeds by using any remaining extracted DNA to attempt mitochondrial typing. It has been noted previously (17) that the efficiency and quality of this analysis procedure could be improved substantially by obtaining reliable estimates of the amounts of human nuclear and mitochondrial DNA in these samples prior to beginning analysis. Based on such estimates, an optimal analytical approach could be selected at the outset, leading directly to optimal genotyping or haplotyping results and to a concomitant savings in time, labor, reagent/kit costs and extracted DNA.

Several recent publications have addressed exactly these issues for forensic samples, particularly for degraded samples. For example, von Wurmb-Schwark et al. (25) have developed a duplex endpoint PCR assay to detect a nuclear target (164 bp amplicon at betaglobin) and a mitochondrial target (260 bp at *ND1*). In their assay, the endpoint PCR products were resolved and detected by capillary electrophoresis (CE)/fluorescence and, though the results were not claimed to be quantitative, the intensities of the nuclear and mitochondrial signals were shown to be useful for selecting the appropriate forensic analysis tool for the typing of anthropological bone samples. Alonso et al. (19,26) have developed several qPCR assays for degraded and challenging samples, including a duplex assay that simultaneously quantifies human X- and Y-chromosomes (106 bp and 112 bp targets at the amelogenin gene). They have also developed two singleplex qPCR assays for quantifying the human mitochondrial genome, one with a 113 bp HVI target and the other with a 287 bp HVI target. Due to the size difference between the two target sequences, they demonstrated that a comparison of the quantification results from these mitochondrial qPCR assays could provide information about the degree of DNA degradation. Lastly, Andreasson et al. (17) have described a nuclear-mitochondrial duplex qPCR assay. The nuclear portion of this assay quantifies a 79 bp target at the retinoblastoma (*RBI*) gene, and the mitochondrial portion quantifies a 135 bp target spanning the junction of the *tRNA^{lys}* and *ATP8* genes. Their results demonstrated that the duplex qPCR approach provides useful and sensitive DNA quantifications while saving analyst time and often-limited DNA. It has been noted elsewhere (1), however, that the choice of *RBI* as a qPCR target sequence may not be ideal because cross-species sequence homology investigations indicate that the nuclear *RBI* target sequence is relatively conserved. Consequently, though this assay can be expected to provide accurate results for forensic samples of known human origin, it has not been demonstrated to be sufficiently primate-specific to be of general forensic utility.

We describe here a new nuclear-mitochondrial duplex qPCR assay developed following the same general approach of Andreasson et al. (17), except that we have chosen alternative nuclear and mitochondrial target sequences for amplification and quantification. For the nuclear portion of the qPCR assay, we have chosen a target sequence that spans the repeat region of the primate-specific

TH01 STR locus, a locus that has been used widely for forensic applications. This target sequence is of direct interest for quantification, considering that the primary reason for quantifying human nuclear DNA in forensic samples is to determine the amount of extract to amplify with a commercial multiplex STR PCR kit (8). Our results indicate that for degraded samples our choice of the relatively long TH01 target sequence (~170–190 bp) leads to improved STR typing results, compared to results based on quantification of a short target sequence (e.g., 62 bp in the Applied Biosystems Quantifiler™ qPCR kit (16)) or via slot blot hybridization. For the mitochondrial portion the assay, we selected a relatively short target sequence (69 bp) in the mitochondrial *ND1* gene. This selection provides a sensitive means for determining the presence of human mitochondrial DNA, degraded or not, in forensic samples. In addition to describing results on quantifying DNA in degraded samples, we also cover aspects of development of the duplex qPCR assay, and details from forensic validation studies, including studies of precision, reproducibility, sensitivity, species specificity, and applications to casework-type samples.

Materials and Methods

Standards and Samples

Pre-quantified, high molecular weight, human genomic DNA extracts obtained from Promega (Female-#G1521) and from Applied Biosystems (TaqMan® Control DNA) were typically used as qPCR quantification standards. Pre-quantified DNA extracts obtained from Applied Biosystems (genomic DNA standards from the Quantifiler™ and QuantiBlot™ kits), Promega (K562 and Male-#G1471), and ATCC (HL60) were also used as control samples in several studies.

Unless otherwise indicated, DNA extracts from reference, non-probative and simulated casework samples were obtained using the California Department of Justice casework organic extraction protocol (ProK/SDS digestion, phenol/chloroform extraction, then Centricon100 (Millipore) concentration into Tris-EDTA buffer (TE⁻⁴)). For simulated sexual assault samples, differential extractions of sperm and epithelial fractions were performed using a DTT-based digestion protocol (27). For non-root hair shafts (2 cm portions), extractions were performed by either (i) a magnetic-bead-based protocol (28) or (ii) a protocol in which a tissue grinder was first used to homogenize the hair shaft with 150 µL TE⁻⁴, after which the homogenate was extracted into 50 µL of boiling 20% Chelex (Bio-Rad). DNA extracts were stored at -20°C, except for Chelex-extracted samples, which were stored at +4°C.

For degradation studies, DNA samples were fragmented by treatment with DNase I (Invitrogen). Separate tubes of high molecular weight HL60 DNA, each tube containing 1.4 µg of DNA in 5 µL TE⁻⁴, were treated with 0.5 U of DNase I, 10x DNase I Reaction Buffer, and sterile water to bring the reaction volume to 10 µL. Increasing degrees of fragmentation were achieved by allowing the tubes to digest at room temperature for periods of 1, 2, 3, 4, 5, 15, 30, 45, 60 min and overnight. At the end of each digestion period, DNase activity was quenched by adding 2 µL of 25 mM EDTA and heating each tube at 65°C for 15 min. The samples were not further purified or concentrated. A “0 min” sample of intact DNA which contained all reaction components for digestion except DNase I was treated identically to the samples in the degradation series. The degree of DNA fragmentation was assessed by gel electrophoresis (2% agarose, ethidium bromide detection) using HyperLadder I (Bioline) and Ready-Load λ-DNA/Hind III (Invitrogen) size markers.

Non-human DNA samples were purchased as pre-quantified genomic DNA extracts from several vendors: *E. coli*, *C. perfringens* (Sigma), *B. subtilis*, *S. epidermidis*, *C. albicans* (ATCC), mouse (Promega), and cat, chicken, cow, dog, fish, horse, monkey, pig, rat (Zyagen Labs, San Diego, CA).

STR Genotyping

The AmpFISTR® Identifiler™ PCR Amplification kit (Applied Biosystems) was used for STR genotyping. PCR amplifications of 1 ng of nuclear DNA in a 25 µL reaction volume were performed according to vendor instructions on a GeneAmp® 9700 PCR thermocycler (Applied Biosystems). STRs were resolved and detected on a Prism® 3100 Genetic Analyzer (Applied Biosystems) according to vendor instructions, except that the electrokinetic sample injection time was dropped from the default of 10 seconds to 5 seconds, and the data were analyzed with a baseline of 35 rather than the default of 51. Alleles were identified at a minimum threshold of 100 RFU using GeneScan® (v.3.7.1/NT) and Genotyper® (v.3.7/NT) (Applied Biosystems) for data analysis.

Mitochondrial HVI/HVII PCR

The HVI and HVII portions of the mitochondrial DNA control region were amplified in a duplex PCR using reagents from the LINEAR ARRAY™ Mitochondrial DNA HVI/HVII Region-Sequence Typing Kit (Roche Applied Science) (29,30). This kit produces nominal 444 bp (HVI) and 416 bp (HVII) amplicons. Template quantities for PCR were determined either from our nuTH01 duplex qPCR assay (using 100 pg of nuclear DNA per duplex HVI/HVII amplification) or from our mtND1 duplex qPCR assay to quantify the mitochondrial genome (using ~14,000 mitochondrial copies per duplex HVI/HVII amplification). Post-amplification yields and purities of the HVI and HVII PCR products were assessed by gel electrophoresis (4% NuSieve 3:1 (FMC) agarose gel, ethidium bromide staining) of 5 µL of PCR product using Low DNA Mass Ladder (Invitrogen) as a size and quantity marker.

Slot Blot Quantification

The QuantiBlot™ Human DNA Quantification Kit (Applied Biosystems) was used according to vendor instructions. Hybridized probes were detected by chemiluminescence (SuperSignal West Femto (Pierce)) using a CCD camera system (CCDBio 16SC (Hitachi/MiraiBio)). CCD data were analyzed semi-automatically using SlotQuant software running under the GeneTools (SynGene) analysis package. Quantification standards (DNA Standard A from the kit) ranged from 20 ng to 25 pg for each run.

Quantifiler™ qPCR Quantification

The Quantifiler™ Human DNA Quantification Kit (Applied Biosystems), hereafter referred to as “Quantifiler™ qPCR,” was used according to vendor instructions for data collection on an Applied Biosystems 7000 Prism® SDS qPCR instrument.

qPCR Assay Design

Primers and probes (Table 1) for the nuTH01 TaqMan® and mtND1 TaqManMGB® (MGB = Minor Groove Binder) single-

TABLE 1—Oligonucleotide sequences for nuTH01 primers and probe, and for mtND1 primers, probe, and mitochondrial copy number standard.

Oligonucleotide	Sequence {5' → 3'}
nuTH01-F	AGG GTA TCT GGG CTC TGG
nuTH01-R	GGC TGA AAA GCT CCC GAT TAT
nuTH01-probe	FAM-ATT CCC ATT GGC CTG TTC CTC CCT T-BHQ
mtND1-F	CCC TAA AAC CCG CCA CAT CT
mtND1-R	GAG CGA TGG TGA GAG CTA AGG T
mtND1-probe	VIC-CCA TCA CCC TCT ACA TC-MGB-NFQ
mtND1-standard	GAG CGA TGG TGA GAG CTA AGG TCG GGG CGG TGA TGT AGA GGG TGA TGG TAG ATG TGG CGG GTT TTA GGG

plex qPCR assays were designed using Applied Biosystems' PrimerExpress™ v2.0 software. In general, design guidelines were followed as recommended by Applied Biosystems (31), although the software settings in PrimerExpress™ were frequently relaxed to allow amplicon lengths to exceed the recommended maximum length of 150 bp. DNA sequences for design work were downloaded from the GenBank resource at the National Center for Biotechnology Information (NCBI) website (32). Sequence information was also obtained from the STRBase (33) and MitoMap (34) websites.

The nuclear qPCR assay (nuTH01) was designed to span the STR sequence at the human tyrosine hydroxylase (*TH01*) gene on chromosome 11 (11p15.5) using sequence from GenBank locus AF536811. In this design work, we configured PrimerExpress™ to fix the positions of the 3' ends of the qPCR primers so as to match known or deduced primers from commercial STR amplification kits (35,36). This was done so that the qPCR amplifications might more accurately predict quantifications for STR genotyping. The positions of the fluorogenic probe and the 5' ends of the primers were typically unconstrained for optimization by PrimerExpress™. The mitochondrial qPCR assay (mtND1) was designed in the *ND1* gene of the human mitochondrial genome using sequence from Genbank locus HUMMTCG (37). The *ND1* gene expresses subunit 1 of the mitochondrial NADH dehydrogenase protein.

For both the nuclear and mitochondrial assays, prior to ordering any synthetic oligonucleotides for experimental work, potential primer and probe sequences were compared to DNA sequences available through the NCBI website by using the Basic Local Alignment Search Tool (BLASTn) (38). The purpose of these comparisons was to identify and avoid inadvertent homologies of primers and/or probes to non-target genomes that might lead to undesirable cross-species reactivities. In addition, attempts were made to avoid known single nucleotide polymorphisms (SNPs) in the primer and/or probe sequences. SNP information was obtained from several sources: the NCBI SNPdB website (39), the On-Line Mendelian Inheritance in Man (OMIM) website (40), and, for mitochondrial sequences, the MitoMap website (34). Once developed, singleplex and duplex qPCR assays were optimized based on published recommendations (41,42).

nuTH01-mtND1 qPCR Quantifications

Primer and probe sequences for the nuTH01 and mtND1 qPCR assays are provided in Table 1. For nuTH01-mtND1 duplex qPCR amplifications, each assay was run as a 20 µL amplification that included 10 µL of TaqMan® Universal Master Mix 2X, no UNG (Applied Biosystems), 4 µL of sample, with the remaining 6 µL composed to give final concentrations of: 0.16 µg/µL non-acetylated

BSA (Sigma); 600 nM in each nuTH01 primer; 200 nM in nuTH01-probe; 50 nM in each mtND1 primer; 100 nM in mtND1-probe. Primers (Qiagen-Operon (Alameda, CA)) were diluted in sterile, de-ionized water. Probes (Qiagen-Operon for nuTH01, Applied Biosystems for mtND1) were diluted in TE⁻⁴.

For nuTH01 and mtND1 singleplex TaqMan[®]/TaqManMGB[®] qPCR amplifications, assays were run using the same reagents as described for the duplex qPCR assay, except to replace the non-desired primer/probe combination with sterile water or TE⁻⁴. Singleplex SYBR[®] Green I assays were typically run as 20 µL amplifications that included 10 µL of 2X SYBR[®] Green Master Mix (Applied Biosystems) or 10 µL of 2X Brilliant[®] QPCR Master Mix (Stratagene), 4 µL of sample, with the remaining 6 µL composed to give final concentrations of 300 nM in each appropriate primer.

Real-time qPCR data were collected on an Applied Biosystems Prism[®] 7000 SDS instrument controlled by a computer running version 1.0 of the 7000 SDS Collection software. The instrument was typically configured for the following run conditions: 20 µL sample volumes; 9600 emulsion mode; one 10 min 95°C polymerase activation step, followed by 45 cycles of 2-step qPCR (15 s of 95°C denaturation, 60 s of 60°C combined anneal/extension). Well-to-well variations in background fluorescence were corrected for by use of a ROX-labeled passive reference, included as part of the Applied Biosystems qPCR Master Mix for each sample. For runs that used SYBR[®] Green detection, a melt curve was collected after the final cycle of PCR extension by configuring the SDS Collection software to monitor SYBR[®] Green fluorescence as the temperature was increased (~1.8°C/min) from 60°C to 95°C.

Amplification curves were analyzed by using empirically established cycle threshold and baseline settings for each type of assay (for nuTH01, threshold = 0.15, baseline 6–18 cycles; for mtND1, threshold = 0.06, baseline 3–13 cycles). For each qPCR run, the SDS Collection software generated a linear calibration plot of CT (cycle threshold) vs. log C₀ (initial standard DNA concentration) by using amplification results from a freshly prepared dilution series of pre-quantified high molecular weight human genomic DNA standard (Promega Female or Applied Biosystems TaqMan[®]). DNA quantifications for unknown samples were interpolated from the resulting linear calibration curve. These calibration and interpolation steps are semi-automatic features of the SDS Collection software. For the nuclear qPCR assay, calibration plots were constructed using data from standard DNA dilutions containing 25, 5, 1, 0.5, 0.1, and 0.05 (in duplicate) ng of total DNA per sample. For the mitochondrial qPCR assay, the calibration plots were extended to lower quantities of template per sample (0.010, 0.001, and 0.0001 (in duplicate) ng). At least one negative control, 4 µL of TE⁻⁴ or sterile water, was included in each run.

Nuclear DNA copy numbers were estimated using the ratio of one haploid nuclear copy per 3.3 pg genomic DNA (43). Mitochondrial copy numbers were estimated using a ratio of 400 mitochondrial copies per 3.3 pg of Promega genomic standard DNA or 450 mitochondrial copies per 3.3 pg of HL60 genomic standard DNA. These mitochondrial copy number ratios were empirically estimated by running the Promega and HL60 standard DNA samples against a dilution series of quantified (UV-vis absorption) mtND1 synthetic oligonucleotide standard (see Table 1 for oligonucleotide sequence of the mitochondrial copy number standard). For each qPCR run, we used the high molecular weight genomic DNA dilution series to generate two linear calibration plots, one for the nuclear portion and one for the mitochondrial portion of the duplex assay.

