



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

---

FACULTAD DE ESTUDIOS SUPERIORES IZTACALA

“DETERMINACIÓN DEL PERFIL DE LAS  
VARIACIONES ALÉLICAS EN EL *locus*  
COL4A3 DEL COLÁGENO TIPO IV EN NIÑOS  
MEXICANOS CON SÍNDROME DE ALPORT”

TESIS

QUE PARA OBTENER EL TÍTULO DE

BIÓLOGO

PRESENTA

GARCÍA VÁZQUEZ RAÚL

DIRECTORA DE TESIS

BIÓL. ADRIANA SANCHÉZ BOISO



Tlalnepantla, Estado de México

Agosto 2011



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## JURADO ASIGNADO

**Presidente:** Dr. Santiago Martínez Calvillo

**Vocal:** Dr. Sergio Vaca Pacheco

**Secretario:** Biól. Adriana Sánchez Boiso

**1er Suplente:** Dr. Diego Julio Arenas Aranda

**2do Suplente:** M. en C. Gloria Luz Paniagua Contreras

El presente estudio fue desarrollado en el área de biología molecular del Departamento de Genética y Nefrología. Hospital Infantil de México “Federico Gómez”





“La libertad no reside en una imaginaria independencia respecto de las leyes de la naturaleza, sino el conocimiento de dichas leyes y en la posibilidad, basada en este conocimiento, de lograr planificadamente que las leyes de la naturaleza actúen en una dirección determinada”

**F. Engels. Anti-Duhring, 1878.**



# ÍNDICE

<b>1</b>	<b>Introducción</b> .....	<b>1</b>
1.1	Filtración renal.....	1
1.1.1	Glomérulo.....	4
1.1.2	Membrana Basal Glomerular .....	5
1.1.3	Colágeno tipo IV .....	7
1.2	Gen <i>COL4A3</i> .....	15
1.3	Nefritis hereditaria.....	20
1.3.1	Síndrome de Alport.....	21
<b>2</b>	<b>Antecedentes</b> .....	<b>32</b>
2.1	Estudios clínico-patológicos .....	32
2.2	Estudios moleculares.....	35
<b>3</b>	<b>Justificación</b> .....	<b>45</b>
<b>4</b>	<b>Hipótesis</b> .....	<b>46</b>
<b>5</b>	<b>Objetivos</b> .....	<b>46</b>
5.1	General .....	46
5.2	Particulares.....	46
<b>6</b>	<b>Metodología</b> .....	<b>48</b>
6.1	Métodos .....	48
6.1.1	Población de estudio .....	48
6.1.2	Toma de muestra .....	48
6.1.3	Extracción de ADN .....	48
6.1.4	Cuantificación del ADN.....	50

6.1.5	Integridad del ADN.....	50
6.1.6	Amplificación de los fragmentos del gen COL43 mediante reacción en cadena de la polimerasa (PCR) .....	51
6.1.7	Electroforesis en gel de agarosa .....	54
6.1.8	Secuenciación automatizada mediante el principio de terminación de la cadena “didesoxi” .....	55
6.1.9	Detección de variantes alélicas mediante cromatografía líquida desnaturalizante de alto rendimiento (DHPLC).....	60
6.1.10	Pruebas estadísticas .....	64
6.2	Diseño metodológico .....	66
<b>7</b>	<b>Resultados y discusión .....</b>	<b>67</b>
7.1	ADN genómico de la población de estudio .....	67
7.1.1	Cuantificación y pureza de la extracción de ADN .....	67
7.1.2	Integridad de la extracción de ADN.....	68
7.2	Validación de los 17 fragmentos amplificados mediante PCR .....	69
7.3	Detección de las variantes alélicas mediante cribado por DHPLC .....	70
7.3.1	Distribución de las variantes alélicas en la población de estudio .....	71
7.4	Análisis molecular y bioquímico.....	74
7.4.1	Fragmento UTR 5'-Exón 1 .....	74
7.4.2	Fragmento 2.....	78
7.4.3	Fragmento 3.....	84
7.4.4	Exón 7 .....	89
7.4.5	Exón 9 .....	92
7.4.6	Exón 21 .....	94
7.4.7	Exón 22 .....	99
7.4.8	Fragmento 23.....	101

7.4.9	Exón 52 y fragmentos UTR 3' A-I .....	103
7.5	Predicciones PolyPhen y SIFT .....	111
7.6	Frecuencias genotípicas .....	113
<b>8</b>	<b>Conclusión .....</b>	<b>117</b>
<b>9</b>	<b>Perspectivas .....</b>	<b>119</b>
<b>10</b>	<b>Bibliografía .....</b>	<b>121</b>
<b>11</b>	<b>Anexo .....</b>	<b>130</b>
	Características físicas de los cebadores para las condiciones de PCR .....	130

## Resumen

---

El gen *COL4A3* es uno de los seis genes que componen al colágeno tipo IV (COL(IV)), su modificación genética mediante polimorfismos o mutaciones conlleva a una serie de configuraciones estructurales que compromete la funcionalidad de su producto codificante, es decir la cadena  $\alpha 3$  del colágeno tipo IV ( $\alpha 3$ (IV)), dichas modificaciones resultan en fenotipos heterogéneos y pueden ser patológicos de acuerdo al tipo de variante alélica implicada. El caso autosómico de Síndrome de Alport (SA) es un ejemplo de ello. El SA es una enfermedad hereditaria del las membranas basales (MBs) debido a mutaciones en alguno de los genes que componen el COL (IV). Clínicamente está caracterizada por nefropatía hereditaria progresiva, y a veces sordera sensorial y/o lesiones oculares. El uso del método de cromatografía líquida semi-desnaturalizante de alto rendimiento (DHPLC), para buscar variantes alélicas en el gen *COL4A3*, determinó el alto grado polimórfico de la población mexicana (>50%), especulando incluso la posesión de una variabilidad mayor que la de otras poblaciones. Mediante secuenciación se definieron 15 polimorfismos tanto en pacientes como en controles, excepto uno de *novo* en la región no traducible en el extremo 3' (UTR 3'), debido a que solo se localizó en controles. Seis variantes alélicas fueron localizadas en la región exónica: un cambio sinónimo y los demás de sentido erróneo; tres suponen cambios (p.R408H, p.H451R y p.E162G), sin embargo no implican consideraciones significantes en el efecto fenotípico; otro tiene mención especial por especularse benéfico, debido a su posible aportación en la estabilidad estructural y térmica de la cadena  $\alpha 3$ (IV) (p.L141P), el último refleja un impacto potencialmente dañino al reducir la actividad estructural en el sitio de cambio (p.G43R). Ocho variantes alélicas fueron definidas en la región UTR 3', 3 de ellas de *novo* (c.5352A/G, c.5436C/T y c.7214G/C). Por otro lado, fueron encontradas dos mutaciones de *novo* en el dominio colagenoso, las cuales abarcan cuatro nucleótidos consecutivos [c.IVS3 -1G/A] + [c.307-309G/A]; ambas mutaciones concuerdan con la incidencia que tiene el SA autosómico (alrededor del 20% de los casos). Sumado a esto, de acuerdo al análisis auxiliar de herramientas bioinformáticas (SIFT y PolyPhen), existe la posibilidad de que el polimorfismo codificante p.G43R, contribuiría al estado progresivo de SA, característica clave que podría ser enriquecida y utilizada para auxiliar la genética molecular de la enfermedad.

**Palabras clave:** Síndrome de Alport (SA), gen *COL4A3*, cadena  $\alpha 3$  del colágeno tipo IV ( $\alpha 3$ (IV)), polimorfismos y mutación.