Where appropriate, qPCR amplification efficiencies were determined from the slopes of the linear calibration curves (% PCR efficiency = 100[(10^(-1/slope) - 1)] (44).

Oligonucleotide Melting Profile Calculations

Melting profiles for selected TH01 alleles were calculated using MELT94, a DOS-based program available on-line (45). This program uses the theory and equations of Poland (46) and of Fixman and Freire (47) to calculate variations in thermal stability along the sequence of a DNA fragment (48). Stacked melting profiles were constructed by exporting MELT94 output data (temperatures for 50% helical:50% melted states vs. sequence number) into an Excel[®] spreadsheet.

Results and Discussion

Design and Development of the nuTH01-mtND1 Duplex qPCR Assay

In the preceding Materials and Methods section of this report we included a brief outline of the procedures used to design our singleplex nuclear (nuTH01) and mitochondrial (mtND1) qPCR assays. In this section, we provide a more detailed discussion of selected aspects of assay design and development, including: (i) choices of target DNA sequences; (ii) experimental results to establish that the singleplex nuTH01 and mtND1 qPCR assays work successfully in a duplex amplification; and (iii) the observation of unusual, but predictable, SYBR[®] Green melt curves for the TH01 qPCR target.

Selection and Design of the Nuclear TH01 Amplification Target

Because the primary reason for quantifying nuclear DNA in forensic samples is to determine the amount of template to use as input for STR genotyping, we anticipated that the quantification of an STR target would be directly predictive of success for STR genotyping. Furthermore, STR loci have been well characterized with respect to cross-species reactivity (49,50) and DNA mutation rates (51) because of their widespread use in forensic DNA analysis and in paternity investigations. Although initially we did not focus solely on the TH01 locus for assay development, an examination of the DNA sequences at each of the CODIS STR loci, in conjunction with assay design work using PrimerExpress[™], quickly identified the TH01 locus as a promising target. The TH01 locus contains suitable sequence in both STR flanking regions for design and placement of a TaqMan[®] detection probe. The TH01 STR target was also attractive because the amplicons are of sufficient lengths to place them roughly in the middle of the size range produced by the commercial STR kits, but are not long enough to overly compromise the TaqMan[®] qPCR efficiency. During development, we designed and tested a number of potential qPCR assays at the TH01 STR locus. These assays were evaluated by comparing their qPCR properties (e.g., PCR efficiency (>90%), sensitivity (low CT), and precision (low variance in CT, especially at low template quantities)) in order to identify an optimal assay, designated here as nuTH01. The target sequence for the nuTH01 qPCR assay is shown in Fig. 1, which displays the relative positions of the PCR primers, the STR repeat region, and the 5'-FAM/BHQ fluorogenic probe.

Selection and Design of the Mitochondrial NDI Amplification Target

For design of the mitochondrial qPCR assay, we selected a region of the NDI (NADH dehydrogenase subunit 1) gene corresponding to bases 3485–3553 of the Cambridge Reference Sequence (CRS) (52). This target was chosen because previous cross-species

```

GGGCAAAATT CAAAGGGTAT CTGGGCTCTG GGGTGATTCC CATTGGCCTG
CCCCGTTTTAA GTTTCACATA GACCCGAGAC CCCACTAAGG GTAACCGGAC

TTCCTCCCTT ATTTCCCTCA TTCATTCAAT CATTCAATCA TTCATTCAAC
AAGGAGGGAA TAAAGGGAGT AAGTAAGTAA GTAAGTAAGT AAGTAAGTGG
                                     ↑
ATGGAGTCTG TGTTCCCTGT GACCTGCACT CGGAAGCCCT GTGTACAGGG
TACCTCAGAC ACAAGGGACA CTGGACGTGA GCCTTCGGGA CACATGTCCC

GACTGTGTGG GCCAGGCTGG ATAATCGGGA GCTTTTCAGC CCACAGGAGG
CTGACACACC CGGTCCGACC TATTAGCCCT CGAAAAGTGG GGTGTCTCTC

```

FIG. 1—Sequence information for the nuTH01 portion of the duplex qPCR assay showing relative positions of the forward and reverse primers (horizontal arrows), the TaqMan[®] detection probe (underlined on top strand), and the (CATT)₇ STR region (underlined on bottom strand). The small, vertical arrow marks the sequence position that approximately separates the nuTH01 amplicon into low and high melting temperature domains. The position of this arrow corresponds to Sequence Number 96 in Fig. 5.

sequence homology studies (53, and personal communication with Cummings MP) had indicated this sub-region of the *ND1* gene to be significantly non-conserved, a promising feature for developing a species-specific assay, and because this region of *ND1* has few known SNPs, especially when compared to the HVI/HVII control sub-regions. Using a number of web-based searching tools and avoiding any well-known disease-associated SNPs such as the LHON-associated SNP at CRS 3460, we designed and evaluated several assays at the *ND1* target. Using the same criteria as for the nuTH01 qPCR assay, we selected an optimal mitochondrial qPCR assay, designated mtND1 (Table 1).

The mtND1 assay quantifies a relatively short (69 bp) target sequence, one strand of which is shown in Table 1 as the sequence for our mtND1 copy number standard. One advantage to selecting an assay with a short target sequence was that it allowed purchase of a synthetic oligonucleotide which could then be used as a copy number standard for the mitochondrial assay, rather than preparing a standard by cloning or PCR product purification. Also, we anticipated that a short qPCR assay would detect more sensitively the presence of human mitochondrial DNA, even in degraded DNA samples.

Development of the nuTH01-mtND1 Duplex qPCR Assay

In order to develop a successful duplex qPCR assay, it is necessary to identify reaction conditions that effectively allow two amplifications to occur independently in the same tube. The goal is to avoid the predicament in which one of the amplifications reduces the PCR efficiency of the second amplification. Such a situation can lead to a delayed amplification for the second assay, an artificially large CT value, and a corresponding underestimation of the quantity of target DNA measured by that assay. One strategy for avoiding this situation is to develop duplexed qPCR assays to run under conditions that limit PCR amplification of the more abundant target sequence, for example by limiting the primer and/or probe concentrations for this amplification (42).

For the nuTH01-mtND1 duplex qPCR assay, we expected that the ratio of mitochondrial genome copies to haploid nuclear genome copies would normally exceed 100 for relevant forensic samples. This expectation was based on literature reports of mitochondrial-to-nuclear copy number ratios for various tissue types (54,55). Consequently, our approach for developing a duplex assay relied on finding appropriate limiting conditions for the mtND1 qPCR portion of the assay, while running the nuTH01 portion under optimized conditions. The development was accomplished in several steps (42). The nuTH01 singleplex qPCR assay was first examined at a range of primer (50–900 nM) and probe (50–300 nM) concentrations to determine optimal conditions for the assay. These experiments indicated that optimal sensitivity (low CT) and precision (low standard deviation of CT) were obtained by running the nuTH01 assay with 600 nM primer and 200 nM TaqMan[®] probe concentrations. We similarly examined the mtND1 singleplex qPCR assay at a range of primer (25–900 nM) and probe (25–200 nM) concentrations to determine that adequate sensitivity and precision could be obtained by running this assay with 50 nM primer and 100 nM TaqManMGB[®] probe concentrations. Our goal of running the mtND1 assay at limiting conditions is evident from the very different primer concentrations identified for the two singleplex qPCR assays. In order to determine if these reaction conditions were suitable for duplex qPCR amplifications, we compared results obtained by quantifying identical samples in both singleplex and duplex qPCR formats (Fig. 2). This comparison indicated that the qPCR assays, whether run in singleplex or in duplex, gave nearly

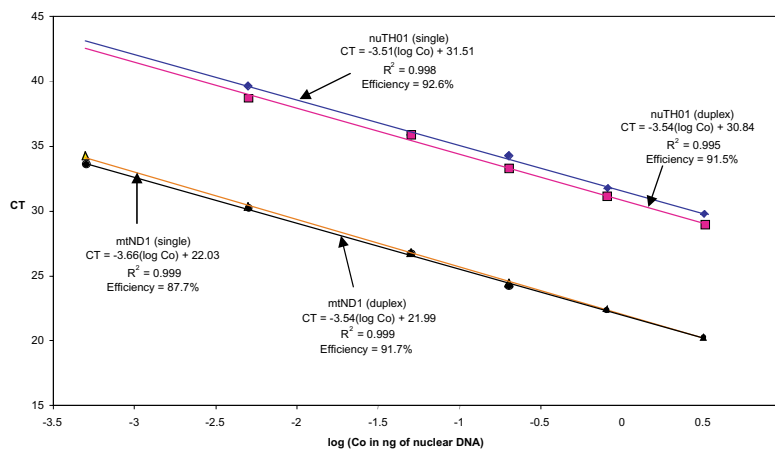


FIG. 2—Standard curves for nuTH01 and mtND1 qPCR assays run in singleplex and in duplex modes. For the nuTH01 qPCR assays, the data points shown are averages of duplicate measurements on samples containing 3.2, 0.8, 0.2, 0.05, and 0.005 ng of Promega genomic DNA (female). For the mtND1 qPCR assays, duplicate measurements on samples containing 0.0005 ng of DNA were also included. The PCR efficiency for each assay was calculated using the slope of the CT v. log C₀ standard curve.

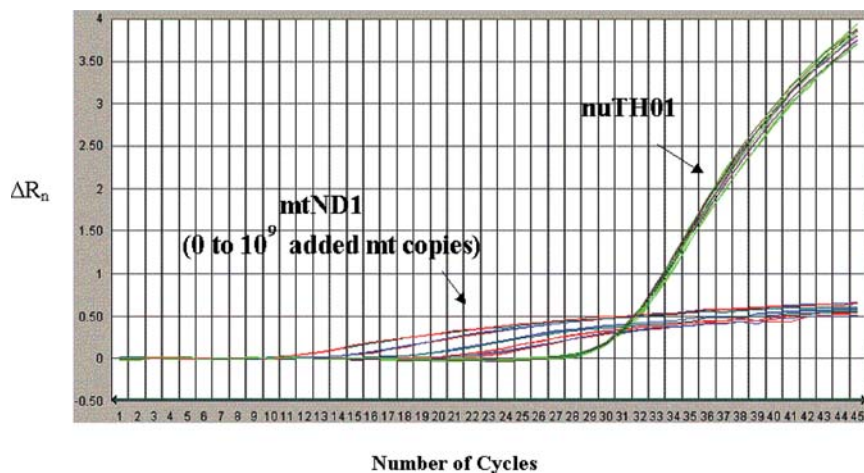


FIG. 3—qPCR amplification curves for nuTH01-mtND1 duplex assays with 3 ng of control DNA plus 0, 10^6 , 10^7 , 10^8 , and 10^9 excess copies of single-stranded mtND1-standard. Assays were run in duplicate.

identical quantifications and efficiencies. The nuTH01 qPCR assay neither lost efficiency nor showed a delayed CT when run in duplex vs. singleplex. (Due to a small amount of “leakage” of the VIC fluorescence signal into the FAM detection channel, the nuTH01 qPCR assay actually appeared to amplify with a slightly lower CT when run in duplex than when run in singleplex format.) As a final developmental step, we challenged the duplex assay by spiking a 3 ng control sample of DNA with increasing amounts of synthetic oligonucleotide mtND1 copy number standard, adding up to a one-billion-fold excess number of effective mitochondrial copies to the original DNA sample. As shown in Fig. 3, the nuTH01 amplification curves and the corresponding CT values did not change significantly, even in the presence of an unrealistic excess of mitochondrial copies.

Because we designed the duplex qPCR assay to limit the mtND1 amplification conditions while running the nuTH01 amplification under robust conditions, the duplex assay was not optimized for samples in which the number of nuclear genome copies is nearly equivalent to or greater than the number of mitochondrial genome copies. For such samples, the more robust nuTH01 amplification can lower the efficiency of the mtND1 amplification, resulting in an underestimation of the quantity of mitochondrial DNA. So far, we have encountered only one type of forensic sample—the largely tail-less sperm cells from differential extraction pellets—in which the ratio of mitochondrial genome copies to nuclear genome copies is not large. Although such samples are not commonly used for mitochondrial haplotyping, this limitation of the assay should be kept in mind when interpreting mtND1 quantification data.

TH01 SYBR[®] Green Melt Curves

The selection of a qPCR target sequence that spans the repeat region of an STR locus introduces the possibility that the resulting assay will depend in some way upon the STR genotype, for example, that the nuTH01 assay will give measurably different results for a 6,6 TH01 genotype than for a 9,9.3 TH01 genotype. We have not, within the precision and accuracy of our quantification experiments, seen any such effects for our nuTH01 qPCR assay. The only allele-dependent effect that we have seen was the observation of unusual SYBR[®] Green melt curves. While this effect has no detrimental impact on the quality of TH01 qPCR assay, it is novel and requires an explanation.

In initial developmental experiments, each potential assay was evaluated by performing qPCR runs using SYBR[®] Green I detection. This detection method allows for early identification of sub-optimal primer combinations, avoiding use of the more expensive, dye-labeled TaqMan[®] detection probes. As part of these initial assessments, we evaluated the specificity of the PCR amplifications by gel electrophoresis (4% NuSieve 3:1 agarose gel with EtBr staining) of the post-run qPCR reaction mixtures and by using the qPCR instrument to generate SYBR[®] Green melt curves. During this developmental work, for each of our potential TH01-based qPCR assays, we observed unusual SYBR[®] Green melt curves (Fig. 4).

For qPCR assays detected with SYBR[®] Green, a melt (or dissociation) curve can be generated at the end of the final PCR cycle by configuring the qPCR instrument to monitor the SYBR[®] Green fluorescence of each sample as the temperature is slowly increased from $\sim 60^\circ\text{C}$ (the extension temperature of the final PCR cycle) to $\sim 95^\circ\text{C}$. At temperatures corresponding to the melting of PCR products there will be concomitant decreases in SYBR[®] Green fluorescence due to de-intercalation of the reporter dye. In order to easily visualize these changes in fluorescence, melt curves are typically plotted as the negative first derivative of the change in fluorescence ($-dF/dT$) so that each “melt transition” will appear to be a single peak in the plot. A well-designed qPCR assay is expected to produce a single amplicon and, typically, a single transition in the SYBR[®] Green melt curve.

Figure 4, which shows the SYBR[®] Green melt curves collected for several different samples using the TH01 qPCR assay, indicates that each sample shows not one, but two melt transitions—a high melting temperature major peak ($T_m \sim 81^\circ\text{C}$) and a lower melting temperature “shoulder” ($T_m \sim 78\text{--}79^\circ\text{C}$). This observation of two melt transitions was not anticipated because the TH01 qPCR assay was otherwise unremarkable; it showed evidence for only one PCR product in post-amplification yield gels, and there was no evidence for spurious melt transitions in any negative controls. The two-step appearance of the melt transitions suggests the presence of two sequence-dependent melting domains in each TH01 amplicon. This suggestion is supported by calculations of allele-dependent melting profiles for the amplicons performed with a software package, MELT94, which uses sequence information to estimate the melting temperature of a DNA fragment along its double-stranded length (Fig. 5). Such calculations separate the TH01 amplicon into two distinct melting domains—a low melting temperature

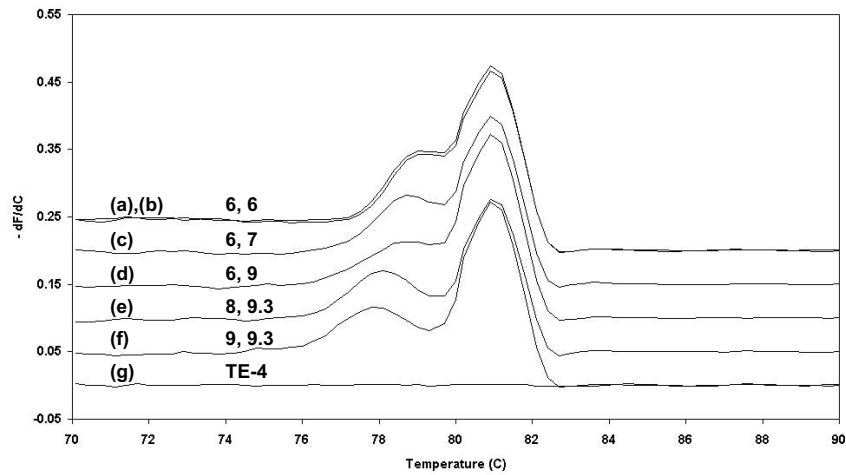


FIG. 4—SYBR[®] Green melting curves obtained for DNA extracts from six different individuals and a negative control sample using the TH01 qPCR assay. The TH01 genotype for each sample is indicated. Traces (a) and (b) are from two different individuals, both genotype 6, 6 at TH01.