## Abreviaturas

---

**7S:** dominio N-terminal no colagenoso.

**aa:** aminoácidos.

**AN:** acetonitrilo (DHPLC).

**ARN(m):** ácido ribonucleico (mensajero).

**ARNi:** ARN de interferencia.

**ATEA:** acetato de trietilamonio (DHPLC).

**cadena  $\alpha 3$ (IV):** proteína codificada por el gen *COL4A3*.

**COL(IV):** colágeno tipo IV.

**CpG:** islas de nucleótidos enlazadas por un fosfato.

**C-terminal:** extremo carboxilo terminal.

**DHPLC:** técnica cromatografía líquida semi-desnaturalizante de alto rendimiento.

**Gly-X-Y:** glicina-prolina-hidroxiprolina.

**HFB:** hematuria familiar benigna.

**HSPGs:** proteoglicanos de heparan sulfato.

**IRCT:** insuficiencia renal crónico terminal.

**MB:** membrana basal.

**MBG:** membrana basal glomerular.

**miARN:** micro ARN.

**NC1:** dominio globular no colagenoso C-terminal.

**NMBD:** nefropatía de la membrana basal delgada.

**ADN:** ácido ribonucleico.

**N-terminal:** extremo amino terminal.

**PCR:** técnica reacción en cadena de la polimerasa.

**RE(R):** retículo endoplasmático (rugoso).

**RT-PCR:** técnica retrotranscriptasa inversa mediante PCR.

**SA:** síndrome de Alport.

**SAAD:** síndrome de Alport autosómico dominante.

**SAAR:** síndrome de Alport autosómico recesivo.

**SALX:** síndrome de Alport ligado al cromosoma X.

**SNP:** polimorfismo de un solo nucleótido.

**SSCP:** técnica de polimorfismos de conformación de cadenas sencillas.

**T<sub>m</sub>:** temperatura semi-desnaturalizante.

**UTR:** región del gen no traducible.

**$\Delta G^\circ$ :** principio termodinámico de energía libre.

**$\alpha$ (IV):** uno de seis posibles productos codificantes del colágeno tipo IV.

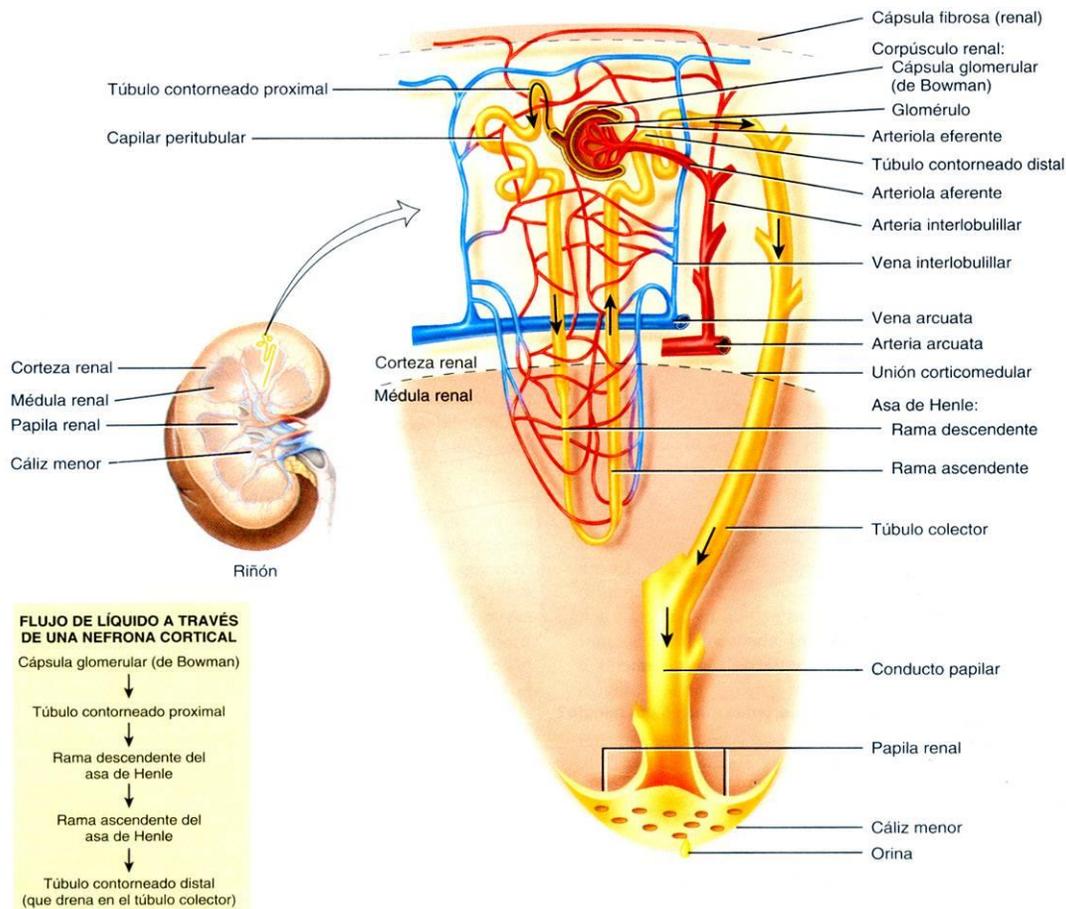
**$\alpha 3$ – $\alpha 4$ – $\alpha 5$ (IV):** protómero conformado por tres cadenas  $\alpha$ (IV).

# 1 Introducción

## 1.1 Filtración renal

Los riñones son órganos pares que se disponen en la pared posterior del abdomen, fuera de la cavidad peritoneal. Su función radica principalmente en regular en la sangre: la composición iónica, el volumen y pH (equilibrio hídrico/electrolítico); la presión arterial (sistema renina-angiotensina-aldosterona, en general autacoides); la producción de hormonas (vitamina D, eritropoyetina); la regulación de la concentración de glucosa (glucogenia); la excreción de productos metabólicos de desecho (urea, creatinina, ácido úrico, bilirrubina, y metabolitos de varias hormonas) y sustancias químicas extrañas (fase I y II del metabolismo xenobiótico). Estas múltiples funciones permiten mantener el ambiente celular estable y necesario para que las células desempeñen sus diversas actividades (Tortora and Derrickson, 2006). De este modo, los riñones poseen una gran irrigación e inervación que les permite realizar el filtrado del plasma con una intensidad variable, dependiendo de las necesidades del cuerpo. El riego sanguíneo de los dos riñones en adultos es normalmente alrededor del 22% del gasto cardiaco, o 1,100 ml/min; sin embargo en niños es variable, aunque generalmente mayor a 90 ml/min (Guyton y Hall, 2009).

La unidad funcional del riñón es la nefrona y existen aproximadamente un millón en cada riñón. En la nefrona se reconocen dos regiones funcionalmente bien definidas: un corpúsculo renal donde es filtrado el plasma sanguíneo, incluye la capsula de Bowman y el glomérulo; mientras que la otra región la constituyen los peritúbulos renales, éstos pasan el líquido filtrado, consisten en el túbulo contorneado proximal, el asa de Henle y el túbulo contorneado distal (fig. 1) (Tortora y Derrickson, 2006).



**Figura 1.** Estructura de una nefrona cortical (color amarillo oscuro) y sus vasos sanguíneos asociados en la irrigación del flujo sanguíneo a través del corpúsculo renal y los peritúbulos (Tomada de Tortora and Derrickson, 2006).

Los lechos capilares glomerulares y peritubulares se encuentran dispuestos en serie y están separados por las arteriolas eferentes, las cuales ayudan a regular la presión hidrostática en los dos grupos de capilares. La presión hidrostática alta en los capilares glomerulares (de unos 60 mm Hg) da lugar a una filtración rápida, mientras que una presión hidrostática mucho menor en los capilares peritubulares (de unos 13 mm Hg) permite una reabsorción lenta del líquido; al ajustar la resistencia de las arteriolas aferente y eferente, los riñones pueden regular la presión hidrostática en los capilares glomerulares y peritubulares, cambiando el filtrado glomerular, la reabsorción tubular o ambas en respuesta a las demandas homeostáticas del cuerpo (Guyton y Hall, 2009).

El líquido filtrado de los capilares glomerulares entra en el espacio capsular de Bowman, el cual sigue su camino por los capilares tubulares hacia el túbulo renal, para que finalmente la orina formada llegue a la uretra y sea eliminada por micción. La intensidad con la que se excretan diferentes sustancias en la orina representa la suma de tres procesos renales: 1) filtración glomerular; 2) reabsorción de las sustancias de los túbulos renales hacia la sangre, y 3) la secreción de sustancias desde la sangre hacia los túbulos renales.

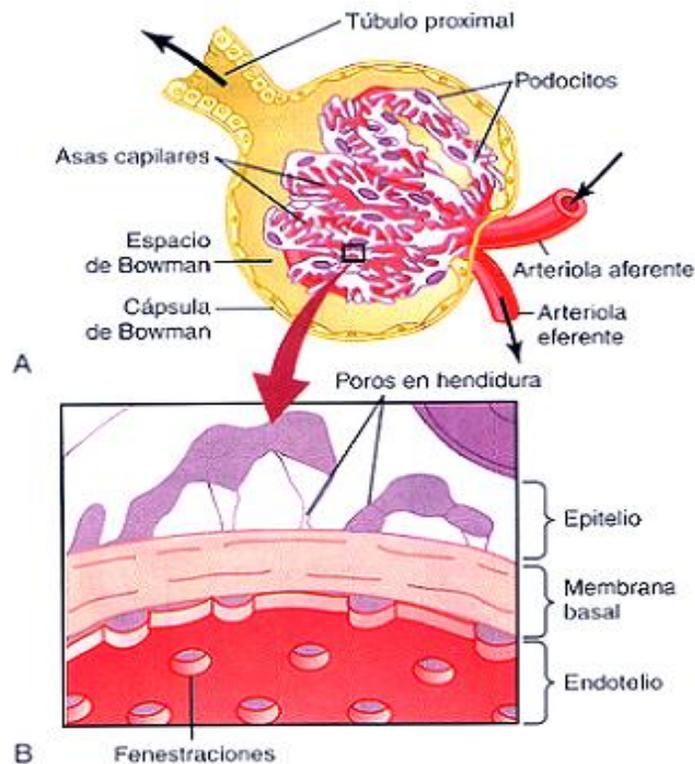
La formación de orina comienza cuando una gran cantidad de líquido que casi no dispone de proteínas se filtra desde los capilares glomerulares a la capsula de Bowman. La mayor parte de las sustancias del plasma, excepto las proteínas, se filtran libremente, de manera que su concentración en el filtrado glomerular de la capsula de Bowman es casi la misma que en el plasma. A medida que el líquido abandona la capsula de Bowman y pasa a través de los túbulos, se modifica por la reabsorción de agua y solutos específicos de nuevo hacia la sangre o por la secreción de otras sustancias desde los capilares peritubulares hacia los túbulos (Tortora y Derrickson, 2006; Guyton y Hall, 2009).

La filtración y reabsorción de grandes cantidades de soluto es debido a dos ventajas que ofrece el filtrado glomerular. En la primera, permite eliminar con rapidez productos de desecho del cuerpo que depende de la filtración glomerular para su excreción; la mayoría de los productos de desecho se absorben mal en los túbulos y, por ello, depende de un filtrado glomerular alto para extraerlos eficazmente del cuerpo. En la segunda, un filtrado glomerular alto permite que el riñón filtre y procese todos los líquidos corporales muchas veces al día; considerando un volumen plasmático de 3 L, mientras que el filtrado glomerular es de 125 ml/día ó 180 l/día y la reabsorción tubular de 178.5 l/día, todo el plasma puede filtrarse y procesarse unas 60 veces al día, lo que deja 1.5 l/día de líquido que se excreta en la orina. Este filtrado glomerular alto permite

a los riñones autoregular de forma rápida y precisa el volumen y la composición de los líquidos corporales del cuerpo (Kriz y Kaissling, 2008).

### 1.1.1 Glomérulo

El glomérulo es el componente de filtración principal de la nefrona, comprende un ovillo capilar de tres capas principales (en lugar de las dos habituales): 1) el endotelio del capilar; 2) una membrana basal glomerular (MBG), y 3) una capa de células epiteliales con proyecciones en forma de pie ó podocitos rodeando la superficie externa de la membrana basal capilar. Juntas, estas capas forman la barrera de filtración que, a pesar de sus tres capas, filtran cientos de veces más agua y solutos que la membrana capilar habitual. El endotelio está perforado por cientos de pequeños agujeros relativamente amplios, llamados fenestraciones; dichas células están dotadas de muchas cargas negativas que dificultan el paso de las proteínas plasmáticas. Rodeando al endotelio está la MBG, la cual contiene grandes espacios a través de los cuales pueden filtrarse grandes cantidades de agua y solutos; generalmente posee cargas negativas más fuertes que las del endotelio. Recubriendo la superficie externa del glomérulo a partir de largas prolongaciones, se encuentran los podocitos; estas células están separadas por espacios llamados poros de hendidura a través de los cuales se mueve el filtrado glomerular; también están provistos de carga negativa (fig. 2). De este modo, todas las capas de la pared capilar glomerular integran una barrera de filtración, que principalmente es impermeable a las proteínas plasmáticas (Kriz y Kaissling, 2008; Guyton y Hall, 2009).



**Figura 2.** Representación de la membrana capilar glomerular. A. Ultraestructura básica de los capilares glomerulares. B. Sección transversal de la membrana capilar glomerular y sus principales componentes: endotelio, membrana basal y podocitos (Tomada de Guyton y Hall, 2009).

### 1.1.2 Membrana Basal Glomerular

La MBG consiste en una matriz acelular especializada que constituye la compartimentalización del glomérulo, separando componentes celulares del mesénquima subyacente, además representa el esqueleto de soporte entre las capas celulares epiteliales y podocíticas del ovillo glomerular (Pöschl *et al*, 2003). Su síntesis, desde el inicio del desarrollo embrionario hasta la madurez del individuo, la llevan a cabo estas mismas células, posteriormente es exclusiva de los podocitos (Kriz y Kaissling, 2008).

La MBG constituye una barrera polianiónica electrodensa que es principalmente impermeable para la albúmina. A menudo la MBG está sujeta a un recambio continuo,

debido a la demanda mecanismo de reflujo sanguíneo, sin embargo pocos detalles son conocidos de cómo y cuándo sus componentes son incorporados, removidos y/o degradados (Agtmael y Tuderman, 2010).

Los componentes substanciales de la MBG incluyen al colágeno tipo IV (COL(IV)), la laminina tipo 11 (laminina 521), nidogeno/enactina, proteoglicanos de heparan sulfato (HSPGs; principalmente agrina y perlecano) y fibronectina (Pöschl *et al*, 2003). Diferentes isoformas de estas grandes moléculas son producidas, las cuales en conjunto establecen una única lamina densa (electrodensa), que se encuentra unida directamente a las bases del endotelio y epitelio (Kriz y Kaissling, 2008).

La laminina 11, fibronectina y enactina son glicoproteínas de la MBG. La principal es la laminina-11, debido a su abundancia y por ser el formador inicial de la MBG, está constituida de un heterotrímero de cadenas  $\alpha_5\text{-}\beta_2\text{-}\gamma_1$  (unidos por puentes disulfuro), este forma segundas redes mediante sus dominios globulares en el extremo amino terminal (N-terminal); mientras que en el extremo carboxilo terminal (C-terminal), une la MBG a la superficie celular del endotelio y podocitos, esto por medio de interacciones con receptores de adhesión de superficie celular, tales como integrina  $\alpha_2\text{-}\beta_1$ , distroglicanos, sindecano y heparina; en tanto por vía nidogeno/enactina se une en sitios específicos a la red de COL(IV), formando de esta forma complejos ternarios. La red proteica que se conjetura entre laminina-11 y COL(IV), provee la fuerza mecánica de la MBG, además juntas establecen un andamio para que se incorporen los componentes de la matriz extracelular mencionados, así como los que influyen en la señalización y función celular (principalmente, receptores acoplados a proteínas G, GPCR; receptores proteincinasas, RTKs; proteincinasa dependiente de adenosin monofosfato cíclico, PKA; proteincinasa C, PKC; vía receptores cinasas treonina/serina, RKTre/Ser). Por otra parte, los proteoglicanos de la MBG consisten de glicosaminoglucanos unidos entre sí de forma covalente en toda la lamina electrodensa. Los HSPGs forman un agregado físicamente

geloso en forma de “malla” que se mantiene altamente hidratado por moléculas de agua atrapadas en el intersticio de la matriz, dicha estructura sirve de reservorio para factores de crecimiento (factor de crecimiento transformador  $\beta$ , TGF $\beta$ ; y discoidina). La carga electronegativa de la MBG es impartida sobre todo por los componentes agrina y perlecano de los HSPGs; principalmente el carbohidrato heparan sulfato y la sustancia rica en ácido siálico llamada podocalixina (sintetizada por los podocitos), los cuales actúan como agentes anticoagulantes para prevenir puentes de hidrógeno y adsorción de proteínas aniónicas del plasma circundante y mantener un flujo efectivo de agua a través de las membranas celulares y acelulares; estas propiedades establecen la identidad de la MBG, ya que además de conferir filtración física, también establece ultrafiltración mediante cargas selectivas (Gubler, 2008; Kriz y Kaissling, 2008; Agtmael, 2010 ). El COL(IV) integra todos estos componentes moleculares mencionados en una alta organización de arquitectura supramolecular, estableciendo principalmente una barrera física frente al paso de grandes moléculas desde la sangre hacia el espacio urinario (Miner, 1999; Pöschl *et al*, 2003).

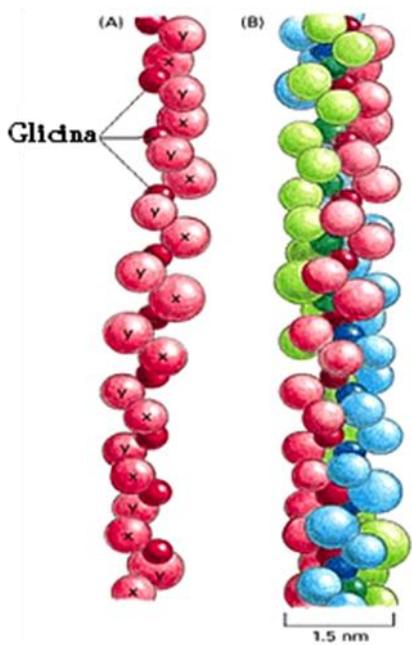
La importancia de la MBG ha sido demostrada por la identificación de enfermedades glomerulares hereditarias vinculadas a anomalías estructurales en alguna de estas biomoléculas, principalmente del COL(IV) (Stevens y Lower, 2006).

### 1.1.3 Colágeno tipo IV

El COL(IV) es el principal componente colagenoso de las MBs, pertenece de acuerdo a su función y dominio homológico, a una familia de proteínas fibrosas de colágenos formadores de redes; su síntesis la llevan a cabo células endoteliales y podocíticas (Alberts *et al*, 2008; Shoulders y Raines, 2009). El COL(IV) forma redes poligonales que interactúan con otros constituyentes glicoproteicos. Es encontrado principalmente en

músculo, hueso, riñón, cóclea, testículos y en la lente del ojo (Kashtan, 2000; Srinivasan *et al*, 2009).

Las características primarias que rigen al COL(IV) son: su longitud, rigidez y estructura en triple hélice constituida por tres cadenas polipeptídicas, denominadas cadenas  $\alpha$  (Hudson, 2004; Alberts *et al*, 2008). Dichas cadenas comúnmente contienen uno o más dominios colagenosos que consisten en repeticiones ricas de los aminoácidos (aa) glicina (Gly)-X-Y; donde X y Y refieren a otro aminoácido, generalmente prolina (Pro) e hidroxiprolina (Hyp), respectivamente; dichos aa rigen la propiedad estructural y funcional en la conformación de la triple hélice del COL(IV). El aa Gly regula el espacio estereoquímico de la molécula cada tercer aa en toda la región central de cada cadena  $\alpha$ (IV), además permite que tres cadenas  $\alpha$ (IV) proteicas se correlacionen para establecer una torsión estrecha, ayudando con esto a formar la estructura final del colágeno plegado en triple hélice; por otro lado, la estructura en anillo del aa Pro estabiliza la conformación helicoidal de cada cadena  $\alpha$ (IV), debido al carbono en la posición cuatro dentro del anillo de este aa, el cual establece la torsión en un ángulo fijo  $\phi$  y  $\psi$  de alrededor de  $-60^\circ$ , permitiendo el pliegue que necesitan las cadenas  $\alpha$  para la configuración estérica de torsión en forma de triple hélice (fig. 3); mientras que la hidroxilación del aa Pro (por la enzima prolil 3-hidroxilasa) -solamente cuando ocupa la posición Y- incrementa dramáticamente la estabilidad térmica de la proteína (superior a  $38^\circ\text{C}$ ), debido a la incorporación de un radical hidroxilo, por otro lado también contribuye a la formación del trímero proteico, debido a que permite enlaces mediante puentes de hidrogeno exactamente en el punto de torsión de las tres cadenas  $\alpha$ (IV) (Shoulders and Raines, 2009; Agtmael y Tuderman, 2010).



**Figura 3.** Estructura típica del colágeno con cada aminoácido representado por una esfera. (A) Cadena  $\alpha(IV)$  con el triplete de repetición Gly-X-Y. (B) Conformación en triple hélice de tres cadenas  $\alpha(IV)$  enrolladas de forma levógira (Tomada de National Center for Biotechnology Information, 2011).