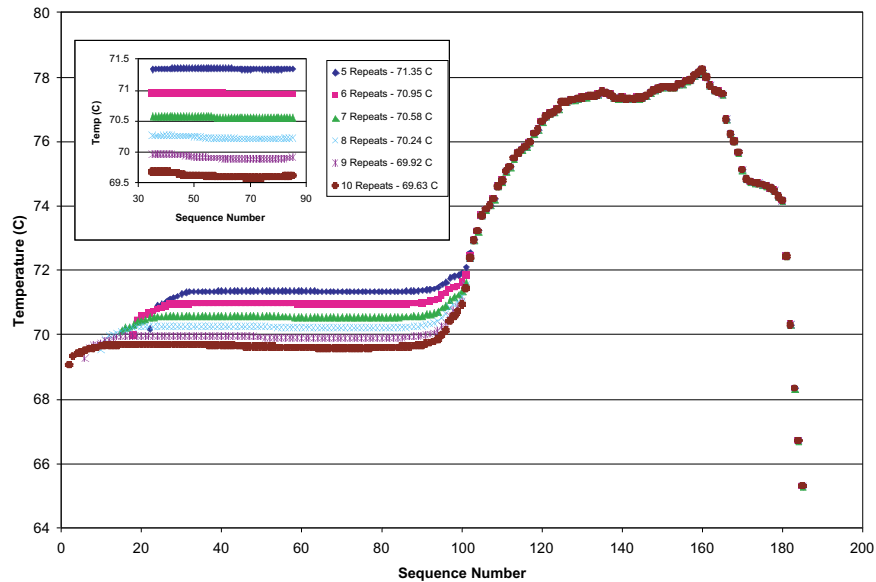


FIG. 5—Calculated melting profiles for TH01 amplicons (STR alleles 5–10) as generated by MELT94 software. Sequence Number 96 corresponds to the position of the small, vertical arrow in Fig. 1. The inset figure shows an expansion of the melting profiles in the STR region.

domain that includes the STR region (Sequence Number 0–~96) and a high melting temperature domain (Sequence Number ~96–~180). These calculations predict that the melting temperature of the low-melt domain will decrease as the number of STRs increases (Fig. 5, inset), a prediction that is in qualitative agreement with the trend seen for the experimental melting curves shown in Fig. 4. This trend is due to an increase in the AT-percentage of the low-melting domain sequence as the number of CATT repeats increases. The melting profile of the high-melt domain is predicted to be independent of the TH01 genotype (Fig. 5), consistent with the allele-independent, overlapping melt transitions ($T_m \sim 81^\circ\text{C}$) shown in Fig. 4.

The stepwise melting of small DNA fragments due to sequence-dependent melting domains has been reported (56), and similar sequence-dependent denaturation phenomena form the basis for separations by such techniques as denaturing gel gradient electrophoresis (DGGE) (57). To our knowledge, however, genotype-dependent stepwise melt transitions have not been previously reported for STR amplicons. It is intriguing that the SYBR[®] Green

melt curves for the TH01 qPCR assay might form the basis for a fairly simple and rapid, albeit low-resolution, means for STR genotyping at this locus, as well as at other STR loci that show similar allele-dependent melt curves.

nuTH01-mtND1 Duplex qPCR Validation: Precision, Sensitivity, and Reproducibility

Precision of the nuTH01-mtND1 duplex qPCR assay was assessed by running 96 identical assays in a single wellplate using 4 ng of pre-quantified standard DNA (Promega Human Genomic Female DNA) per quantification. For the nuTH01 portion of the duplex assay, we measured an average CT of 28.94 cycles with a standard deviation (SD) of 0.13 cycles and a CT range of 0.59 cycles. For the mtND1 portion of the assay, we measured an average CT of 22.63 cycles with a SD of 0.21 cycles and a CT range of 0.98 cycles. No systematic deviations in CT across rows or columns of the 96-well plate for either portion of the duplex assay were observed. Although we did not include a standard DNA dilution series in this

TABLE 2—Intra-run sensitivity and precision results for the nuTH01-mtND1 duplex qPCR assay. Results are based on five replicate quantifications of serially diluted Promega Female Genomic DNA standards. Quantities are per 4 μ L of sample. Standard deviations are in parentheses.

Input DNA Quantity (ng)	nuTH01 qPCR		Approx. Input Quantity (mt copies)	mtND1 qPCR	
	Average Quantity (ng)	RSD (%)		Average Quantity (mt copies)	RSD (%)
100	95.8 (7.7)	8.0	1.2×10^7	$8.6 (0.97) \times 10^6$	11
10	10.7 (0.90)	8.4	1.2×10^6	$1.3 (0.15) \times 10^6$	12
5	4.7 (0.088)	1.9	6.1×10^5	$7.6 (0.95) \times 10^5$	12
1	1.2 (0.034)	3.0	1.2×10^5	$1.3 (0.24) \times 10^5$	18
0.5	0.48 (0.057)	12	6.1×10^4	$7.1 (1.4) \times 10^4$	20
0.1	0.10 (0.021)	21	1.2×10^4	$1.0 (0.20) \times 10^4$	20
0.05	0.051 (0.013)	26	6.1×10^3	$7.2 (1.5) \times 10^3$	21
0.01	0.0057* (0.0017)	30*	1.2×10^3	$1.3 (0.21) \times 10^3$	17
0.005	—	—	6.1×10^2	$7.1 (2.0) \times 10^2$	28
0.001	—	—	1.2×10^2	$1.1 (0.39) \times 10^2$	35
0.0001	—	—	1.2×10^1	$1.1 (0.41) \times 10^1$	39

* Based on four replicates, because one quantification dropped out at this template quantity.

run, the standard deviation of CT can be used to estimate the relative standard deviation (RSD or %CV) of genome copy number as $100(2^{(SD \text{ of } CT)} - 1)$; this estimation assumes 100% PCR efficiency for the assay. By this means, we estimate the intra-plate RSD of genome copy number to be $\sim 10\%$ for the nuTH01 assay and $\sim 16\%$ for the mtND1 assay.

Sensitivity of the duplex qPCR assay was assessed by quantifying Promega Human Genomic Female standard DNA using template quantities ranging from 100 ng to 1 fg. All quantifications were performed in replicates of five in order to estimate the minimum level of input DNA at which stochastic effects (e.g., qPCR signal dropout and reduced quantification precision) can be expected to become significant. Results from these runs are summarized in Table 2. As expected, these results indicate that as the number of template copies diminishes, there is a general decrease in precision for both the nuTH01 and mtND1 portions of the duplex assay. For the nuTH01 portion of the qPCR assay, adequate levels of precision ($<30\%$ RSD) are seen down to approximately 50 pg of template DNA (~ 15 haploid nuclear copies), and we have generally seen similar levels of precision down to approximately 32 pg or ~ 10 copies (data not shown). At lower quantities of nuclear template, loss of precision, as well as an increase in the proportion of “dropped out” amplification curves occurred (e.g., 20% of the nuTH01 qPCR assays dropped out and were undetected at 10 pg of template, 30% at 5 pg, 80% at 1 pg, and 100% at <1 pg). This is probably due to stochastic effects in sampling and in amplification at low levels of template. For the mtND1 portion of the qPCR assay, adequate levels of quantification precision ($<40\%$ RSD) were seen down to at least 100 fg of nuclear template, which corresponds to ~ 12 mitochondrial template copies for the Promega genomic female standard DNA. All mtND1 qPCR amplification curves dropped out at 10 fg of nuclear template.

Reproducibility of the duplex assay was assessed by comparing quantifications of six different high molecular weight control DNA samples obtained from three independent qPCR runs as performed by two different analysts on two different days. Samples were quantified as single replicates in each run. The results are summarized in Table 3, where we estimate “reproducibility” as the effective RSD for each sample by pooling the quantification results from the three runs. These inter-plate RSD values are, on average, in reasonable agreement with the intra-plate RSD values reported in our discussion of assay precision and sensitivity. A notable exception is the

TABLE 3—Inter-run reproducibility results for nuTH01-mtND1 duplex qPCR quantifications of six commercial high molecular weight genomic DNA standards (samples A–F). Slot blot quantities are based on single replicate measurements for each sample. qPCR quantities are reported as averages, based on pooling the results from the three single-replicate qPCR runs. Standard deviations are in parentheses. Promega Genomic Female DNA was used as the quantification standard for all three qPCR runs.

Sample	Slot Blot Quantity (ng/ μ L)	nuTH01 qPCR		mtND1 qPCR	
		Average Quantity (ng/ μ L)	RSD (%)	Average Quantity (1000 mt copies/ μ L)	RSD (%)
A	0.85	0.64 (0.08)	12	130 (79)	62
B	0.80	0.96 (0.10)	10	880 (150)	17
C	0.34	0.53 (0.10)	20	310 (35)	11
D	1.31	1.58 (0.31)	19	430 (31)	7
E	0.97	1.51 (0.17)	11	790 (170)	21
F	0.40*	0.45 (0.05)	12	360 (73)	21

* Sample F represents the calibration standard used for the slot blot quantifications; the value for sample F (0.40 ng/ μ L) is a defined, rather than a measured quantity.

single large RSD (62%) observed for the mtND1 quantification of sample A.

nuTH01-mtND1 Duplex qPCR Validation: Species Specificity

To assess species specificity of the nuTH01 and mtND1 qPCR amplifications, the duplex assay was run using template DNA from each of fifteen different non-human species. Duplicate assays, each containing 1 ng of template, were used for each species. No cross-reactivity to non-human DNA was observed for either the nuTH01 or mtND1 portions of the duplex qPCR assay. We further challenged the duplex assay by amplifying 100 ng portions of each of the microbial DNA samples, and we observed no amplifications above the level of background in either the nuTH01 or mtND1 quantifications. These results indicate that the nuTH01-mtND1 duplex qPCR assay is sufficiently specific for forensic applications. Although we did not test the assay with DNA from higher primates, based on previous work with the TH01 STR locus (49,50) and on sequence homology investigations, we

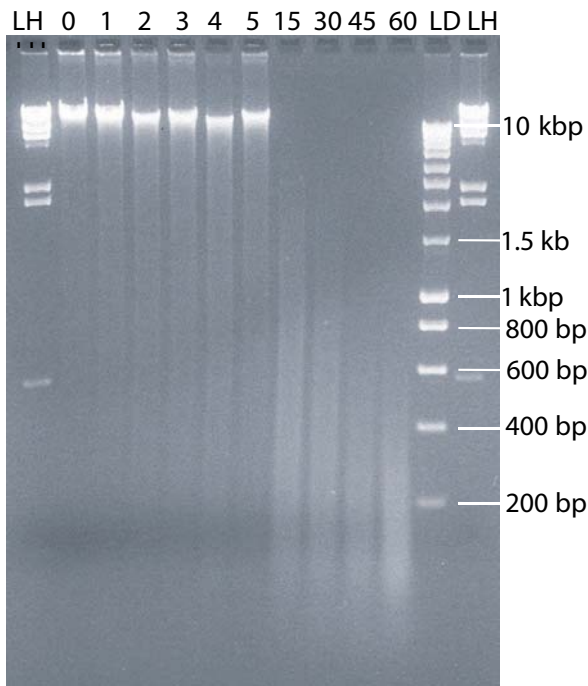


FIG. 6—HL60 DNase degradation series (EtBr-stained). Lanes labeled 0 to 60 indicate DNase treatment time in minutes. Lanes labeled LH and LD represent λ -HindIII and low-mass DNA ladder, respectively.

anticipate that the nuTH01 qPCR assay would amplify and detect such samples.

nuTH01-mtND1 Duplex qPCR Validation: Degraded DNA

Forensic evidence samples often contain DNA that has been degraded by environmental and/or microbial exposures. To assess the ability of the nuTH01-mtND1 duplex qPCR to quantify DNA in such samples, a DNA degradation series was prepared by treating aliquots of high molecular weight genomic DNA (HL60) with DNase I for increasing periods of time ranging from one minute to overnight. Figure 6 shows a gel illustrating the degree of DNA fragmentation for the DNase-treated samples. An increase in the degree of DNA fragmentation during the first 5 min of DNase treatment is evident from increased intensity of the low molecular weight “smear” in lanes 3 to 7 of the gel. During this time, the extent of degradation is best described as “moderate” given the consistent appearance of a fairly intense high molecular weight band in these lanes. For samples with at least 15 min of DNase treatment, however, there is no evidence for the high molecular weight band. For the purpose of this discussion, we describe these samples as “highly degraded.”

Samples from the DNA degradation series were quantified by three different methods: slot blot hybridization, nuTH01-mtND1 duplex qPCR, and the Applied Biosystems Quantifiler™ Human DNA Quantification kit. Because the nuTH01 assay amplifies a ~170–190 bp target sequence, while the Quantifiler™ kit amplifies a 62 bp target sequence, we included the Quantifiler™ assay in our comparison as a means to evaluate the significance of amplicon size for quantifying degraded DNA. For the qPCR assays, 2 μ L of diluted sample (1:20 in TE⁻⁴) were quantified in duplicate for each point in the degradation series. For the slot blot assays, single quantifications were obtained on 4 μ L of the same diluted samples. Quantification results (nuclear quantifications only) for the

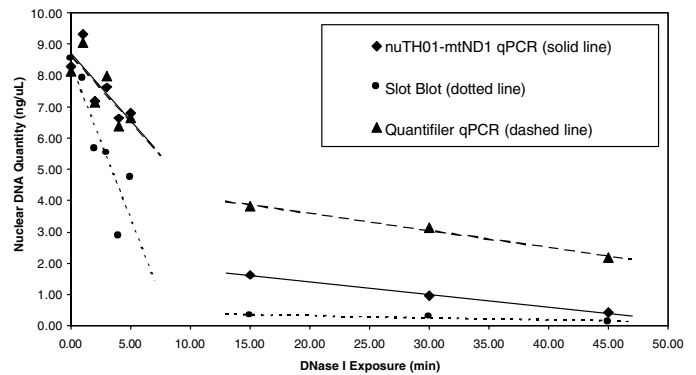


FIG. 7—Nuclear quantifications of DNase-degraded HL60 samples by nuTH01-mtND1 duplex qPCR, Quantifiler™ qPCR, and slot blot hybridization. For the qPCR assays, each point represents the average of duplicate quantifications. For slot blot, each point represents a single quantification. Lines represent linear least-squares fits to early (0–5 min) and late (15–45 min) time points in the degradation series.

three methods are summarized graphically in Fig. 7, from which a number of conclusions can be drawn. First, for undigested, high molecular weight DNA (0 min sample), the three methods provide quantifications that are in good agreement with each other, as well as in reasonable agreement with the concentration (~6 ng/ μ L) estimated for a 1:20 dilution of the DNA stock. Second, for the moderately degraded samples (1–5 min of DNase treatment), the two qPCR-based quantification methods are in good agreement, but both give quantification values significantly higher than those obtained from the slot blot method. Third, for highly degraded samples (>15 min of DNase treatment), none of the quantification methods agree. Relative to both qPCR-based methods, the slot blot method detects significantly less DNA. Even the nuTH01-mtND1 duplex and Quantifiler™ qPCR assays disagree, the latter assay detecting significantly more DNA than the former.