El COL(IV) se compone de seis cadenas  $\alpha(IV)$  que configuran desde  $\alpha 1(IV)$  hasta  $\alpha 6(IV)$ ; estas cadenas muestran una elevada homología en su secuencia del ácido desoxirribunucleico (ADN), con lo cual es entendible que comparta características estructurales básicas: un residuo colagenoso flanqueado por residuos no colagenosos (Stevens y Lowe, 2006). Mediante análisis homológicos de secuencias en varias especies de mamíferos, se ha determinado la similitud de las seis isoformas de COL(IV); definiendo la relación de un gen ancestral común, que padeció tres duplicaciones consecutivas para dar como resultado las seis isoformas evolutivamente relacionadas. Esta similitud relaciona a las cadenas  $\alpha(IV)$  en dos grupos:  $\alpha 1$ , que incluye a los *locis* *COL4A1*, *COL4A3* y *COL4A5*; y  $\alpha 2$ , que incluye a *COL4A2*, *COL4A4* y *COL4A6*. Esta organización es ubicua a COL(IV), de tal forma que un miembro del grupo  $\alpha 1$  es par de alguna cadena  $\alpha(IV)$  del grupo  $\alpha 2$  (Khoshnoodi *et al*, 2008). Estas cadenas interactúan entre sí, ensamblándose de manera específica para formar uno de los tres distintos

heterotrímeros posibles:  $\alpha 1-\alpha 1-\alpha 2(\text{IV})$ ,  $\alpha 3-\alpha 4-\alpha 5(\text{IV})$  y  $\alpha 5-\alpha 5-\alpha 6(\text{IV})$ . Las cadenas  $\alpha 5(\text{IV})$  y  $\alpha 6(\text{IV})$  son abundantes en las membranas basales del músculo liso del esófago, útero, vejiga, epidermis y algunos vasos sanguíneos; las cadenas  $\alpha 3(\text{IV})$  y  $\alpha 4(\text{IV})$  se restringen más hacia algunos tejidos, como el cerebro, pulmones, ojos, testículos y riñones; en este último órgano, las cadenas  $\alpha 3(\text{IV})$  y  $\alpha 4(\text{IV})$  conforman la MBG, y en otros tejidos, la MB de la cóclea, MB del cristalino, membrana de Descemet y membrana de Bruch; sin embargo, están ausentes en el mesangio, MB vascular, MB de la capsula de Bowman y MB epidérmica. Es común encontrar que todas las MBs que expresan  $\alpha 3(\text{IV})$  y  $\alpha 4(\text{IV})$ , también expresan  $\alpha 5(\text{IV})$  y  $\alpha 6(\text{IV})$ , pero no viceversa (Sado, *et al*, 1998; Kobayashi y Uchiyama, 2003; Hudson, 2004).

La biosíntesis y la expresión de los genes que codifican para las cadenas  $\alpha(\text{IV})$  están sujetas a una regulación temporal de la MBG. Durante el desarrollo y maduración glomerular (es decir, desde el estadio embrionario hasta el inicio de la infancia), la síntesis de las cadenas  $\alpha(\text{IV})$  la llevan a cabo todas las células constituyentes del glomérulo (células endoteliales, podocíticas y en mucho menor grado las mesangiales), expresando las cadenas  $\alpha 1(\text{IV})$  y  $\alpha 2(\text{IV})$ ; sin embargo, posteriormente ocurre una transición gradual en la sustitución de la síntesis, siendo exclusiva de los podocitos, es decir las cadenas  $\alpha 1-\alpha 1-\alpha 2(\text{IV})$  son sustituidas por las cadenas  $\alpha 3-\alpha 4-\alpha 5(\text{IV})$ ; el cambio resulta fundamentalmente crítico para la maduración de la MBG como barrera especializada de filtración del plasma en el riñón. Las mutaciones identificadas en  $\alpha 1(\text{IV})$  y/o  $\alpha 2(\text{IV})$  sugieren una alteración substancial de las células endoteliales que en el feto suelen ser letales, mientras que las mutaciones dirigidas hacia cualquiera de las cadenas  $\alpha 3(\text{IV})$ ,  $\alpha 4(\text{IV})$  y/o  $\alpha 5(\text{IV})$  implican una alteración de los podocitos, “arrestando” el interruptor de expresión de estos genes para la maduración de la MBG, aunque en este caso las mutaciones resultan ser más benignas. Aun se desconoce el mecanismo que relaciona esta transición de cambio en una u otra expresión génica de

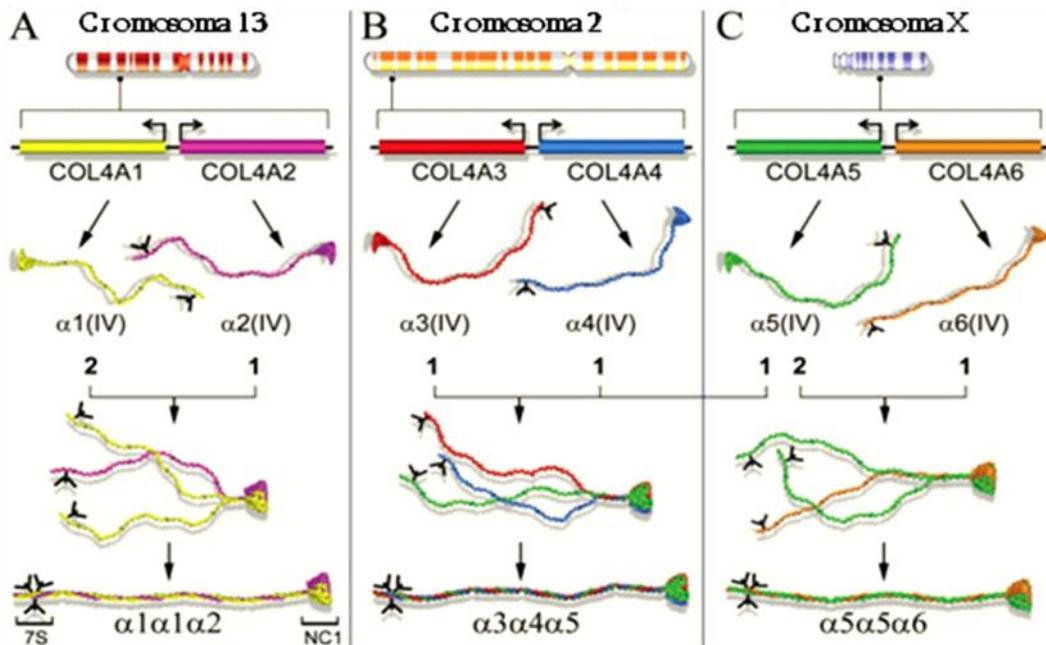
las proteínas trimericas participantes en los diferentes estadios durante el desarrollo del individuo (Boutaud *et al*, 2000; Abrahamson *et al*, 2009).

Alteraciones encontradas en cualquiera de los tres aa mencionados que conforman el triplete Gly-X-Y, generalmente se conjeturan en modificaciones estructurales en la cadena proteica; principalmente mutaciones que recaen sobre el aa de Gly (Voskarides *et al*, 2008), debido a su capacidad de encajar en la estructura interna cuando se conforma el plegamiento en triple hélice entre cadenas  $\alpha$ , ya que establece ángulos de rotación  $\phi$  y  $\psi$  de alrededor de  $-90^\circ$  (Voet *et al*, 2006); por otro lado mutaciones en los residuos X y Y, desestabilizan la fuerza mecánica, debido a que no se efectúan de manera correcta los enlaces de hidrogeno necesarios para plegar tres cadenas  $\alpha$ , provocando regiones de baja estabilidad que pueden ocasionar el rompimiento de la triple hélice. De este modo, patologías asociadas a estos tres aa, tienden a fenotipos más severos (Kuhn, 1994; Voskarides *et al*, 2008).

Los genes que conforman al COL(IV) se encuentran organizados en tres pares, orientados cabeza a cabeza de sus extremos 5'. Están separados por una corta región promotora que comparten ambos genes: *COL4A1* y *COL4A2* en el cromosoma 13, *COL4A3* y *COL4A4* en el cromosoma 2, y *COL4A5* y *COL4A6* en el cromosoma X. Estos pares de genes son transcritos en direcciones opuestas (característica única de COL(IV)), y posteriormente se traducen en cadenas  $\alpha$ (IV) individuales. Cada cadena  $\alpha$ (IV) contiene un dominio N-terminal no colagenoso (7S), un largo dominio colagenoso y un dominio globular no colagenoso en el C-terminal (NC1) (Boutaud *et al*, 2000).

Las características estereoquímicas que presentan cada cadena  $\alpha$ (IV) permiten determinar las cadenas que interactúan para formar un heterotrímero, este reconocimiento específico es proporcionado por el dominio NC1; seguido por un superenrollamiento de la triple hélice que culmina hasta el dominio 7S.

Aproximadamente 56 posibles heteotrímeros podrían formarse, pero solamente existe especificidad para tres combinaciones posibles (fig. 4) (Khoshnoodi *et al*, 2008).



**Figura 4.** Localización, organización, productos proteicos y ensamblaje de las seis diferentes isoformas de las cadenas  $\alpha(IV)$  (Tomada de Khoshnoodi *et al*, 2008).

El COL(IV) está directamente relacionado principalmente en la patogénesis de tres enfermedades:

- Síndrome de Goodpasture. Consiste en un desorden autoinmune caracterizado por hemorragias pulmonares y/o rápida glomerulonefritis progresiva; es causado por anticuerpos anti- $\alpha 3$  (IV) que se unen al dominio NC1 de la MBG (Badenas *et al*, 2002).
- Nefropatía de la membrana basal delgada (NMBD). Radica en una nefropatía hereditaria recesiva o rara vez dominante, caracterizada por un adelgazamiento de la MBG y que no lleva a insuficiencia renal (Heidet *et al*, 2001).

- Síndrome de Alport (SA). Esta enfermedad presenta una nefropatía hereditaria progresiva, caracterizada por un irregular adelgazamiento, engrosamiento y fragmentación de la MBG. A menudo se asocia con sordera y síntomas oculares anormales (Heidet *et al*, 2001; Badenas *et al*, 2002).