These results, which indicate that the quantity of DNA measured in degraded samples depends upon the quantification method used, were not entirely unexpected. There have been previous reports that the slot blot method underestimates the quantity of nuclear DNA in degraded or compromised samples (14,23,24), and it has been suggested that this effect is due to a lowered binding efficiency of the degraded DNA fragments to the slot blot membrane (14). Some degree of difference in quantifying degraded DNA by qPCR was also to be expected, considering the size difference between the target sequences for the nuTH01 and Quantifiler™ assays. The Quantifiler™ qPCR assay effectively measures the concentration of nuclear DNA fragments at least 62bp long, whereas the nuTH01 qPCR assay measures the concentration of fragments at least 170–190 bp long. For high molecular weight DNA samples these two concentrations are practically the same, but for highly degraded DNA samples there is a greater concentration of smaller than larger fragments. This concentration difference is indicated by the yield gel shown in Fig. 6 and is reflected in the divergent DNA quantities measured by the Quantifiler™ and nuTH01 qPCR assays.

The primary reason for accurately quantifying nuclear DNA in forensic samples is to ensure the correct amount of template is used for STR amplification. When too little template DNA is amplified, incomplete STR profiles can result, while amplification of excess template can cause such problems as poor inter-locus balance, increased stutter peak intensities, incomplete non-template-directed nucleotide addition, off-scale signals, and signals due to cross-dye “pull-up,” each of which can complicate or preclude accurate STR genotyping. For degraded samples, as we have seen, different

quantification methods can provide very different estimates of DNA quantity. To investigate the implications of these differences, we attempted to determine which of the three quantification methods would provide the most suitable estimates of nuclear DNA quantity for STR typing of degraded samples. To this end, we used the quantification results shown in Fig. 7 to prepare nominal “1ng” portions of nuclear DNA for each time-point in the DNase-degradation series based on each of the three quantification methods. These “1ng” portions were then amplified and genotyped using the AmpF/STR[®] Identifiler[™] STR kit.

Figure 8 provides a qualitative, graphical overview of the STR genotyping results for the entire set of DNase-degraded samples. Panel (A) in this figure suggests that for untreated and moderately degraded DNA samples (0–5 min DNase treatment), all three methods were suitable for quantifying nuclear DNA for STR genotyping. At these moderate levels of degradation, regardless of the method of quantification, all STRs were detected with signals greater than 490 RFU per locus, and there was no evidence of artifacts that would complicate or preclude accurate genotyping. For the more highly degraded DNA samples (15–60 min DNase treatment), however, success rates for STR genotyping were seen to depend upon the method of DNA quantification.

Panel (C) of Fig. 8 shows that for amplifications based on Quantifiler[™] quantifications, a relatively large proportion of alleles were undetected for all of the AmpF/STR[®] Identifiler[™] reporter dyes. As the extent of DNA fragmentation increased, the larger STR alleles for each reporter dye were the first that failed to reach the 100 RFU analytical threshold. The implications for using a short qPCR target sequence are evident. In these highly degraded samples, there were more 62 bp than longer STR-sized DNA fragments. As a consequence, the Quantifiler[™] assay overestimated the quantity of longer template fragments, leading to under-amplification of the longer STR alleles and, ultimately, to inadequate quantities of PCR product for detection. With this quantification method, at very high levels of degradation (45–60 min of DNase treatment) less than half of the Identifiler[™] STR alleles were detected, and the overestimation of DNA quantity was so pronounced that even the shorter STR alleles lost significant intensity due to under-amplification (Figs. 8C and 9B, and Table 4).

By contrast, quantification of these same samples using the nuTH01-mtND1 duplex qPCR method led to an improvement in

the success rate for STR genotyping (Fig. 8). Even for the most highly degraded samples (60 min of DNase treatment), more than 75% of the HL60 alleles were detected. For the purpose of STR genotyping highly degraded samples, these results indicate that the ~180 bp nuTH01 qPCR target is more appropriate for quantifying DNA than is the shorter 62 bp target. Data provided in Table 4 show that by basing DNA quantification on the medium-sized nuTH01 target, the intensities of the detected TH01 STRs were maintained at a fairly consistent level (~1800–3100 RFU/locus) despite the varying extent of DNA fragmentation. Even for highly degraded samples, the short (~121 bp) D19 STRs were not over-amplified to the extent that they resulted in off-scale peaks, nor were the long (~293 bp) D18 STRs under-amplified to the extent that they went undetected.

STR amplifications based on slot blot quantifications resulted in very few undetected alleles, even for the most highly degraded samples (Fig. 8). However, the electropherograms shown in panels (C) and (D) of Fig. 9 indicate the presence of several artifacts due to the over-amplification of excess template DNA. These artifacts include shoulder peaks due to incomplete non-template-directed nucleotide addition, cross-dye “pull-up” signals due to off-scale peaks in adjacent detector channels, and stutter peaks above 100 RFU. For some slot-blot quantified samples in the DNA degradation series, off-scale peaks led to pull-up signals as intense as 2700 RFU (data not shown). Off-scale peaks are evident even for samples exposed to DNase for only 30 minutes (Table 4). Inter-locus balance, as represented by the D19/D18 ratios, was significantly worse for the slot-blot quantified samples than for the qPCR-quantified samples (Table 4). Although some of these artifacts (e.g., off-scale peaks, “pull-up” signals) are possibly remedied by reducing CE sample injection times, other artifacts (e.g., shoulders due to incomplete nucleotide addition, poor inter-locus balance) are repairable only by re-amplification.

Highly degraded samples present analytical challenges for DNA quantification and for subsequent STR genotyping. The quantity and quality of DNA in such samples, which are typically comprised of fragments with a wide range of sizes, cannot be represented by any single qPCR assay, which is fundamentally biased to detect and quantify the subpopulation of fragments that are at least as long as the specific target sequence being amplified. In principle, the most accurate estimate of overall DNA quantity in degraded samples

TABLE 4—Selected AmpF/STR[®] Identifiler[™] STR intensities for DNase-degraded HL60 samples. All amplifications used nominal 1 ng nuclear template amounts as quantified by nuTH01-mtND1 duplex qPCR, Quantifiler[™] qPCR, and slot blot hybridization. STR intensities are provided as RFU per locus. For samples exposed to DNase for 0–5 min, the qPCR-based STR intensities are identical because only one STR amplification was performed based on nearly identical quantifications of these samples by the nuTH01 and Quantifiler[™] assays (see Fig. 7). Off-scale peaks, as identified by GeneScan[®] software, are indicated in bold. For HL60, the average Identifiler[™] STR allele sizes for TH01, D19, and D18 are 177 bp, 121 bp, and 293 bp, respectively (60).

DNase Exposure (min)	nuTH01-mtND1 Duplex qPCR				Quantifiler [™] qPCR				Slot Blot Hybridization			
	TH01 (RFU)	D19 (RFU)	D18 (RFU)	D19/D18	TH01 (RFU)	D19 (RFU)	D18 (RFU)	D19/D18	TH01 (RFU)	D19 (RFU)	D18 (RFU)	D19/D18
0	1824	2809	3677	0.76	1824	2809	3677	0.76	2724	2671	4780	0.56
1	1998	1851	2730	0.68	1998	1851	2730	0.68	2861	2360	3747	0.63
2	2883	2489	3253	0.77	2883	2489	3253	0.77	3636	2789	3849	0.72
3	2697	2400	3210	0.75	2697	2400	3210	0.75	3122	2465	3183	0.77
4	2417	2043	2519	0.81	2417	2043	2519	0.81	3791	2880	2637	1.09
5	2469	2119	3133	0.68	2469	2119	3133	0.68	2894	2469	3332	0.74
15	2393	3434	1537	2.23	955	1472	584	2.52	6389	7053	1566	4.50
30	3108	5156	1329	3.88	1036	1786	543	3.29	7797	9103	1766	5.15
45	2540	5408	738	7.33	534	1153	115*	...	6563	9011	903	9.98
60	1941	5249	473	11.10	349	825	5916	8845	543	16.29

* Only one STR allele of heterozygous pair detected.

A)	DNase Treatment	Percent of STR Alleles Detected	Number of STR Artifacts*	Blue (FAM) STRs								Green (VIC) STRs					Yellow (NED) STRs				Red (PET) STRs					
				D8		D21		D7		CSF		D3	THO1	D13	D16	D2	D19	vWA	TPOX	D18		AML	D5	FGA		
				12	13	29	30	11	12	13	14	16	7	8	8	11	11	17	14	16	8	11	14	15	X	12
	All Assays																									
	0 min	100	- none -																							
	1 min	100	- none -																							
	2 min	100	- none -																							
	3 min	100	- none -																							
	4 min	100	- none -																							
	5 min	100	- none -																							
B)	nuTH01-mtND1 qPCR																									
	15 min	100	- none -																							
	30 min	96	- none -																							
	45 min	84	2 P																							
	60 min	76	- none -																							
C)	Quantifiler qPCR																									
	15 min	84	- none -																							
	30 min	76	- none -																							
	45 min	48	- none -																							
	60 min	28	- none -																							
D)	Slot Blot Hybridization																									
	15 min	100	1 O, 2 P, 2 S, 3 A																							
	30 min	100	1 O, 3 P, 2 S, 3 A																							
	45 min	92	2 O, 3 P, 1 S, 3 A																							
	60 min	92	2 O, 3 P, 3 S, 4 A																							

* O = off-scale peak; P = detected pull-up signal; S = detected stutter signal; A = detected shoulder due to incomplete A addition

FIG. 8—Overview of AmpFISTR® Identifiler™ STR results obtained for DNase-degraded HL60 DNA samples. Each shaded rectangle indicates the detection of an STR allele with intensity at least 100 RFU. Un-shaded rectangles indicate alleles that were undetected at the 100 RFU analytical threshold. For each reporter dye, the Identifiler™ alleles are presented left to right in order of increasing size. STR amplifications used nominal 1 ng input levels of nuclear DNA, as quantified by the nuTH01-mtND1 duplex qPCR assay (panels A and B), by the Quantifiler™ assay (panels A and C), or by slot blot hybridization (panels A and D).

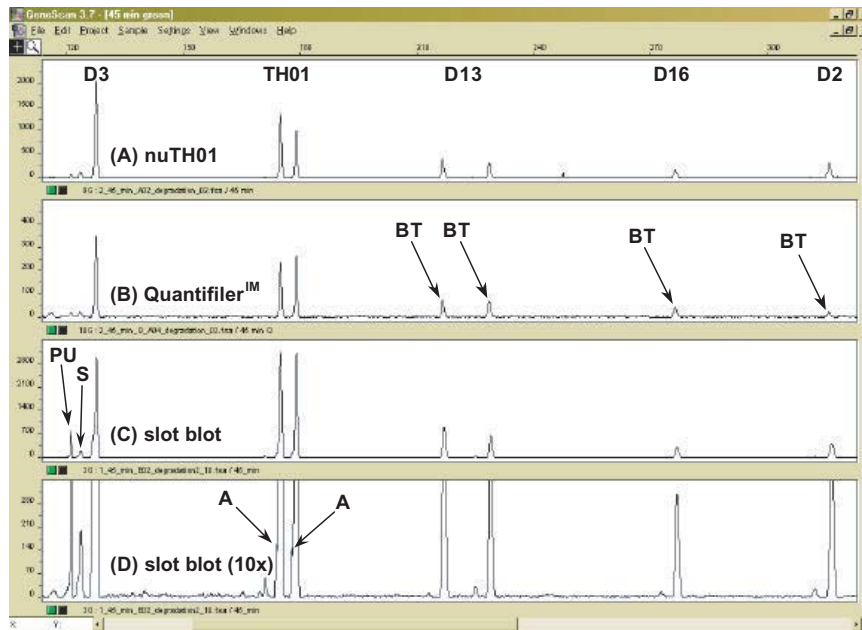


FIG. 9—AmpFISTR® Identifier™ VIC-labeled STRs from DNase-degraded (45 min) HL60 samples. Panels show STR alleles for amplifications of “1 ng” amounts of nuclear template as determined by: (A) nuTH01-mtND1 duplex qPCR, (B) Quantifiler™ qPCR, and (C) slot blot hybridization. Panel (D) is a 10-fold vertical scale expansion of panel (C). Detection and/or amplification artifacts are indicated as: BT = below detection threshold; PU = “pull-up” signal from adjacent detection channel; S = detected stutter signal; A = detected shoulder due to incomplete non-template nucleotide addition. Full vertical-scale RFU values for the panels are: (A) 2500 RFU; (B) 500 RFU; (C) 3500 RFU; (D) 350 RFU.

should be provided by qPCR assays that use very short target sequences. However, for the specific purpose of STR genotyping, the overall DNA quantity in a degraded sample is not necessarily the measurement of interest. Our results indicate that if qPCR methods are to be used to quantify degraded DNA for STR genotyping, there are advantages in selecting a qPCR target sequence that is of appropriate length to detect those DNA fragments that are most relevant to the method of analysis. This is particularly true for the highly multiplexed commercial STR genotyping kits that are used in most forensic DNA labs, kits that recommend the use of a fairly narrow range of template DNA to give optimal success rates for STR amplification and detection. For the AmpFISTR® Identifier™ STR kit, which amplifies STRs ranging from ~100 bp to ~400 bp, our genotyping results indicate that “1 ng” quantities of highly degraded DNA are better estimated by using a ~180 bp nuclear qPCR target sequence, as in the nuTH01-mtND1 duplex assay, than by using a 62 bp qPCR target sequence, as in the Quantifiler™ qPCR assay. Although this discussion is based on our experiments with a specific set of DNase-degraded samples, the general conclusion that the target length of a qPCR assay is an important consideration for quantifying DNA in degraded samples is sound. We are aware, however, that the degree of advantage gained by using a long target sequence for qPCR will likely depend on precisely how the sample is extracted and purified.

In addition to quantifying the amount of nuclear DNA, we also used the mtND1 portion of our duplex qPCR assay to estimate the quantity of mitochondrial genome copies in this same set of DNase-degraded samples. These mtND1 quantification results are provided graphically in Fig. 10. In this figure, the mitochondrial quantification results shown at each time-point in the degradation series are normalized relative to the quantification result for the untreated sample (1.3 million mt copies/ μ L). For comparison, the figure also shows normalized nuclear quantifications, as estimated by the nuTH01 and Quantifiler™ qPCR assays for the same set of

DNase-degraded samples. Notice that the mtND1 and Quantifiler™ assays generally measured higher normalized quantities of DNA than did the nuTH01 assay. For example, both the mtND1 and Quantifiler™ assays indicated that ~25% of the initial DNA remained after 45 minutes of DNase exposure, while the nuTH01 assay detected only ~5% of the initial DNA concentration. It is likely that because the mtND1 qPCR assay amplifies a relatively short target sequence (69 bp), the normalized quantifications for the mtND1 and Quantifiler™ assays were very similar, despite quantifying entirely different genomes. Because we have already seen that the selection of a short nuclear target has implications for STR genotyping, we were motivated to learn if the selection of a short mitochondrial quantification target would have a similar detrimental impact on HVI/HVII amplification success rates.

To address this issue, we performed two separate sets of HVI/HVII duplex PCR amplifications on samples from the DNA degradation series. In the first set, for each time-point in the series we amplified 100 pg of nuclear DNA, as measured by the nuTH01 portion of the duplex qPCR assay. This approach, to use the quantity of nuclear DNA to indirectly estimate the amount of template for HVI/HVII PCR amplifications, is commonly used in the forensic analysis of mitochondrial DNA, if there is sufficient nuclear DNA to be quantified (30,58). In the second set, for each time point we amplified ~14,000 mitochondrial genome copies, as estimated directly by the mtND1 portion of the duplex qPCR assay. (For pristine samples of HL60, we had previously determined that 100 pg of nuclear DNA represented ~14,000 mitochondrial genome copies.) The post-amplification yield gel results for both sets of HVI/HVII duplex PCR amplifications are shown in Fig. 11.

A “successful” HVI/HVII duplex PCR amplification was considered to produce sufficient quantities of the HVI and HVII PCR products for successful cycle sequencing. For cycle sequencing, our laboratory has validated a protocol that uses the Applied Biosystems Terminator BigDye® v1.1 Cycle Sequencing System

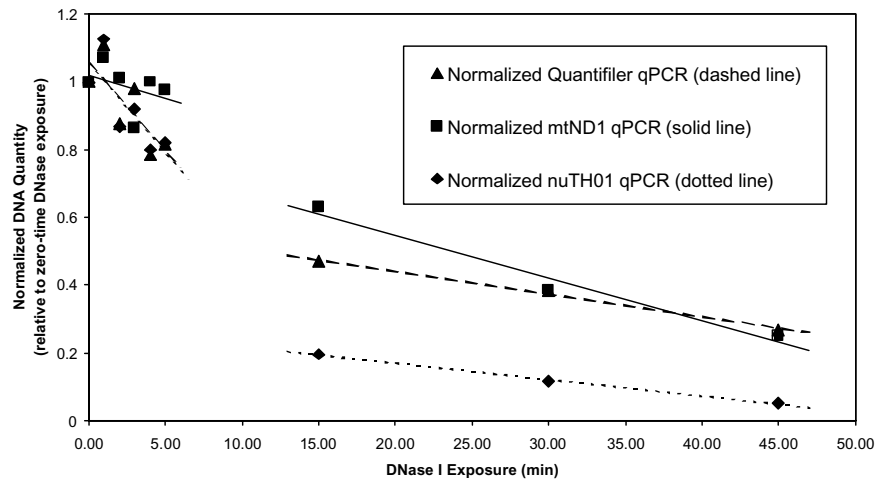
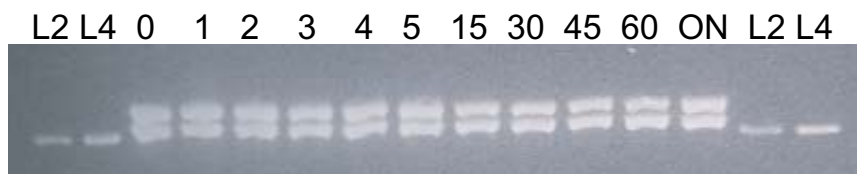


FIG. 10—Normalized quantifications of DNase-degraded HL60 samples by mtND1 (duplex) qPCR, QuantifilerTM qPCR, and nuTH01 (duplex) qPCR. Each point represents the average of duplicate quantifications. For each assay, the DNA quantities are normalized relative to the measured quantity at 0 min of DNase-treatment, i.e., to 8.1 ng/ μ L for the QuantifilerTM assay, to 8.3 ng/ μ L for the nuTH01 assay, and to 1.3×10^6 mt copies/ μ L for the mtND1 qPCR assay. Lines represent linear least-squares fits to early (0–5 min) and late (15–45 min) time points in the degradation series.

(A) Amplification of 100 pg nuclear DNA



(B) Amplification of ~14,000 mitochondrial copies



FIG. 11—HVI/HVII post-amplification yield gel results for DNase-degraded HL60 samples. Gel is EtBr stained. Each pair of bands represents the HVI (~444 bp) and HVII (~416 bp) PCR products from 34 cycles of the duplex amplification. Numerical lane designations (0–60) represent DNase digestion times in minutes; ON represents overnight DNase digestion; L2 and L4 represent, respectively, 20 ng and 40 ng ladder bands (400 bp). In panel (A), amplifications used 100 pg of nuclear DNA as estimated by the nuTH01 portion of the duplex qPCR assay, except for the ON sample, which used 20 μ L of digest. In panel (B), amplifications used 14,000 mitochondrial genome copies as estimated by mtND1 portion of the duplex qPCR assay.

(30). Although the manufacturer recommends using 3–10 ng/20 μ L of PCR product in each 50 μ L cycle-sequencing reaction (59), local validation studies indicate that successful sequencing can be routinely achieved even with as little as 250 pg of PCR product per reaction. Even using the more conservative recommendation of 3–10 ng/20 μ L, we can deduce that a “successful” HVI/HVII amplification is one that produces at least 0.15 ng/ μ L of each PCR product. This concentration is well below the visual detection limit of our ethidium bromide-stained gel. Consequently, our criterion for a successful HVI/HVII amplification is simply that the post-amplification yield gel shows visible evidence, even weak visible evidence, for two appropriately sized product bands. Referring again to Fig. 11, the visual appearance of two such bands for all of the amplifications indicates that both targets of the duplex qPCR assay successfully quantified mitochondrial DNA for HVI/HVII

duplex PCR and that all amplifications produced more than enough of each PCR product for successful cycle sequencing. It is, however, evident that there were implications for selecting a short target sequence for the mtND1 qPCR assay. In particular, Fig. 11 (panel B)) shows that for highly degraded samples (15 min—overnight DNase treatment) increasingly weak HVI/HVII bands are seen in the post-amp yield gel. This trend indicates that the “short” mtND1 qPCR assay overestimated the number of ~400 bp mitochondrial genome copies, just as the QuantifilerTM qPCR assay overestimated the quantity of 100–400 bp nuclear genome copies in these same samples. However, because the mitochondrial cycle sequencing protocol is capable of obtaining successful results over a very wide input range of PCR product (250 pg–10 ng), these overestimations of the quantity of mitochondrial DNA by the mtND1 qPCR assay did not lead to any failed HVI/HVII amplifications.

TABLE 5—Summary of *nuTH01-mtND1* qPCR quantification results for selected reference and casework-type samples. Where indicated, NSF and SF represent, respectively, the non-sperm fraction and sperm fraction from a differential extraction protocol. All quantification results, qPCR and slot blot, are from single replicate experiments. IS indicates that insufficient sample was available for amplification.

Sample	Number and Description	nuTH01 qPCR (ng/4 μ L)	Slot Blot (ng/4 μ L)	mtND1 qPCR (1000 copies/4 μ L)	Genome Copy Ratio— Mitochondrial: Nuclear	Mitochondrial HVI/HVII Post- Amplification Yield Gel	Identifiler™ STR Data (RFUs per locus)				Approximate Mitochondrial Copies Amp'ed for Limited Samples
							TH01	D19	D18	D19/D18	
1	liquid blood	0.42	0.60	200	1600	±	1213	5199	255	20.4	
2	liquid blood	1.22	0.95	380	1000	+	1349	1065	980	1.1	
3	dried blood on black sateen	19.0	11.6	1400	240	+	1814	1444	1403	1.0	
4	dried blood in potting mix	3.86	9.11	480	410	+	7829	6079	1033	5.9	
5	dried blood on newspaper	0.51	0.45	60	390	+	2522	1586	1363	1.2	
6	dried blood on denim	0.12	0.10	37	1000	+	4486	3562	775	4.6	
7	dried blood on wool	2.25	1.50	300	440	+	2096	1335	1230	1.1	
8	dried blood on carpet	0.46	0.39	86	620	+	2747	2204	1429	1.5	
9	dried blood on leather	0.15	0.10	18	400	+	2562	1468	1292	1.1	
10	dried blood on toothpick	1.48	0.66	290	650	+	2466	1399	1145	1.2	
11	femur (powdered)	0.24	0.05	170	2300	+	2412	3231	378	8.6	
12	non-root hair shaft (Chelex)	0	0	9.1	—	+	IS	IS	IS	IS	
13	non-root hair shaft (DNA-IQ)	0	0	2.0	—	±	IS	IS	IS	IS	5000
14	toenail	4.83	0.88	9700	6600	+	2319	3038	459	6.6	
15	toenail	1.06	0.22	1400	4400	+	1810	1523	333	4.6	
16	semen on denim (NSF)	0.53	0.13	50	320	+	3459	2335	1323	1.8	
17	semen on boxer shorts (NSF)	0.23	0.09	990	14000	±	2867	4702	456	10.3	
18	vaginal swab (NSF)	32.1	19.9	880	90	+	3046	1869	927	2.0	
19	electric razor shavings	1.83	0.80	2900	5200	±	1661	1920	484	4.0	
20	swab of razor	0.42	0.07	210	1700	+	1218	915	534	1.7	
21	bloody bandage	52.4	10.1	1700	110	+	1747	1319	1055	1.3	
21	saliva from cigarette butt	0.77	0.29	470	2000	+	1814	1553	671	2.3	
22	saliva from envelope	0.52	0.35	160	1000	+	1890	1308	847	1.5	
24	toothbrush	1.85	1.06	390	700	+	1608	1314	749	1.8	
25	saliva from coffee cup	0.84	0.29	270	1100	+	3017	2051	936	2.2	
26	buccal swab	16.0	5.10	4800	990	+	1677	1182	1130	1.0	
27	buccal swab	17.8	6.70	4900	910	+	1722	1476	937	1.6	
28	buccal swab	5.49	2.13	2400	1400	+	2059	1335	1134	1.2	
29	saliva from envelope	0.04	0	3.4	280	±	IS	IS	IS	IS	1300
30	vaginal swab (SF)	0.04	0	0.036	2.7	0	1204	1417	1005	1.4	300
31	vaginal swab (SF)	4.68	2.99	0.71	0.5	+	2464	1555	1101	1.4	
32	semen on denim (SF)	2.03	1.61	0.014	0.02	+	1580	1067	1038	1.0	700
33	semen on leather (SF)	0.27	0.09	0	0	+	3579	2635	1888	1.4	(20 μ L)
34	semen on cotton fabric (SF)	39.4	28.5	0.16	0.01	+	2164	1592	1272	1.3	7000
35	semen/blood on swab (SF)	0.74	1.06	0.046	0.2	+	1857	1832	1304	1.4	2300
36	semen on boxer shorts (SF)	0.40	0.15	38	310	±	3963	3068	552	5.6	
37	muscle tissue	133	31.4	20000	500	+	1323	620	696	0.9	
38	baby tooth	0.26	0.23	38	480	±	2017	1846	1354	1.4	
39	baby tooth	10.6	2.19	800	250	+	2432	1780	1052	1.7	
40	tooth	53.1	34.1	5300	330	+	1723	1197	1198	1.0	

nuTH01-mtND1 Duplex qPCR Validation: Reference and Casework-type Samples

In order to further evaluate the nuTH01-mtND1 duplex qPCR assay, we quantified DNA extracts from a variety of reference and simulated casework samples for subsequent STR genotyping and mitochondrial HVI/HVII amplifications. A summary of results for a subset of 40 samples from this study is presented in Table 5. For STR genotyping, each amplification used 1 ng of nuclear template DNA as determined by the nuTH01 portion of the duplex qPCR assay. STR results are represented in the table by the RFU values per locus for three of the 15 AmpF/STR® Identifier™ STR loci, TH01(163–202 bp), D19(102–136 bp), and D18(262–346 bp) (60). Data for these particular loci were selected because their amplicons represent intermediate, short, and long STR alleles, respectively, in the genotyping kit. The D19/D18 RFU ratios are tabulated to provide a simple STR-based metric to indicate DNA degradation. For HVI/HVII amplifications, ~14,000 mitochondrial copies were used per reaction, as quantified by the mtND1 portion of the duplex qPCR assay. For some samples, limited quantities required amplification of fewer than 14,000 copies (last column in the table). HVI/HVII amplifications were evaluated by post-amplification gel electrophoresis, where the gel results are represented by either a “+” (to indicate two visible HVI/HVII product bands with expected intensities), a “±” (to indicate two visible product bands with lower than expected intensities), or a “0” (to indicate that no product bands were seen). Both “+” and “±” results indicate that the amplifications provided more than enough HVI/HVII product for cycle sequencing.

For nearly all of these casework-type samples, quantifications provided by the nuTH01-mtND1 duplex qPCR assay led to successful STR and HVI/HVII amplifications. With one exception, each sample that contained a sufficient quantity of nuclear DNA was amplified to give a full STR profile. For the single exception, sample #1, only the CSF alleles were undetected, a result consistent with the very high degree of DNA degradation evident in this sample (D19/D18 > 20). Sample #4 was inhibited, as indicated by nuTH01 and mtND1 qPCR amplification curves with shallow slopes and reduced final plateau values (data not shown). Due to the presence of co-extracted inhibitors in this sample, the nuTH01 qPCR assay underestimated the quantity of nuclear DNA, leading to STR amplification of excess template and to unusually intense TH01 and D19 allele signals. Sample #30 gave an unsuccessful HVI/HVII duplex amplification; this sample was extracted from a sperm cell fraction that contained very little nuclear or mitochondrial DNA. Several other sperm cell extracts (samples 31–35) gave successful HVI/HVII duplex amplifications despite having very low or even undetected quantities of mitochondrial DNA, indicating that the mtND1 portion of the assay significantly underestimated the mitochondrial quantity in these particular samples. As discussed earlier in this paper, such underestimations of DNA by the mtND1 qPCR assay are expected when the ratio of mitochondrial to nuclear genome copies is low, as is the case for these sperm fraction samples. For non-sperm-cell-fraction samples, the empirical ratios of mitochondrial to nuclear genome copies (mt#:nu#) were seen to range from 90 to 14,000. The breadth of this range is likely due, in part, to intrinsic differences in the mt#:nu# for the various tissue types and, in part, due to differences in the extent of DNA fragmentation in these samples, as the short mitochondrial ND1 qPCR target sequence detects smaller DNA fragments than does the longer nuclear TH01 target sequence. It is worth noting that for any sample exhibiting an empirical mt#:nu# greater than 2000, there was corresponding evidence for DNA degradation in

the STR intensity ratios (D19/D18 > ~4), suggesting that genome copy number ratios estimated from the duplex qPCR assay can provide useful information about DNA fragmentation.

Conclusions

We have developed a duplex qPCR assay for the specific quantification of human nuclear and mitochondrial genomes in a variety of forensic-type samples. The nuclear portion of the assay (nuTH01) quantifies DNA in samples with concentrations ranging from ~10 pg/μL to at least 25 ng/μL, while the mitochondrial portion of the assay (mtND1) provides quantifications over the range ~4 copies/μL to ~2 million copies/μL. The quantifications were suitably accurate and precise for determining template quantities for nuclear STR genotyping and mitochondrial HVI/HVII sequencing applications.

For quantifying DNA in highly degraded samples, our results indicate that the length of the qPCR target sequence is an important consideration when selecting an assay for forensic use. For the purpose of STR genotyping, the nuTH01 portion of the duplex qPCR assay, which uses a ~170–190 bp target sequence, was found to be well suited for estimating nuclear DNA quantities in highly degraded samples. Because this assay detects and quantifies DNA fragments that are of the same general length as the STR alleles amplified in commercial genotyping kits, it provided an optimal estimate of the quantity of nuclear template to amplify for successful STR genotyping. By contrast, less than optimal results were obtained when these same highly degraded samples were quantified using a nuclear qPCR assay with a much shorter target sequence (62 bp). This assay, by detecting shorter DNA fragments, overestimated the quantity of STR-sized fragments, resulting in a relatively high proportion of under-amplified and undetected STR alleles. Slot blot quantifications of the same highly degraded samples resulted in underestimated quantities of nuclear DNA, leading to over-amplification of excess template, off-scale STR peaks, and undesired artifact signals.