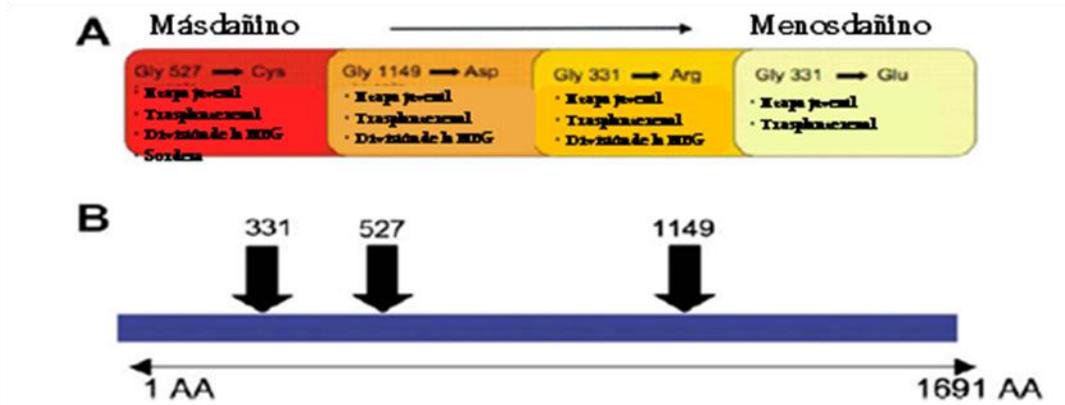
### 1.1.3.1 Biosíntesis del colágeno tipo IV

La biosíntesis del COL(IV) presenta en el retículo endoplasmático (RE) y después en el aparato de Golgi, modificaciones enzimáticas altamente coordinadas en cada cadena  $\alpha$ (IV), interacciones que en un inicio son dirigidas en gran medida por los dominios no colagenosos localizados en los extremos proteicos. Dichos cambios establecen la eliminación del péptido señal post-traducciona l cuando la cadena  $\alpha$ (IV) ingresa al RE, además se configura el reconocimiento y plegamiento mediante la chaperona Hsp47 de las tres cadenas  $\alpha$ (IV) para conformar una molécula heterotrimérica, denominada protómero (como en el caso de las cadenas  $\alpha$ 3- $\alpha$ 4- $\alpha$ 5(IV) en la MBG); la formación inicia a través de la interacción de los dominios NC1 de cada cadena  $\alpha$ (IV) mediante puentes disulfuro de 12 residuos de cisteína (Cys) conservados, estableciendo de este modo la afinidad relativa de reconocimiento del tipo de cadena  $\alpha$ (IV) interactuante. Posteriormente en el Aparato de Golgi, se llevan a cabo una serie de modificaciones a lo largo de cada cadena  $\alpha$ (IV), los cuales incluyen la hidroxilación de residuos de lisina (Lys) y Pro, glicosilación de residuos específicos de hidroxilisina (Hyl), adición de oligosacáridos ricos en manosa (Khoshnoodi *et al*, 2008; Agtmael y Tuderman, 2010).

Una vez plegado y estabilizado el protómero, es secretado mediante vesículas de transporte hacia la matriz extracelular para ser parte de la MBG. En este sitio, se establece un conjunto complejo de interacciones, donde conjuntos de protómeros tienen lugar para dar forma a una red de COL(IV) que asemeja una “rejilla”. Esto es logrado debido a que dos protómeros interactúan para formar un dímero (constituyendo un

hexámero de dominios NC1), y a su vez, cuatro protómeros se asocian a través de los dominios 7S, constituyendo un tetramero (Boutaud *et al*, 2000; Khoshnoodi *et al*, 2008); posteriormente ocurren las siguientes modificaciones: oligosacáridos se enlazan en el dominio 7S, mientras que disacáridos se unen a las hidroxilisinas a lo largo de toda la molécula, la conservación evolutiva de la glicosilación ligada a los sitios 7S en todas las cadenas, sugiere que esta modificación es fundamental para el ensamblaje y alineamiento correcto de la red del COL(IV) en las MBs. Posteriormente, la red de COL(IV) obtiene mayor estabilidad mediante la introducción de otros enlaces: el dominio 7S une inter-cadenas  $\alpha$ (IV) a través de puentes disulfuro y enlaces covalentes entre residuos Lys e Hyl (Agtmael y Tuderman, 2010).

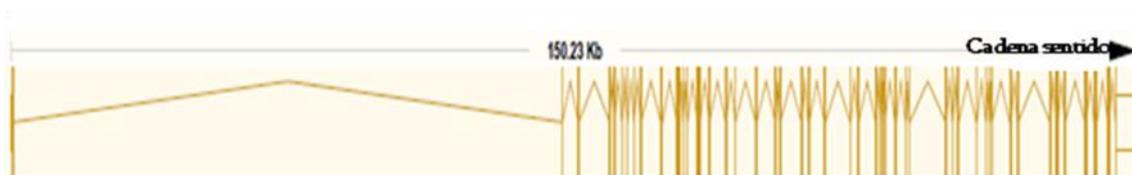
Alteraciones mutacionales asociadas a la MBG, ocasionan una serie de efectos significativos en la geometría y propiedades mecánicas del protómero (conformado por las cadenas  $\alpha$ 3- $\alpha$ 4- $\alpha$ 5(IV)), estos cambios establecen deformidad con grandes aberturas en el trímero colagenoso, ocasionando un desequilibrio estructural en su rigidez cuando es ensamblado en la MBG. Modificaciones mutacionales o incluso ciertos polimorfismos que recaen en la modificación Gly-X-Y en cualquiera de las cadenas  $\alpha$ 3- $\alpha$ 4- $\alpha$ 5(IV), se correlacionan con severidad fisiológica y patológica. 4 tipos de mutaciones de acuerdo a la repuesta estructural del protómero son consideradas (fig. 5) (Srinivasan *et al*, 2009).



**Figura 5.** Severidad nanomecánica del protómero en SA. A. Diferentes tipos de mutaciones de sentido erróneo. B. Localización de estas mutaciones en una cadena  $\alpha$ (IV) (Tomada de Srinivasan *et al*, 2010).

## 1.2 Gen *COL4A3*

El gen *COL4A3* (clasificación ENSG00000169031; [www.ensembl.org/index.html](http://www.ensembl.org/index.html), 2011) se localiza en el cromosoma 2 en la región q36-37 (228,029,281-228,179,508), contiene 52 exones y 52 intrones, los cuales abarcan una longitud de 8,097 pares de bases (pb). Los exones e intrones se encuentran flanqueados por regiones 5' y 3' no traducibles (UTR). Se han identificado hasta 10 transcritos diferentes; sin embargo, generalmente este gen codifica para 1,670 residuos aa -de acuerdo al transcrito ENST00000396578- (fig. 6). La proteína tiene un peso molecular de 161.812 kilodaltones (Kda) y su punto isoeléctrico es de 9.6497 (<http://www.ensembl.org/index.html>, 2011).



**Figura 6.** Modelo esquemático de la cadena sentido 5' del gen *COL4A3*, en el cual muestra los exones (líneas horizontales), intercalados con los intrones (líneas en forma de pico) (Tomado de Ensembl Genoma Browser, 2011).

El gen *COL4A3* contiene en el extremo 5' UTR regiones ricas de GC que abarcan más de 1000 nucleótidos, con un contenido de GC de 60-80% y con muchas islas de dinucleótidos CpG (la “p” indica que están enlazados por un fosfato), en estas islas se encuentra embebido el primer exón de este gen. La región promotora que comprende cajas con contenido GC, CTC y CCAAT, se encuentran separadas a 5 pb de la primera región promotora del gen *COL4A4*. La actividad promotora de cada caja genera transcritos alternativos que delimitan el ensamblaje de las MBs de manera específica en diferentes órganos y tejidos (Momota *et al*, 1998).

En el gen *COL4A3*, el resultado completo de la traducción principalmente es de 1,670 aa; sin embargo y posteriormente, el número de residuos se reduce a 1,642 aa (161,753 KDa), cuando la proteína es incorporada a la MBG. La proteína contiene un péptido señal de 28 aa ricos en leucina (Leu), un dominio colagenoso de 1,410 aa que comienza con una secuencia no colagenosa 7S de 14 aa que incluye 4 Cys y un dominio no colagenoso de 232 aa en el dominio NC1. Existen 23 interrupciones cortas no colagenosas en la secuencia repetida Gly-X-Y del dominio colagenoso. Por otro lado, a lo largo de toda la proteína se encuentran 5 secuencias que consisten en Arg-Gly-Asp (arginina, glicina y asparagina; respectivamente), las cuales atribuye su función de unión a las proteínas de adhesión integrinas (fig. 7) (Moriyama *et al*, 1994).



**Figura 7.** Secuencia de nucleótidos que muestra los aa codificantes del gen *COL4A3*. Sobre la línea se muestra la secuencia de nucleótidos; y por debajo de la misma, la de aA. El número de residuos comienza por el iniciador de la traducción (metionina (M)). El sitio de la peptidasa señal está indicado por una flecha. Los residuos conservados de Cys se encuentran encerrados en un círculo, y las 23 interrupciones de la secuencia repetida Gly-X-Y están encerradas (desde el N-terminal), en cuadros numerados indicados con números romanos. Tanto las secuencias no colagenosas de los 14 residuos en el N-terminal, como las que se encuentran en el dominio NC1, se muestran encerradas. (Tomada de Mariyama *et al*, 2001).

El tamaño de los exones dentro del gen *COL4A3* resulta variable, siendo por ejemplo, de 27 pb para el exón 8 y de 210 pb en el caso del exón 48 (excluyendo las regiones UTR 5' y 3'); por otro lado, el tamaño del intrón 1 queda por ser determinado con precisión debido a su gran longitud. Los 28 aa que forman el péptido señal son codificados por un solo exón de 249 pb, y que también codifica el primer aminoácido del dominio 7S. Los exones 2 al 47, y la parte 5' del exón 48, codifican para el dominio colagenoso. Los exones 2 al 19 inician con un codón de Ala, seguido de más de cuatro repeticiones que consisten en el tripéptido Gly-X-Y. Los 232 aa del dominio NC1 es codificado por el extremo 3' del exón 48, completando la estructura los exones 49-52 (Mariyama *et al*, 1994; Heidet *et al*, 2001).

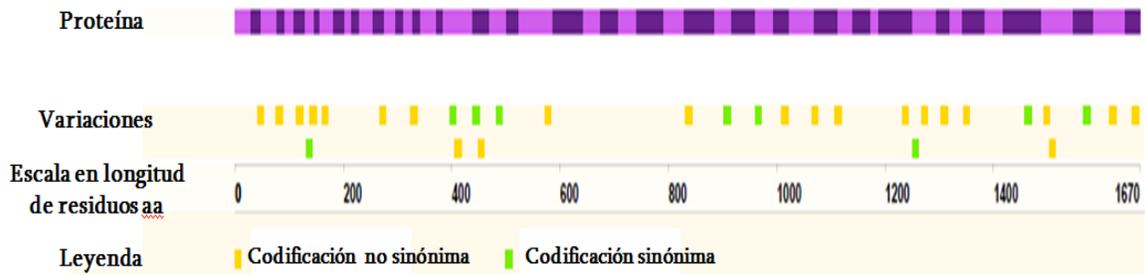
En la actualidad, la expresión diferencial del gen *COL4A3* se ha demostrado en 237 experimentos en 118 diferentes partes del organismo humano que involucran a sistemas, órganos y tejidos (fig. 8) (<http://www.ebi.ac.uk/>, 2011).



**Figura 8.** Comparación promedio entre el número de estudios publicados en donde *COL4A3* muestra sobreexpresión (sección de los cuadros en rojo), y aquellos en los que se observa baja expresión (sección de los cuadros en azul) en distintos niveles de organización en el ser humano (Tomada de European Bioinformatics Institute, 2011).

En algunas alteraciones, el empalme alternativo que presenta el dominio NC1, puede afectar el marco de lectura o introducir prematuramente codones de paro, esta última característica se encuentra intrínsecamente relacionada con SA, afectando la porción de la proteína que es fundamental en el reconocimiento para el ensamblaje específico con las cadenas  $\alpha 4(\text{IV})$  y  $\alpha 5(\text{IV})$  (Feng *et al*, 1994).

Hasta el momento se han identificado 1190 variaciones de polimorfismos de un solo nucleótido (SNPs) en todo el gen, 31 en la región exónica, las cuales a su vez constituyen 22 variantes de sentido erróneo y 9 raras (fig. 9) (<http://www.ensembl.org/index.html>, 2011 y <http://www.ebi.ac.uk/>, 2011).



**Figura 9.** Representación esquemática de los SNPs localizados a lo largo de la proteína (empalme de los 52 exones) (Tomada de Ensembl Genoma Browser, 2011).

### 1.3 Nefritis hereditaria

La nefritis hereditaria es una de las múltiples patologías renales que causan daño glomerular, debido a un grupo de desordenes genéticos de los genes que involucran la composición de los capilares glomerulares. Clínicamente se caracteriza por comienzo de hematuria persistente durante la infancia, disminuyendo de manera progresiva la tasa de filtración glomerular hasta el momento en que se inicia el tratamiento de reemplazo renal (Wein *et al*, 2008).

El hallazgo de hematuria microscópica o proteinuria puede representar la presencia de patologías serias del sistema renal o urinario y debe ser foco de una evaluación extensa. La hematuria se presenta en enfermedades que incluyen los síndromes de Epstein y de Fechtner, riñones policísticos, hipercalcinuria y otras formas de urolitiasis hereditarias; por otra parte, el SA, hematuria familiar benigna (HFB) y la NMBD son las principales formas de hematuria familiar persistente en al menos 1% de la población (Kashtan, 2009). Por otro lado, la hematuria asintomática se presenta en menos del 1% de los individuos entre la 3ª y 4ª década de la vida, y del 13% para adultos mayores; en cambio la proteinuria es menos frecuente (Polo *et al*, 2004).

En todos los pacientes con SA, alrededor del 50% de los pacientes con NMBD y cerca del 10% de los casos con HFB llevan mutaciones en alguna de las cadenas  $\alpha(IV)$ .

Los individuos afectados con NMBD y HFB (además de hematuria microscópica), también presentan signos similares de proteinuria mínima, función renal normal y adelgazamiento uniforme de la MBG con división en sus segmentos ó lamelación. Por lo general, el curso clínico suele ser benigno y de buen pronóstico. En ocasiones la HFB y la NMBD suele ser el inicio del desarrollo de SA, sobre todo en SA autosómico dominante (SAAD) y síndrome de Alport autosómico recesivo (SAAR), respectivamente (Tryggvason y Patrakka, 2006 y Savige *et al*, 2003). Algunos reportes enfatizan que la NMBD es una causa de HFB que co-segrega mutaciones en los genes *COL4A3/COL4A4* en el 40% de los casos; haciendo hincapié en que la NMBD es una condición benigna con excelente pronóstico, solo asociado ocasionalmente con falla renal; ya que en pacientes en todas las edades, aproximadamente el 38% que portan mutaciones heterocigotas en *COL4A3/COL4A4* desarrollan falla renal y 19.5% progresa a insuficiencia renal crónico terminal (IRCT). Esto sugiere que el término “benigno” es un concepto equivoco por la heterogeneidad clínica y genética que presentan los pacientes afectados de la MBG, lo cual engloba a la afección autosómica dominante o recesiva del SA (Voskarides *et al*, 2008). En SA se presentan episodios macroscópicos de hematuria, proteinuria e insuficiencia renal progresiva con complicaciones extra renales que incluyen sordera y lesiones oculares (Thorner, 2007).

En el 2005, un estudio de la Asociación Pediátrica de Trasplante Renal de Norte América, reporta que el 2.2% de los trasplantes renales pediátricos, tuvieron un diagnóstico de nefritis hereditaria, y de los cuales, el 1.8% se encontraban en diálisis crónica y el 1.4% con insuficiencia renal crónica (Kashtan, 2006).

### 1.3.1 Síndrome de Alport

En 1927, Cecil Alport fue el primero en describir la enfermedad en una familia con nefritis hereditaria, resaltando que la sordera era una característica marcada en casi

todos los casos. En el género masculino observó nefritis y sordera; mientras que en el femenino sordera y hematuria, pero con mayor sobrevida. En 1973, Gerald Spear propuso que las anormalidades renales, oculares y cocleares, debían ocurrir en un *locus* que rigiera la estructura del colágeno en la MB, y que varias formas de las nefropatías podían ser el reflejo fenotípico de la heterogeneidad genética de la enfermedad (Kashtan y Michael, 1996).

Los pacientes quienes presentan anormalidades estructurales de la MBG y por lo tanto, hematuria aislada, pueden ser razón de una de las tres etiologías diagnósticamente diferenciales, aunque estas compartan una homología clínica. Éstas son: SA, el cual puede mostrarse con episodios macroscópicos de hematuria durante la infancia, y es mejor explicado por un historial familiar crónico renal; nefropatía por inmunoglobulina A (IgA) o anti-MBG (anticuerpos contra la MBG), el cual no presenta un historial familiar, pero también puede presentarse con episodios de hematuria macroscópica; y por último, la NMBD, que en muestras de biopsias renales analizadas por microscopia electrónica, muestran una atenuación difusa de la MBG que es idéntica a la visualizada en el inicio de SAAR; sin embargo, aproximadamente dos tercios de los pacientes con NMBD tiene una forma autosómica dominante de la enfermedad, designado como HFB, ya que el patrón de herencia de esta última afección, como ya fue mencionado, es dominante y también suele ser de excelente pronóstico durante el transcurso de la enfermedad (Buzza *et al*, 2001; Badenas *et al*, 2002; Tazón *et al*, 2003 y Deltas, 2009).

La alteración de la MBG corresponde de forma similar a las lesiones de las MBs de otros órganos: se han descrito anomalías de la capsula de la cámara anterior del ojo y del epitelio pigmentado de la retina, lo cual sugiere alteraciones de la MB de Bruch; estas MBs comparten similitudes autogénicas, ultraestructurales e inmunohistoquímicas con la MBG; lo mismo ocurre con la membrana de Descemet de la cornea, la cual posee antígenos relevantes para la enfermedad de SA. Las anomalías ultraestructurales

de la MB, muestran el cambio de la estructura normal trilaminar por una lamina densa homogénea, mas la presencia de roturas y fragmentación de la misma con adelgazamientos y lamelaciones irregulares, lo que supone una modificación de la composición bioquímica de las glicoproteínas de la MBG, lo cual podría explicar las manifestaciones clínicas renales, así como la presencia de manifestaciones auditivas y oculares por provenir de órganos con MBs semejantes (McCarthy y Maino, 2000).

Los subtipos de nefropatías relacionadas al COL(IV) en SA, son: SA vinculado al cromosoma X (SALX), SAAR y SAAD. El SA está caracterizado por complicaciones renales, oculares y en la cóclea. La enfermedad progresiva renal va de hematuria a proteinuria, insuficiencia renal progresiva e IRCT en todos los hombres con SALX, y en todos los hombres y mujeres con SAAR y SAAD. La sordera neurosensorial progresiva está usualmente presente durante la infancia o al comienzo de la adolescencia. Los resultados oculares incluyen lenticono anterior (prácticamente patognómico), maculopatía (manchas blancuzcas o amarillentas o granulaciones en la región perimacular), vesículas endoteliales corneales (distrofia polimorfa posterior), y constante erosión corneal (Kashtan, 2010).

Las características clínicas y familiares que suponen el diagnostico del SA son:

**Renal.** Hematuria. En SALX, el 100% de los hombres y más del 90% de las mujeres afectadas tienen microhematuria; en SAAR, el 100% de ambos sexos presentan hematuria; mientras que en SAAD, el 90-95% de hombres y mujeres se observa hematuria benigna. Los episodios de macrohematuria no son inusuales. Proteinuria, hipertensión y desarrollo de insuficiencia renal progresiva en todos los hombres con SALX, y en todos los hombres y mujeres con SAAR.

**Ocular.** Lenticono anterior. Ocurre en el 15%-20% con SALX ó SAAR, y de manera general aparece antes de la adolescencia o al comienzo de la edad adulta. Cataratas subcapsulares. Ocurren en la gran mayoría de los casos

**Coclear.** Sordera neurosensorial bilateral a altos tonos. Tanto en hombres y mujeres con SALX como en SAAR, se presenta una alta frecuencia de sordera neurosensorial bilateral durante la infancia y al comienzo de la adolescencia.

**Historia familiar.** Es frecuentemente positivo para hematuria, sordera y/o IRCT. Hematuria o falla renal pueden resultar negativos debido a que el 10%-15% de los hombres con SALX presentan mutaciones de *novó* y aproximadamente el 15% de los individuos afectados tienen SAAR y en un 5% SAAD.