Acknowledgments

We appreciate advice and assistance from a number of colleagues at the State of California Jan Bashinski Department of Justice DNA Lab, including Mavis Date Chong, Ram Kishore, Sonja Klein, Jeanette Wallin, Eva Steinberger, and Cecilia von Beroldingen. For providing DNA extracts from a variety of sample types, we acknowledge the following colleagues: Angelynn Moore, Amy Hoover, Christopher Gale, Jessica Battaglia, Dianne Burns, Deanna Kacer and Jeanette Wallin.

References

1. Nicklas JA, Buel E. [Quantification of DNA in forensic samples](#). *Anal Bioanal Chem* 2003;376:1160–7. [\[PubMed\]](#)
2. Wayne JS, Presley LA, Budowle B, Shutler GG, Fournery RM. A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. *Biotechniques* 1989;7:852–5. [\[PubMed\]](#)
3. Walsh PS, Varlaro J, Reynolds R. A rapid chemiluminescent method for quantification of human DNA. *Nucleic Acids Res* 1992;20:5061–5. [\[PubMed\]](#)
4. Mandrekar MN, Erickson AM, Kopp K, Krenke BE, Mandrekar PV, Nelson R, et al. Development of a human DNA quantitation system. *Croat Med J* 2001;42:336–9. [\[PubMed\]](#)
5. Hayn S, Wallace MM, Prinz M, Shaler RC. [Evaluation of an automated liquid hybridization method for DNA quantification](#). *J Forensic Sci* 2004;49:87–91. [\[PubMed\]](#)

6. Sifis ME, Both K, Burgoyne LA. A more sensitive method for the quantitation of genomic DNA by Alu amplification. *J Forensic Sci* 2002;47:589–92. [PubMed]
7. Nicklas JA, Buel E. Development of an Alu-based, QSY7-labeled primer PCR method for quantitation of human DNA in forensic samples. *J Forensic Sci* 2003;48:282–91. [PubMed]
8. Fox JC, Cave CA, Schumm JW. Development, characterization, and validation of a sensitive primate-specific quantification assay for forensic analysis. *Biotechniques* 2003;34:314–22. [PubMed]
9. Mackay I, Arden KE, Nitsche A. Real-time PCR in virology. *Nucleic Acids Res* 2002;30:1292–305. [PubMed]
10. Walker NJ. A technique whose time has come. *Science* 2002; 296:557–8. [PubMed]
11. Bustin SA. Quantification of mRNA using real-time reverse transcript PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 2002;29: 23–39. [PubMed]
12. Cockerill FR. Application of rapid-cycle real-time polymerase chain reaction for diagnostic testing in the clinical microbiology laboratory. *Arch Pathol Lab Med* 2003;127:1112–20. [PubMed]
13. Ginzinger DG. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol* 2002;30:503–12. [PubMed]
14. Nicklas JA, Buel E. Development of an Alu-based, real-time PCR method for quantitation of human DNA in forensic samples. *J Forensic Sci* 2003;48:936–44. [PubMed]
15. Richard ML, Frappier RH, Newman JC. Developmental validation of a real-time quantitative PCR assay for automated quantification of human DNA. *J Forensic Sci* 2003;48:1041–46. [PubMed]
16. Applied Biosystems. Quantifiler™ kits user's manual (part number 4344790B). Foster City, CA: Applied Biosystems, 2003. [PubMed]
17. Andreasson H, Gyllensten U, Allen M. Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. *Biotechniques* 2002;33:402–11. [PubMed]
18. von Wurmb-Schwark N, Higuchi R, Fenech AP, Elfstroem C, Meissner C, Oehmichen M, Cortopassi GA. Quantification of human mitochondrial DNA in real time PCR. *Forensic Sci Int* 2002;126:34–9. [PubMed]
19. Alonso A, Martin P, Albarran C, Garcia P, Primorac D, Garcia O, et al. Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies. *Croat Med J* 2003;44:273–80. [PubMed]
20. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res* 1996;6:986–94. [PubMed]
21. Tyagi S, Kramer FR. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 1996;14:303–8. [PubMed]
22. Whitaker JP, Clayton TM, Urquhart AJ, Millican ES, Downes TJ, Kimpston CP, Gill P. Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. *Biotechniques* 1995;4:670–7. [PubMed]
23. Alonso A, Andelinovic S, Martin P, Sutlovic D, Erceg I, Huffine E, et al. DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. *Croat Med J* 2001;42:260–6. [PubMed]
24. Prinz M, Schmitt C. Effect of degradation on PCR based DNA typing. *Adv Forensic Haemogenetics* 1994;5:375–8. [PubMed]
25. von Wurmb-Schwark N, Schwark T, Harbeck M, Oehmichen M. A simple duplex-PCR to evaluate the DNA quality of anthropological and forensic samples prior to short tandem repeat typing. *Leg Med (Tokyo)* 2004;6: 80–8. [PubMed]
26. Alonso A, Martin P, Albarran C, Garcia P, Garcia O, de Simon LF, et al. Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. *Forensic Sci Int* 2004;139:141–9. [PubMed]
27. Gill P, Jeffreys AJ, Werrett DJ. Forensic application of DNA 'fingerprints'. *Nature* 1985;318:577–9. [PubMed]
28. Promega Corporation. Tissue and hair extraction kit protocol (TB307). Madison, WI: Promega Corporation, 2002. [PubMed]
29. Gabriel MN, Calloway CD, Reynolds RL, Primorac D. Identification of human remains by immobilized sequence-specific oligonucleotide probe analysis of mtDNA hypervariable regions I and II. *Croat Med J* 2003;44:293–8. [PubMed]
30. Chong MD, Calloway CD, Klein SB, Orrego C, Buoncristiani MR. Optimization of a duplex amplification and sequencing strategy for the HVI/HVII regions of human mitochondrial DNA for forensic casework. *Forensic Sci Int*. In press. [PubMed]
31. Applied Biosystems. PrimerExpress™ software v2.0 user's manual (part number 4329500A). Foster City, CA: Applied Biosystems, 2001. [PubMed]
32. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank: update. *Nucleic Acids Res* 2004;32:Database issue:D23–6. (<http://www.ncbi.nlm.nih.gov/Genbank/GenbankOverview.html>) [PubMed]
33. Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res* 2001;29:320–2. (<http://www.cstl.nist.gov/biotech/strbase/>) [PubMed]
34. Kogelnik AM, Lott MT, Brown MD, Navathe SB, Wallace DC. MITOMAP: A human mitochondrial genome database. *Nucleic Acids Res* 1996;24:177–9. (<http://www.mitomap.org/>) [PubMed]
35. Masibay A, Mozer TJ, Sprecher C. Promega corporation reveals primer sequences in its testing kits. *J Forensic Sci* 2000;45:1360–2. [PubMed]
36. Butler JM, Becker CH. Improved analysis of DNA short tandem repeats with time-of-flight mass spectrometry. *NIJ Science and Technology Research Report*, 2001: Report NCJ 188292. (<http://www.ncjrs.org/pdffiles1/nij/188292c.pdf>) [PubMed]
37. Crews S, Ojala D, Posakony J, Nishiguchi J, Attardi G. Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. *Nature* 1979;277:192–8. [PubMed]
38. <http://www.ncbi.nlm.nih.gov/blast>
39. <http://www.ncbi.nlm.nih.gov/SNP/>
40. <http://www.ncbi.nlm.nih.gov/omim/>
41. Applied Biosystems. TaqMan® universal PCR master mix protocol (part number 4304449C). Foster City, CA: Applied Biosystems, 2002. [PubMed]
42. Applied Biosystems. User bulletin #5 ABI 7700 sequence detection system: multiplex PCR with TaqMan® VIC probes (part number 4306236B). Foster City, CA: Applied Biosystems, 2001. [PubMed]
43. Creating standard curves with genomic DNA or plasmid DNA templates for use in quantitative PCR. Applied Biosystems. (http://www.appliedbiosystems.com/support/tutorials/pdf/quant_pcr.pdf) [PubMed]
44. Efficiency of PCR reactions. Stratagene Application Note #10 for the Mx4000 Multiplex Quantitative PCR System. (http://www.stratagene.com/lit_items/appnotes10.pdf) [PubMed]
45. <http://web.mit.edu/osp/www/melt.html>
46. Poland D. Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers* 1974;13:1859–71. [PubMed]
47. Fixman M, Freire JJ. Theory of DNA melting curves. *Biopolymers* 1977;16:2693–704. [PubMed]
48. Lerman LS, Silverstein K. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. *Methods Enzymol* 1984;155:482–501. [PubMed]
49. van Oorschot RAH, Gutowski SJ, Robinson SL. HUMTH01: amplification, species specificity, population genetics and forensic applications. *Int J Legal Med* 1994;107:121–6. [PubMed]
50. Crouse CA, Schumm J. Investigation of species specificity using nine PCR-based human STR systems. *J Forensic Sci* 1995;40: 952–6. [PubMed]
51. Annual report summary for testing in 2002. American Association of Blood Banks, 2003. (http://www.aabb.org/About_the_AABB/Stdts_and_Accred/ptannrpt02.pdf) [PubMed]
52. Anderson S, Bankier AT, Barrell BG, deBruijn MH, Coulson AR, Drouin IC, et al. Sequence and organization of the human mitochondrial genome. *Nature* 1981;290:457–65. [PubMed]
53. Cummings MP, Otto SP, Wakeley J. Sampling properties of DNA sequence data in phylogenetic analysis. *Mol Biol Evol* 1995;12: 814–22. [PubMed]
54. Szuhai K, van den Ouweland JM, Dirks RW, Lemaitre M, Truffert JC, Janssen GM, et al. Simultaneous A8344G heteroplasmy and mitochondrial DNA copy number quantification in myoclonus epilepsy and ragged-red fibers (MERRF) syndrome by a multiplex molecular beacon based real-time fluorescence PCR. *Nucleic Acids Res* 2001;29:e13. [PubMed]
55. Gahan ME, Miller F, Lewin SR, Cherry CL, Hoy JF, Mijch A, et al. Quantification of mitochondrial DNA in peripheral blood mononuclear cells and subcutaneous fat using real-time polymerase chain reaction. *J Clin Virol* 2001;22:241–7. [PubMed]
56. Li W, Xi B, Yang W, Hawkins M, Schubart UK. Complex DNA melting profiles of small PCR products revealed using SYBR Green I. *Biotechniques* 2003;35:702–6. [PubMed]

57. Abrams ES, Murdaugh SE, Lerman LS. Intramolecular DNA melting between stable helical segments: melting theory and metastable states. *Nucleic Acids Res* 1995;23:2775–83.
58. Budowle B, Smith J, Moretti T, DiZinno T. DNA typing protocols: molecular biology and forensic analysis. Natick, MA: Eaton Publishing, 2000.
59. Applied Biosystems. BigDye[®] Terminator v1.1 cycle sequencing kit protocol (part number 4337036A). Foster City, CA: Applied Biosystems, 2002.

[\[PubMed\]](#)

60. Applied Biosystems. Application Note: AmpF/STR[®] Identifier[®] PCR amplification kit (publication number 112AP02–01). Foster City, CA: Applied Biosystems, 2003.

Additional information and reprint requests:
Martin R. Buoncristiani, M.P.H.
California Dept. of Justice Jan Bashinski DNA Laboratory
1001 W. Cutting Blvd., Suite 110
Richmond, CA 94804

Carmen Entrala

Laboratorio de ADN forense, Depto. de Medicina Legal

Universidad de Granada, España

TÉCNICAS DE ANÁLISIS DEL ADN EN GENÉTICA FORENSE

Indice:

1. Evolución de las técnicas de estudio de los polimorfismos de ADN

2. Análisis por la reacción en cadena de la polimerasa (PCR)

2.1 Generalidades

2.2 Componentes de la PCR

3. Futuras tecnologías: biochips

3.1 Aplicaciones de los biochips

Sumario:

La identificación con ADN o “huella genética” se basa en el estudio de una serie de fragmentos de ADN presentes en todos los individuos pero que poseen la característica de ser altamente variables o polimórficos entre los mismos.

El análisis de un determinado número de estas secuencias o fragmentos de ADN permite identificar a un individuo con una probabilidad muy cercana al 100%.

Además de ser muy polimórfico, el ADN que se utiliza para la identificación en Genética Forense es un ADN no codificante o no expresivo, por lo que no revela características fenotípicas de los individuos; este hecho es de gran importancia a la hora de considerar la creación de las bases de datos genéticas.

Para analizar dichos polimorfismos del ADN, los laboratorios de Genética Forense utilizan una serie de técnicas que están en continua evolución, consiguiendo que cada vez la identificación por medio del ADN sea más precisa y rápida.

1. Evolución de las técnicas de estudio de los polimorfismos de ADN

La *Hemogenética Forense* nace a principios de siglo, cuando Karl Landsteiner describe el sistema ABO de los hematíes y Von Dürgen y Hirschfeld descubren su transmisión hereditaria. Esta ciencia surgió como una rama de la Criminalística cuyo objetivo era la identificación genética tanto en casos de investigación criminal como en estudios biológicos de la paternidad. Inicialmente, las investigaciones se centraban en el estudio de antígenos eritrocitarios (sistema ABO, Rh, MN), proteínas séricas, enzimas eritrocitarias y sistema HLA. Con el estudio de dichos marcadores podía incluirse o excluirse una persona como posible sospechoso por poseer una combinación genética igual o diferente a la del vestigio biológico hallado en el lugar de los hechos.

Pero fue a mediados de siglo cuando gracias al descubrimiento del ADN y de su estructura y al posterior avance en las técnicas de análisis de dicha molécula la Hemogenética Forense evolucionó considerablemente hasta el punto de que hoy en día puede hablarse de una nueva subespecialidad dentro de la Medicina Forense: la *Genética Forense*. Dicha ciencia estudia básicamente unas regiones del ADN que presentan variabilidad entre los distintos individuos, es decir, estudia regiones polimórficas del ADN. Así, analizando un determinado número de regiones polimórficas, la probabilidad de que dos individuos sean genéticamente iguales es prácticamente nula (excepto en el caso de gemelos univitelinos).

Aunque la Ciencia poseía las herramientas necesarias para el estudio del ADN, su aplicación en la resolución de casos judiciales no se produjo hasta 1985, cuando el Ministerio del Interior Británico solicitó la ayuda de Alec J. Jeffreys, profesor de Genética de la Universidad de Leicester. Los primeros casos de Criminalística fueron resueltos gracias a la técnica de los **RFLPs** (Fragmentos

de Restricción de Longitud Polimórfica). Jeffreys descubrió la existencia de unas regiones minisatélites hipervariables dispersas por el genoma humano que al ser tratadas con enzimas de restricción generaban fragmentos de longitud variable. Estudios posteriores realizados el mismo Jeffreys demostraron que las diferencias en el tamaño de estos fragmentos se debían a que estas regiones consistían en un determinado número de repeticiones en tándem de una secuencia central, el cual variaba de unos individuos a otros.