La historia natural que comprende el espectro de fenotipos que conllevan a insuficiencia renal progresiva con anormalidades extrarenales, son:

**Renal.** La característica distintiva de este síndrome es la microhematuria. Hombres con SALX tienen microhematuria persistente al comienzo de la vida. Episodios de macrohematuria suelen ocurrir, especialmente durante la infancia. Más del 90% de las mujeres con SALX tienen hematuria microscópica, aunque puede ser intermitente. Individuos con SAAR tienen microhematuria persistente, sin diferencia de género; los portadores heterocigotos con mutaciones en *COL4A3* o *COL4A4* tienen una incidencia estimada del 50% de microhematuria persistente o intermitente. Todos los hombres con SALX desarrollan proteinuria y, eventualmente, insuficiencia renal progresiva, el cual lleva a IRCT. En general, se estima que un 60% llega a IRCT a los 30 años de edad, y 90% a los 40 años. La tasa de progresión a IRCT es influenciada por la naturaleza de las mutaciones en *COL4A5*. Aproximadamente el 12% de las mujeres con SALX desarrolla IRCT antes de los 40 años, incrementando 30% a la edad de 60 años y 40% a la de 80 años. La mayoría de los individuos con SAAR desarrollan proteinuria significativa

durante la infancia o al comienzo de la adolescencia e IRCT antes de los 30 años. El progreso a IRCT ocurre a un ritmo lento en individuos con SAAD que aquellos con SALX o SAAR (Jais *et al*, 2000; Jais *et al*, 2003; Kashtan, 2006 y Kashtan, 2010).

**Ocular.** Lesiones oculares son comunes en este síndrome, tiene una ocurrencia entre 30-40% en individuos con SALX. El espectro de lesiones oculares parece ser similar tanto en SALX como en SAAR, mientras que en SAAD parecen ser relativamente poco comunes. En la aparición de lenticono anterior, la porción central de la lente sobresale de la cámara anterior, esto es prácticamente patognómico de SA. Cuando está presente, el lenticono anterior tiene una consecuencia bilateral en aproximadamente 75% de los individuos. Está ausente al nacimiento, usualmente aparece durante la segunda o tercera década de la vida. La distorsión progresiva de la lente puede ocurrir en compañía del incremento de miopía. La opacidad del cristalino puede presentarse junto al lenticono, resultando ocasionalmente de la ruptura en la capsula anterior del cristalino. En todos los individuos reportados con lenticono anterior quienes han sido adecuadamente examinados han mostrado evidencia de nefritis crónica y sordera neurosensorial. Es mucho más común en hombres afectados, aunque también puede ocurrir en mujeres. La frecuencia de lenticono anterior en hombres con SALX fue del 13% en una serie grande, y su aparición está relacionada con la mutación causante. La maculopatía consiste de manchas blancuzcas o amarillas, o granulaciones en la región perimacular que fueron encontradas en aproximadamente el 14% de los hombres con SALX en una gran serie. La maculopatía no parece estar asociada con ninguna anomalía visual. Las vesículas endoteliales corneales (distrofia polimorfa posterior) y erosión corneal recurrente también pueden ser vistos en individuos con SA (Kashtan, 2010).

**Aneurismo de la aorta torácica y abdominal.** Ha sido descrito en un pequeño número de hombres con SA (Lyons *et al*, 2007 y Kashtan, 2010)

**Coclear.** La sordera en este síndrome nunca es congénita. La disminución de la audición es usualmente detectable durante la infancia o al comienzo de la adolescencia en varones con SALX. En sus primeras etapas, el déficit de audición es detectable solo por audiometría, con reducción bilateral sensibilidad a los tonos entre 2000-8000 Hz. En hombres afectados, la sordera es progresiva y eventualmente se extiende a otras frecuencias, incluyendo aquellas de conversación coloquial. La sordera es frecuentemente identificable por evaluación normal de la audición durante la infancia, pero en algunas familias no es detectable relativamente durante el transcurso de la vida. La sordera neurosensorial progresiva se desarrolla en el 80-90% de los hombres a la edad de 40 años. El curso de la sordera depende de la mutación causativa. La discapacidad auditiva en miembros de la familia está siempre acompañada por evidencia de complicación renal, pruebas no convincentes de que hombres sordos carentes de enfermedad renal pueden transmitir el SA a su descendencia. En mujeres con SALX la sordera es menos frecuente y tiende a ocurrir durante el transcurso de la vida. Es indistinta la diferencia de géneros en la incidencia o curso de sordera en SAAR. Individuos con SAAR exhibe típicamente pérdida auditiva al comienzo de la etapa juvenil. En recientes estudios histológicos de cócleas en individuos afectados con este síndrome sugiere defectos de adhesión de el órgano de Corti a la membrana basilar que tienen como consecuencia la pérdida auditiva (Jais *et al*, 2000; Merchant *et al*, 2004 y Kashtan, 2010).

De acuerdo a estas últimas características, la historia natural de la nefropatía por SA puede dividirse de acuerdo a la progresión de la enfermedad:

- I. Puede abarcar desde el nacimiento hasta la adolescencia temprana. La alteración más evidente es la hematuria, con cambios histopatológicos limitados a una proliferación mesangial y adelgazamiento de la MBG; con áreas focales de laminación y podocitos normales.

II. Además de hematuria, también presentan proteinuria; aunque el filtrado glomerular es estable. En la biopsia renal se encuentra proliferación mesangial y pueden existir áreas con esclerosis focal y segmentaria; por otro lado, el área tubulointersticial se encuentra intacta. Por microscopia electrónica se puede identificar engrosamiento difuso y laminación de la MBG con áreas extensas de fusión de pedicelos.

III. El deterioro de la función renal es notorio; en los hallazgos histopatológicos existe daño intersticial y atrofia tubular.

IV. La etapa final conduce a IRCT, lo cual requiere trasplante renal.

La supervivencia del injerto en pacientes trasplantados con SA, en algunas ocasiones, es similar a la de los pacientes con otras causas de uremia terminal; aunque en el primer año, el 3-5% de los hombres, presentan anti-MBG post-trasplante del aloinjerto y el riesgo de recurrencia en subsecuentes aloinjertos es muy alta (Medeiros *et al*, 2008).

El tiempo estimado en que un paciente puede pasar de una fase a otra depende del tipo de mutación que presente. Pacientes en los que la producción de cualquier proteína funcional se encuentra frenada por deleciones o mutaciones sin sentido, pasan a través de estas fases más rápido que aquellos en los que las mutaciones permiten una síntesis funcional de una proteína, como en el caso de las mutaciones de sentido erróneo (Kashtan, 2006 y Medeiros *et al*, 2008).

Debido a lo espuestos, el SA puede definirse como una enfermedad renal hereditaria clínica y genéticamente heterogénea (Tazón *et al*, 2003), debido a mutaciones en cualquiera de los genes implicados en esta patología ( $\alpha 3$ - $\alpha 4$ - $\alpha 5$ (IV)), transmitiéndose ya sea en sus formas autosómicas o vinculada al cromosoma X (SAAR y SAAD, y SALX, respectivamente). Las consecuencias recaen en la afección de la MBG,

coclear y del lente, debido a la incapacidad del COL(IV) de formar una red glicoproteica que sirva como barrera para regular la estructuración y remodelación de dichas MBs. Esto resulta estrechamente significativo en el hecho que presenta la MBG de separar y filtrar las diversas biomoléculas del plasma sanguíneo, y su alteración puede conllevar a un estadio de IRCT (Pescucci *et al*, 2004 y Medeiros *et al*, 2008).

La terapia farmacológica funciona únicamente para retrasar la progresión de la enfermedad renal del SA. Tal es el caso del inmunosupresor Ciclosporina A, el cual ha demostrado retrasar el daño en la MBG, disminuyendo la proteinuria y manteniendo la función renal hasta 10 años más, aunque también puede ocasionar nefrotoxicidad. El uso de inhibidores de la enzima convertidora de angiotensina, también reducen la proteinuria, estos fármacos actúan sobre el sistema renina-angiotensina-aldosterona, reduciendo el volumen sanguíneo y en consecuencia su retorno y presión de la sangre hacia los riñones; este fármaco puede incrementar un 35% la esperanza de vida de los pacientes. La administración de un receptor antagonista del TGF $\beta$  ó un inhibidor de metaloproteinasas de la matriz extracelular (MMP) (proteasas dependientes de zinc que remodelan el tejido conectivo), ha demostrado que mejoran la filtración glomerular y la morfología de la MBG; por ejemplo, la variedad MMP12 fue inhibido por un antagonista del receptor de quimiocina 2, el cual es expresado en podocitos y puede explicar como la inhibición de MMP12 reduce el daño a la MBG (Thorner, 2007). En el caso de los pacientes trasplantados que desarrollan anti-MBG, los tratamientos con plasmaféresis y ciclofosfamida tiene un beneficio limitado (Kashtan, 2006).

### 1.3.1.1 Genética del síndrome de Alport

Existen tres formas genéticas del SA:

1. SALX (clasificación OMIM # 301050). Ocurre aproximadamente en el 80% de los casos, y surge de mutaciones en el gen *COL4A5*. Alrededor de 350 mutaciones

diferentes han sido reportadas, las cuales incluyen: largas deleciones, mutaciones sin sentido y de sentido erróneo, pequeñas deleciones/inserciones causando un corrimiento en el marco de lectura y mutaciones en el sitio de empalme; mutaciones en puntos calientes son desconocidas. En pocas excepciones, se conoce que cada familia lleva una única mutación, pero en aproximadamente el 18% de los casos presentan mutaciones de *novo*.

2. SAAR (OMIM # 203780). Aparece aproximadamente en el 15% de los casos y resulta de mutaciones homocigotas en el gen *COL4A3* y/o *COL4A4*. Cerca de 30 diferentes mutaciones en estos genes han sido identificadas con algún espectro de mutaciones para *COL4A5*.
3. SAAD (OMIM # 104200). Son raros los casos y al igual que en SAAR es causado por mutaciones en los mismos genes, excepto que estos cambios son heterocigotos. Se presenta aproximadamente en el 5% de las familias. Se han identificado alrededor de 10 mutaciones (Thorner, 2007; Medeiros *et al*, 2008 y <http://www.ncbi.nlm.nih.gov/>, 2011).