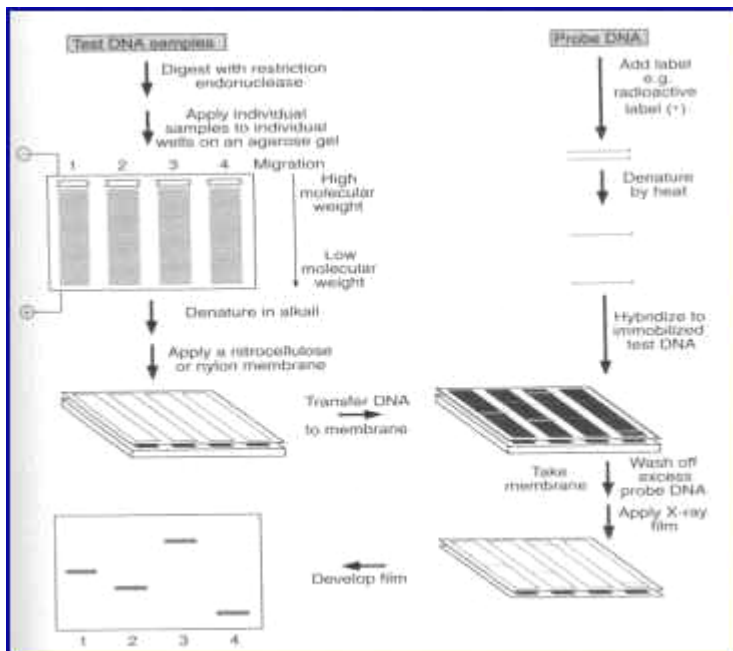
El primer locus de ADN polimórfico fué descubierto por Wyman y White en 1980 usando una sonda de ADN arbitraria. De esta manera observaron fragmentos de más de 15 longitudes diferentes en una pequeña muestra de individuos. Posteriormente se encontraron otros loci hipervariables como en la secuencia del gen de la insulina humana, en el oncogen “ras”, en el pseudogen de la zeta-globina y en el gen de la mioglobina. Estos loci hipervariables constaban de repeticiones en tándem de una secuencia de oligonucleótidos (11 a 60 pb), de manera que las diferentes longitudes de los fragmentos originados dependían del número de dichas repeticiones y se les denominó **VNTR** (“Variable Number of Tandem Repeat”).

Tras el descubrimiento de los primeros VNTRs se vio que éstos podían ser aplicados a la medicina forense y sustituir a los marcadores clásicos.

En un principio la manera de estudiar dichos marcadores se hizo por medio de la técnica llamada **hibridación con sondas** o **Southern blot**. Esta técnica consta básicamente de las siguientes etapas:

1. **Digestión** del ADN con enzimas de restricción tras conseguir extraer un ADN de alta molecularidad.
2. Separación de los fragmentos obtenidos por medio de una **electroforesis** en gel de agarosa.
3. **Desnaturalización** de los fragmentos separados y cortados.

4. **Transferencia** de las cadenas simples a una membrana de nitrocelulosa o nylon y fijación de las mismas por medio de calor (80°C).
5. **Prehibridación** con sondas de ADN inespecífico para bloquear los lugares de unión inespecíficos que pudiera haber en la membrana.
6. **Marcaje de la sonda** con nucleótidos radioactivos (^{32}P normalmente).
7. **Hibridación** de la sonda marcada y desnaturalizada con los fragmentos de ADN fijados a la membrana, y lavado de la membrana para eliminar el exceso de sonda o aquellas que hayan hibridado mal.
8. **Revelado** en placa radiográfica e **interpretación** de los resultados.

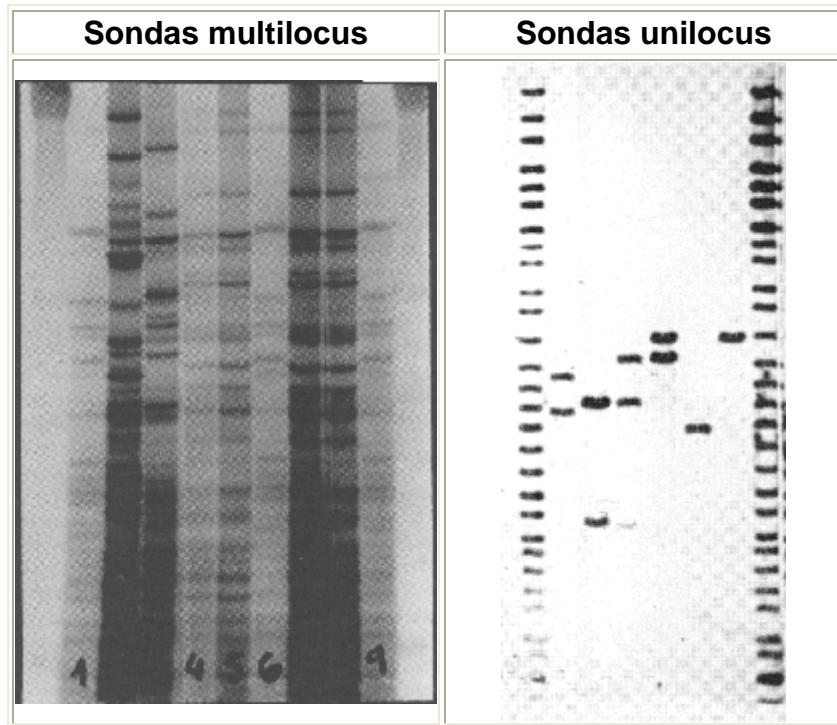


El tipo de sondas utilizadas puede ser de dos tipos:

Sondas Mono-locus (SLP): son específicas para una región de un determinado cromosoma. Se unen a secuencias largas de nucleótidos y presentan mayor variabilidad que las sondas multi-locus. Como resultado se observan una o dos bandas por individuo, según sea homocigoto o heterocigoto. El patrón de bandas obtenido con estas sondas se denomina perfil unilocus de ADN o **“DNA profiling”**

Sondas Multi-locus (MLP): hibridan con secuencias minisatélites

presentes en varios loci de diferentes cromosomas. Son sondas de 10 a 15 nucleótidos que se repiten múltiples veces y tras el revelado se observan de 10 a 20 bandas por persona. Este patrón de múltiples bandas se conoce como huella genética multilocus o **“DNA fingerprint”**.



Las sondas multi y mono-locus presentan una serie de ventajas e inconvenientes con respecto a una serie de parámetros como son:

Información aportada: las sondas multi-locus tienen una mayor capacidad discriminativa al aparecer múltiples bandas. No obstante, las mono-locus son más específicas ya que el fragmento de ADN con el que hibridan es de mayor tamaño.

Cantidad y calidad del ADN: cuando se usan sondas multi-locus se requiere aproximadamente un microgramo de ADN sin degradar mientras que en el caso de las mono-locus se necesita menos de 100 ng y este ADN no necesariamente debe estar en perfecto estado, siempre y cuando el fragmento complementario a la sonda esté intacto.

Especificidad entre especies: las sondas multi-locus permiten su uso

sobre el ADN humano y de cientos de animales superiores, mientras que las mono-locus son exclusivas de ADN humano.

A pesar de que el análisis SLP ha sido y es bastante útil en estudios de paternidad no puede decirse lo mismo de su aplicación a la Criminalística ya que presenta una serie de inconvenientes como son:

La **cantidad de ADN** que se necesita está entre 20 y 100 ng, cantidad difícil de conseguir en casos de criminalística en los que los indicios biológicos encontrados son mínimos.

En cuanto a la **calidad del ADN**, en la práctica forense es muy difícil encontrar en estado no degradado toda la cantidad de ADN que se necesita para un análisis con sondas mono-locus.

El tiempo requerido para este tipo de análisis es de dos o tres días.

El hecho de que se requieran cantidades elevadas de ADN hacen que normalmente, con el primer análisis se consume la totalidad de la muestra, con lo que se dificultan contrapericias y una posterior revisión del caso.

Todas estas limitaciones fueron superadas gracias a la aplicación en Genética Forense de una técnica, la **Reacción en Cadena de la Polimerasa** (“PCR”), que supuso una revolución en muchos campos de la Biología y de la Medicina.

El estudio de indicios biológicos por PCR ha permitido la resolución de un gran número de casos en Criminalística que hasta entonces eran desestimados por no poseer la suficiente cantidad de muestra para su análisis por RFLP. Con el uso de la PCR muestras tan mínimas como pueden ser un *pelo con raíz, una minúscula mancha de sangre o semen e incluso caspa* son suficientes en muchos casos para llevar a cabo un análisis de identificación genética.

2. Análisis por la reacción en cadena de la polimerasa (PCR)

2.1 General

La reacción en cadena de la polimerasa (conocida como **PCR** por sus siglas en inglés, *Polymerase Chain Reaction*) permite amplificar más de un millón de veces un ADN obtenido a partir de una región seleccionada del genoma, siempre y cuando se conozca una parte de su secuencia de nucleótidos. Esta técnica fue ideada en 1989 por Kary B. Mullis que obtuvo el premio Nobel de Química en 1993 por dicho invento.

Para la PCR se utilizan dos oligonucleótidos sintéticos de unos 15-20 nucleótidos que son complementarios a las zonas flanqueantes de la región que se quiere amplificar. Estos oligonucleótidos (habitualmente conocidos por su nombre en inglés, "primers") actúan como cebadores para la síntesis in vitro de ADN la cual está habitualmente catalizada por una enzima llamada *Taq polimerasa*. Dicha enzima se aísla de una bacteria termófila, denominada *Thermus Aquaticus*, que es capaz de crecer a temperaturas elevadas (79 - 85 ° C). A esta temperatura dicha enzima es capaz de mantener una media de extensión de más de 60 nucleótidos por segundo en regiones ricas en uniones G-C. La temperatura óptima a la que actúa la *Taq polimerasa* permite el uso de elevadas temperaturas para la unión de los primers y para la extensión, de esta manera se aumenta el nivel de exigencia de la reacción y se reduce la extensión de los primers unidos inespecíficamente al ADN.

La reacción se lleva a cabo en una serie de ciclos cada uno de los cuales incluye tres fases o pasos:

1. **DESNATURALIZACIÓN:** Para que comience la reacción es necesario que el ADN molde se encuentre en forma de cadena sencilla.

Esto se consigue aplicando temperaturas de 90 a 95°C que producen la rotura de los puentes de hidrógeno intercatenarios y por lo tanto la separación de ambas cadenas. Para conseguir la completa separación de las hebras de toda la muestra esta temperatura debe mantenerse unos minutos. Si el ADN solo se desnaturaliza parcialmente éste tenderá a renaturalizarse rápidamente, evitando así una eficiente hibridación de los primers y una posterior extensión.

2. **HIBRIDACIÓN:** Esta fase se denomina también fase de “annealing” o de emparejamiento. Una vez que el ADN está desnaturalizado se disminuye la temperatura hasta un rango comprendido entre los 40 y los 60°C para que se pueda producir la unión de los primers a las secuencias flanqueantes del fragmento que se va a amplificar. La temperatura de fusión o annealing (T_m , “melting temperature”) depende de varios factores y es relativamente específica para cada primer. La longitud de los primers y la secuencia son críticas en la designación de los parámetros de una amplificación, una fórmula simple para calcular la T_m es la siguiente:

$$T_m = 4(G+C) + 2(A+T).$$

No obstante, cada primer exige una serie de estudios experimentales para determinar su temperatura de annealing específica ya que si la temperatura es muy baja la unión se hará de forma inespecífica y si es muy alta no se producirá una unión completa.

3. **EXTENSIÓN:** Durante este paso la *Taq polimerasa* incorpora nucleótidos en el extremo 3' del primer utilizando como molde la cadena de ADN previamente desnaturalizada. La temperatura a la que se lleva a cabo este paso suele ser de 72°C ya que es la temperatura a la

que la *Taq polimerasa* alcanza su máxima actividad. Normalmente una extensión de 20 segundos es suficiente para fragmentos menores de 500 pb, y 40 segundos para fragmentos por encima de 1.2Kb.

Un factor importante en el transcurso de las diferentes fases es el **tiempo de rampa**. Este se define como el tiempo invertido en pasar de una temperatura a otra y depende del diseño y de las características del aparato donde se realiza automáticamente este proceso, el termociclador. En las nuevas generaciones de termocicladores este factor se ha ido optimizando para hacerlo mínimo.

2.2 Componentes de la PCR

2.2.1 BUFFER DE AMPLIFICACIÓN

Los buffer de PCR que se utilizan normalmente contienen KCl, Tris y MgCl₂. El **MgCl₂** es el componente que más influye en la especificidad y rendimiento de la reacción ya que los iones Mg²⁺ son necesarios para la actividad de la *Taq polimerasa*, es decir, actúan como cofactores de la polimerasa.

La concentración óptima de MgCl₂ está en torno a 1.5 mM si se emplean concentraciones de 200 mM de cada uno de los dNTPs. No obstante, a veces es necesario probar con diferentes cantidades de Mg ya que un exceso del mismo origina una acumulación de productos inespecíficos y una cantidad insuficiente hace que disminuya el rendimiento de la amplificación.

2.2.2 PRIMERS

A la hora de elegir unos primers para amplificar un determinado fragmento de ADN hay una serie de reglas a seguir:

La **longitud de cada uno de los primers** debe estar comprendida entre 18 y 24 bases ya que se ha comprobado que primers de mayor longitud (30-35 bases) no aumentan el rendimiento y los primers cortos carecen de

suficiente especificidad.

Ambos primers deben tener una T_m similar (como mucho la diferencia entre ambas temperatura debe ser de 5°C).

La **relación bases púricas: bases pirimidínicas** debe ser 1:1 (o como mucho 40-60%).

La **secuencia de los primers** debe comenzar y terminar con 1-2 bases púricas.

Para evitar la formación de **dímeros de primers** es necesario comprobar que los primers no contengan secuencias complementarias entre sí.

Los **dímeros de primers** consisten en fragmentos de doble cadena cuya longitud es muy próxima a la de la suma de los primers y se producen cuando un primer es extendido a continuación del otro. El mecanismo exacto por el que se forman estos dímeros no está completamente determinado. La observación de que primers con los extremos 3' complementarios favorecen su formación sugiere que el paso inicial se debe a interacciones transitorias que aproximan los extremos complementarios. Algunas polimerasas, incluida la Taq, han mostrado una débil actividad polimerizadora no dirigida por un ADN patrón, la cual puede unir nucleótidos adicionales al doble extremo apareado. Si esta actividad puede producirse sobre una hebra sencilla de oligonucleótidos, resultaría una buena oportunidad para que la extensión formara un corto solapamiento en el extremo 3' con el otro primer, suficiente para promover la formación del dímero.

2.2.3 DESOXINUCLEÓTIDOS TRIFOSFATOS

Las concentraciones de dNTPs que suelen usarse están en torno a 200 μ M para cada uno de ellos. En un volumen de reacción de 25 μ l con esta concentración de dNTPs se sintetizarían entre 6-6.5 μ g de ADN. La concentración de dNTPs y de $MgCl_2$ va relacionadas ya que el Mg se une a los dNTPs con lo que concentraciones elevadas de dNTPs inhibirían la reacción

al no tener la *Taq polimerasa* suficiente Mg como para incorporar dNTPs. Para una concentración de 200 μM de cada dNTP se suele añadir MgCl_2 a una concentración de 1.5 mM.

2.2.4 TAQ-POLIMERASA

Las cantidades óptimas de *Taq polimerasa* necesarias para la síntesis de ADN están alrededor de 2 unidades en 25 μl de volumen final de reacción. La actividad de este enzima se ve influenciada por la concentración de dNTPs, de Mg^{2+} y de algunos iones monovalentes de manera que concentraciones elevadas de los mismos inhiben dicha actividad.

Por otro lado, pequeñas concentraciones de KCl estimulan la actividad sintética de la Taq en un 50-60% con un máximo aparente cuando su concentración es de 50 mM. Existen algunos datos relacionados con la influencia de ciertos reactivos que se emplean antes de la amplificación y que alteran la actividad de la Taq. Por ejemplo concentraciones 1M de urea estimulan la actividad, el SDS a bajas concentraciones que la inhibe al igual que concentraciones mayores del 10% de etanol.

2.2.5 ADN MOLDE O "TEMPLATE"

Es el ADN del cual queremos copiar un determinado fragmento, es, por tanto, el ADN que la Taq polimerasa utiliza como molde para la síntesis de nuevas cadenas polinucleotídicas. La cantidad de ADN necesaria para la PCR depende de varios factores:

Del **marcador** que se va a amplificar: hay marcadores cuyos primers son más específicos o bien cuyas condiciones de amplificación están mejor optimizadas que las de otros. Por esta razón puede darse el caso de que cierta cantidad de ADN (sobre todo cuando jugamos con cantidades mínimas) amplifique para unos marcadores pero no para otros. Por ello, cuando en un laboratorio se va

a utilizar un nuevo marcador es necesario hacer un estudio de validación en él que se incluye un estudio de sensibilidad. De dicho estudio de sensibilidad puede sacarse como conclusión cuál es la mínima cantidad de ADN que amplifica en condiciones estándar. De manera general, para casi todos los STR utilizados en Genética Forense la cantidad óptima de ADN que asegura un rendimiento adecuado está en torno a los 5 ng.

Calidad del ADN: Cuando se trabaja con ADN cuya calidad es óptima no suele haber problemas en la amplificación y cantidades del mismo por encima e incluso por debajo de los 5 ng rinden buenos resultados. El problema aparece cuando la calidad del ADN obtenido no es la idónea, bien porque esté degradado o bien porque dicho ADN vaya ligado a una serie de contaminantes que pueden inhibir la actividad de la polimerasa. Si el ADN está degradado por la acción de enzimas de restricción el que obtengamos o no resultado en la amplificación va a depender de que el fragmento a amplificar haya sido dañado o no. En el caso en el que tengamos ADN sin degradar pero unido a una serie de contaminantes habría que intentar diluir al máximo la muestra para disminuir dichos contaminantes, pero siempre dentro de un rango de ADN que no esté por debajo del límite de sensibilidad de la PCR. El problema de las cantidades mínimas de ADN y la presencia de contaminantes o inhibidores de la Taq es un hecho habitual en Criminalística y requiere un estudio pormenorizado de la muestra antes de la amplificación.

ADYUVANTES DE LA PCR

Son elementos que mejoran el rendimiento y la especificidad de la PCR.

Aunque algunos autores han recomendado el uso del **DMSO** y del **glicerol**, el adyuvante más extendido y utilizado es el **BSA**. A concentraciones por encima de 0.8 µg/µl el BSA incrementa la eficiencia de la PCR ya actúa como una proteína captadora de iones que pueden ser inhibidores de la Taq polimerasa.

La PCR ofrece una serie de ventajas, frente al uso de las técnicas de análisis genético utilizadas con anterioridad, como son:

Su capacidad para obtener resultados en casos en los que la cantidad de ADN es **mínima** o en casos en los que el ADN esté parcialmente **degradado**.

Genera en un espacio corto de tiempo un **elevado número de copias** de la secuencia de ADN que es objeto de estudio, lo cual permite utilizar técnicas de visualización más sencillas y rápidas que el uso de sondas marcadas radioactivamente.

Permite la determinación y agrupación alélica en **clases discretas**, lo que facilita la elaboración de bases de datos al ser la estandarización inmediata y posibilitar la aplicación de métodos bioestadísticos y programas elaborados.

El uso de marcadores microsátélites de pequeño peso molecular aumenta las probabilidades de obtener resultados positivos de amplificación cuando el ADN se encuentra degradado ya que puede ser que dichos fragmentos no hayan sido digeridos. Esta ventaja es de gran importancia en Criminalística ya que normalmente los indicios biológicos encontrados han estado sometidos a diversos factores (calor y humedad) que favorecen el crecimiento bacteriano.

Sin embargo, una de las grandes ventajas de la PCR que es su elevada sensibilidad puede, en ocasiones, convertirse en un gran problema ya que se podría coamplificar un ADN extraño o ajeno al que nos interesa. No obstante, en los laboratorios de Genética Forense las medidas de precaución que se toman para evitar problemas de contaminación por manipulación son extremas.

Una vez amplificado el ADN, los fragmentos resultantes son separados en función de su tamaño por medio de un proceso de **electroforesis**.

Actualmente se utilizan dos tipos de electroforesis en los laboratorios de Genética Forense:

Electroforesis en geles verticales de poliacrilamida

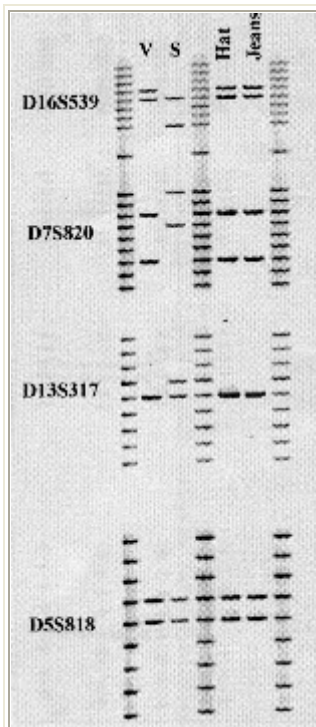
Electroforesis capilar

La **electroforesis capilar** es una técnica relativamente novedosa en el campo de la Genética Forense pero que está poco a poco sustituyendo a los sistemas de electroforesis vertical. En este caso el proceso electroforético es llevado a cabo en un capilar de silica de unas 50 μm de diámetro, lo cual hace que la cantidad de calor generado sea menor y que puedan aplicarse voltajes mayores. Para que puedan ser analizados por electroforesis capilar los primers o los dideoxinucleótidos (en el caso de la secuenciación) deben ser marcados fluorescentemente con unas moléculas denominadas fluorocromos que emiten fluorescencia a una determinada longitud de onda cuando son excitados por láser. El equipo en el que se realiza el proceso lleva acoplado un ordenador encargado de traducir los datos de emisiones fluorescentes en secuencias o fragmentos con sus correspondientes alelos asignados. La electroforesis capilar presenta una serie de ventajas frente a los sistemas de electroforesis vertical como son:

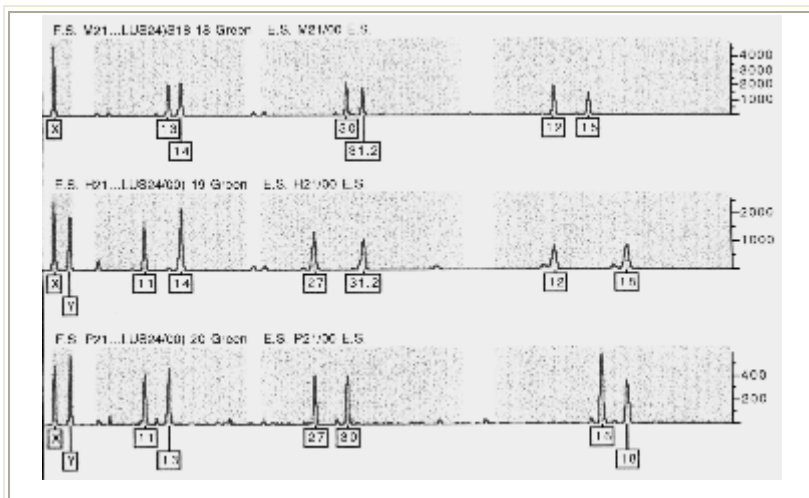
Rapidez: ya que permite el análisis simultáneo de varios loci aunque éstos posean alelos con tamaños solapantes.

Sensibilidad: hace posible detectar cantidades muy pequeñas de ADN amplificado.

Los resultados se obtienen de manera informatizada, lo que evita problemas de interpretación de los resultados y facilita el análisis de los mismos a través de programas informáticos.



Ejemplo de un caso de criminalística resuelto utilizando electroforesis en gel vertical de acrilamida. Si se observan los patrones bandas podrá comprobarse como los perfiles obtenidos en las manchas de sangre encontradas en el sombrero (Hat) y pantalón (Jeans) del sospechoso (S) coinciden con el perfil genético de la víctima (V)



Caso de estudio biológico de la paternidad realizado con electroforesis capilar. Cada pico corresponde a un alelo, el hijo (H) debe poseer un alelo de la madre (M) y otro del padre (P) para que la paternidad sea compatible

3. Futuras tecnologías: biochips

Las técnicas de análisis genético se encuentran hoy en día en continuo desarrollo y evolución. La necesidad de técnicas que permitan el aislamiento y análisis de los casi cien mil genes que componen el genoma humano justifica la existencia de líneas de investigación destinadas al descubrimiento de nuevos métodos que permitan monitorizar elevados volúmenes de información genética en paralelo y que reduzcan tanto el tiempo empleado como el coste por análisis.

Desde el análisis de los primeros polimorfismos de ADN con fines identificativos, la Genética Forense ha sufrido una gran evolución. Los expertos en la materia han sido testigos de cómo el descubrimiento de la PCR revolucionó las técnicas de identificación genética. Es probable que la próxima revolución la constituyan los llamados **biochips o microarrays**.

Los **biochips** surgen como consecuencia de una combinación entre técnicas microelectrónicas empleadas para la fabricación de microprocesadores informáticos y materiales biológicos. En general puede decirse que la principal característica de los chips es su capacidad para generar información en muy poco espacio, ya que posibilitan el procesamiento de multitud de ensayos simultáneamente. Esta característica es la que hace que los biochips sean probablemente la tecnología del futuro en el campo de las investigaciones biomédicas.

La fabricación de los biochips es similar a la de los chips informáticos: por medio de la técnica denominada **fotolitografía** se depositan circuitos microscópicos sobre láminas de silicio. En el caso concreto de los biochips, estas láminas son de vidrio y lo que se deposita en dichas láminas son cadenas de ADN. Las láminas de vidrio presentan una serie de ventajas como son:

Posibilidad de unir las cadenas de ADN a la superficie del cristal, convenientemente tratada, mediante enlaces covalente.

Su capacidad para aguantar altas temperaturas y lavados de elevada fuerza iónica.

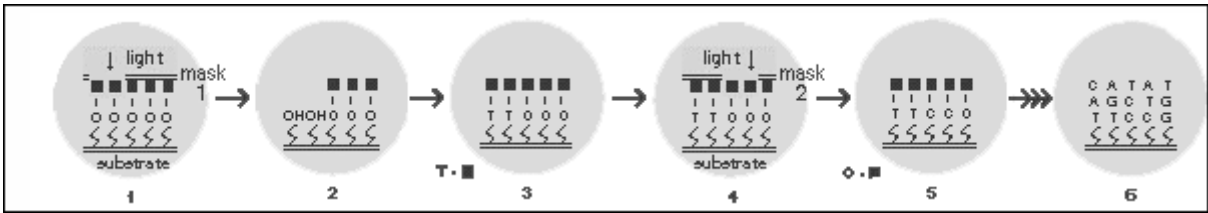
Al ser un material no poroso el volumen de hibridación puede reducirse al mínimo.

Su baja fluorescencia evita ruidos de fondo.

Hay compañías comerciales que han desarrollado otras estrategias para la fabricación de biochips. No obstante, los más usados actualmente y con mayor número de aplicaciones son los basados en *técnicas fotolitográficas*.

Estos chips, como se dijo anteriormente, consisten en una pequeña lámina de vidrio que posee unos grupos reactivos, a los que posteriormente se unirán los nucleótidos, protegidos mediante una película realizada con un agente químico fotodegradable. Mediante una máscara que se sitúan sobre el chip, se logra enfocar un haz de luz hacia unas posiciones y regiones determinadas, degradando al agente químico protector en dichas zonas y dejándolo intacto en las zonas protegidas por la máscara. A continuación se añade al chip un medio que contiene uno de los cuatro nucleótidos que se unirá, mediante enlace covalente, al cristal con los grupos reactivos que hayan quedado desprotegidos. Cada grupo añadido lleva una molécula receptora fotodegradable.

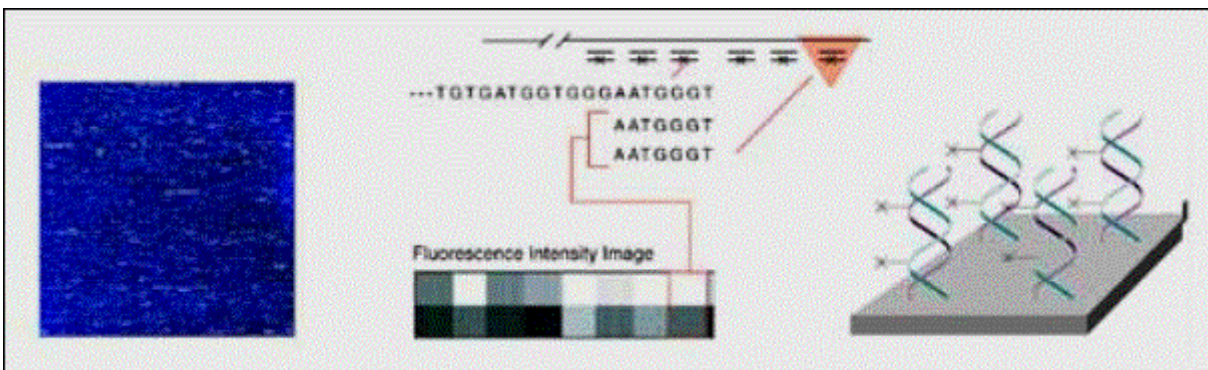
Todo este proceso se va repitiendo con los diferentes nucleótidos y máscaras hasta generar unos oligonucleótidos con las secuencias que nos interesen. El siguiente esquema muestra el todo el proceso explicado anteriormente.



Cada “casilla” del chip posee una cadena de un oligonucleótido de manera que solamente aquel fragmento de ADN que hibride perfectamente con ella permanecerá unido tras los diversos lavados.

Previamente a la hibridación, el ADN de la muestra a estudiar debe haber sido amplificado y marcado fluorescentemente en uno de sus extremos. Una vez marcado se incuba en el recipiente que contiene al chip tras lo cual se lava varias veces para eliminar los fragmentos que no hayan hibridado y se introduce el chip en un escáner en el que se detectan los patrones de hibridación. Esta detección se realiza en base a la fluorescencia emitida por los fluorocromos de la muestra cuando son excitados por luz y en aquellos pocillos en los que la unión haya sido completa la fluorescencia será mayor que los que contengan alguna base desapareada.

Un ordenador conectado al escáner es el que identifica las secuencias sonda por su posición el chip.



3.1 Aplicaciones de los biochips

A pesar de ser una tecnología muy reciente y que, por lo tanto, está aún en vías de experimentación, actualmente los biochips están siendo aplicados en:

Monitorización de expresión génica: permite determinar cual es el patrón de expresión génica y cuantificar el nivel de expresión de manera simultánea para un elevado número de genes. Esto permite realizar estudios comparativos de activación de determinados genes en tejidos sanos y enfermos y determinar así la función de los mismos.

Detección de mutaciones y polimorfismos: Permite el estudio de todos los posibles polimorfismos y la detección de mutaciones en genes complejos.

Secuenciación: Mientras que se han diseñado algunos biochips para secuenciación de fragmentos cortos de ADN, no existe aún en el mercado ningún biochip que permita secuenciar de novo secuencias largas de ADN.

Diagnóstico clínico y detección de microorganismos: Posibilitan la identificación rápida empleando unos marcadores genéticos de los patógenos.

Screening y toxicología de fármacos: el empleo de los biochips permite el analizar los cambios de expresión génica que se dan durante la administración de un fármaco de forma rápida, así como la localización de nuevas posibles dianas terapéuticas y los efectos toxicológicos asociados.

Seguimiento de terapia: los biochips permiten valorar rasgos genéticos que pueden tener incidencia en la respuesta a una terapia.

Medicina preventiva: El conocimiento y posible diagnóstico de ciertos caracteres genéticos asociados a determinadas patologías permite una prevención de las mismas antes de que aparezcan los síntomas.

Copyright © 2000 by Carmen Entrala. Dra. en biología.

Otros documentos del Curso de Doctorado