La información hacia los pacientes afectados sobre el riesgo potencial de heredar la enfermedad a sus descendientes es realizado por asesoramiento genético, por lo que tres modos de herencia son reconocidos: SALX, SAAR y SAAD. En familias con SALX, las madres portadoras tienen un 50% de probabilidad de transmitir la mutación causante de la enfermedad en cada progenie; los hijos que hereden la mutación serán afectados con SA y eventualmente desarrollaran IRCT y, en algunos casos, sordera; las hijas que hereden la mutación serán portadoras y típicamente tendrán hematuria asintomática, pero pueden tener más severidad en la enfermedad renal. Los hombres afectados pasarán la mutación causante de la enfermedad a todas sus hijas y a ninguno de sus hijos. En familias con herencia autosómica recesiva, los padres de un niño afectado son heterocigotos obligados y portan un alelo mutante; en la concepción, cada descendiente

de un individuo afectado tendrá un 25% de probabilidad de ser afectado; un 50% de ser un portador que puede o no ser sintomático, y un 25% de no ser afectado ni portador de la enfermedad. En familias con herencia autosómica dominante, cada hijo de un individuo afectado tendrá un 50% de probabilidad de heredar la mutación y teniendo un desorden en el COL(IV) (SAAD o NMBD) (Kashtan, 2010).

### 1.3.1.2 Epidemiología del Síndrome de Alport

La prevalencia de este síndrome ha sido estimada en aproximadamente 1:50,000 nacidos vivos. Se estima que es la causa de IRCT en el 3% de la población pediátrica mundial, del 1-2% de los pacientes en tratamiento renal sustitutivo en Europa y un 2-3% de los pacientes trasplantados en Estados Unidos de Norteamérica (EUA) (Kashtan, 2010). Aunque por otro lado, también se considera que es la causa de uremia terminal en el 0.6-4.6% en EUA y Europa, respectivamente (Pescucci *et al*, 2004). Datos de varias series sugieren que aproximadamente una quinta parte de los niños evaluados por nefrólogos pediátricos con microhematuria aislada, reciben un diagnóstico de SA (Thorner, 2007). De acuerdo al sistema de datos renales de EUA, alrededor del 0.2% de los adultos y 3% de los niños en EUA con IRCT llevan a un diagnóstico por SA (Avenidaño *et al*, 2006). El síndrome en EUA es considerado como la segunda enfermedad hereditaria más común que causa falla renal. Estas cifras probablemente estén subestimadas debido a la dificultad de diagnóstico de casos aislados o sin sordera (Medeiros *et al*, 2008 y Srinivasan *et al*, 2009).

### 1.3.1.3 Síndrome de Alport en México

En México se han realizado pocos estudios acerca del SA; García y Gordillo (1981), reportaron a cuatro familias mexicanas con SA, provenientes de cuatro estados diferentes (Tabasco, Puebla, Michoacán y Durango) y no relacionadas entre sí, y en todos encontraron que la transmisión hereditaria era autosómica dominante; y Gordillo

(1996), identificó a familias donde la mayoría de las mujeres estaban afectadas y evolucionaron hacia la IRCT; mientras que Robles y colaboradores (1999), reportaron tres mutaciones de *novo* en el gen *COL4A5* en pacientes mexicanos; y por otro lado, Pérez y colaboradores (1999), estudiaron a 8 pacientes sospechosos de SA con hematuria, de los cuales 3 desarrollaron hipocausia neurosensorial y 7 de ellos estuvieron en diálisis peritoneal, por último, 4 de ellos fueron trasplantados y uno permaneció con función renal normal (Tazón *et al*, 2003 y Medeiros *et al*, 2008).

Las diversas mutaciones y modos de transmisión del SA, así como las ocurrencia de casos esporádicos, son responsables de la heterogeneidad clínica y genéticas de esta enfermedad (Tazón *et al*, 2003 y Medeiros *et al*, 2008).

## 2 Antecedentes

### 2.1 Estudios clínico-patológicos

Passwell y colaboradores (1981) describieron a una niña nacida de padres entre primos hermanos, quien en el primer año de vida presentó nefritis y sordera. Posteriormente fue diagnosticada con SA mediante microscopia electrónica, debido a la observación de adelgazamiento y división de la MB del glomerulo y los túbulos renales, con acumulación de partículas electrodensas dentro de la red de COL(IV). Los padres no fueron afectados, sin embargo en la historia familiar se definieron dos tíos maternos con nefritis crónica y sordera neurosensorial.

Feingold y colaboradores (1985) estudiaron a 41 familias diagnosticadas con SA, de las cuales, cuatro familias fueron sintomáticas para SAAR con la peculiaridad de presentar padres consanguíneos y no afectados; dos familias presentaron hematuria y falla renal, sordera neurosensorial en un miembro afectado, y en otro falla renal progresiva.

Mochizuki y colaboradores (1994) reportaron a cuatro familias no relacionadas entre sí y diagnosticadas con SAAR. En una familia, dos hermanos infantes (niña y niño) fueron afectados; la niña desarrollo hematuria y sordera neurosensorial a los cuatro años de edad; a los diez recibió un aloinjerto renal, sin embargo desarrollo anti-MBG, lo cual condujo a nefritis seis meses más tarde. El niño padeció de hematuria, sordera y deterioro de la función renal. En otra familia, una niña de 11 años de edad nacida de padres consanguíneos, desarrollo hematuria y proteinuria a los siete e IRCT hacia los once años, recibió trasplante renal por parte de su madre (no presentó rechazo, ni anti-MBG), posteriormente desarrollo sordera neurosensorial bilateral hacia los 13 años; los padres no fueron afectos. Observando otra familia, dos hermanas alcanzaron IRCT a los 14 años de edad, sin presentar sordera ni anormalidades oculares; previamente mediante biopsia renal en una hermana, se reveló engrosamiento focal y adelgazamiento de la

MBG, mientras que la otra hermana presento hematuria microscópica, posteriormente desarrollo síndrome nefrótico, aunque sin disminución de la función renal hasta los 11 años de edad, también fue excluida de sordera o anormalidades oculares. En la última familia evaluada, ningún miembro presento hematuria o proteinuria.

Kalluri y colaboradores (1995) correlacionaron que del 5 al 10% de los pacientes con SA que reciben trasplante renal por alcanzar IRCT presentan rechazo al órgano, debido a que los pacientes pueden desarrollar anti-MBG contra el dominio NC1 de la cadena  $\alpha 3(IV)$ . En los pacientes que no rechazan el órgano, se vincula que tiene razón en la alta tasa polimórfica en *COL4A3* y *COL4A5* en el dominio NC1, debido a que se observo que son más tolerantes a la pérdida de colágeno.

Colville y colaboradores (1997) estudiaron a una familia diagnosticada con SAAR, donde cuatro de los ocho descendientes pertenecen a un matrimonio consanguíneo; los cuatro miembros familiares presentaron falla renal y sordera alrededor de los 20 años de edad. En un examen oftalmológico, los cuatro afectados padecieron de lenticono anterior, y tres de ellos retinopatía. Una prueba realizada con marcadores intragenicos mostró resultados positivos en los *locus COL4A3* y *COL4A4*. Dichos autores refieren que las manifestaciones oculares del SAAR son idénticas a las visualizadas (en otros estudios) en pacientes afectados con SALX.

Jefferson y colaboradores (1997) reportaron a una familia con aparente SAAD. El padre presento un historial de microhematuria e hipertensión a lo largo de su vida; finalmente alcanzó falla renal antes de morir a la edad de 68 años. En la madre no hubo evidencia de hematuria o disfunción renal; solamente y mediante microscopia electrónica, logro observarse engrosamiento de la MBG. La evaluación de los hijos de ambos padres mostró microhematuria y proteinuria, sin embargo los exámenes audiológicos presentaron audiogramas normales. Por otra parte, en un pariente cercano a esta familia, durante su juventud registro un historial con microhematuria

asintomática e hipertensión, sin embargo su función renal fue deteriorándose hasta alcanzar IRCT a la edad de 35 años, posteriormente presento sordera a tonos altos.

Rhys y colaboradores (1997) observaron en el historial de tres hermanos diagnosticados con SA, ataques espontáneos de erosión corneal persistente (ECR). En dos de ellos sucedieron dos episodios en un periodo de uno a tres años. El otro hermano padeció de lo mismo, aunque en 60 episodios durante diez años. Para evaluar la prevalencia de ECR en el SA, fueron encuestados 41 pacientes con SA con falla renal, además de 67 pacientes (controles) trasplantados para otra forma de nefropatía; los resultados mostraron en siete pacientes con SA y uno de los pacientes control, que un historial por ECR se manifestó entre los 12 y 21 años de edad.

Pérez y colaboradores (1999) observaron desarrollo de anticuerpos circulantes anti-MBG en algunos pacientes con SA que habían sido trasplantados, los cuales no existían al principio de la enfermedad. La alteración implicó glomerulonefritis anti-MBG solamente en el 10% de los casos.

Kobayashi y Uchiyama (2003) encontraron en individuos clínicamente sanos mediante análisis inmunohistoquímicos realizados con biopsias de riñón, que las cadenas  $\alpha 3(IV)$ ,  $\alpha 4(IV)$  y  $\alpha 5(IV)$  mostraban un patrón de tinción similar en la MBG. Por otro lado, en pacientes con SA, mutaciones en los genes que codifican para estas cadenas, conllevaron alteraciones en alguna de las tres cadenas  $\alpha(IV)$ , y por consiguiente, a su ausencia en la MBG. Estos factores ayudaron a determinar que las cadenas  $\alpha(IV)$  forman un heterotrímero en la MBG y cualquier alteración modifica el ensamble anormal de dichas cadenas  $\alpha(IV)$ .

## 2.2 Estudios moleculares

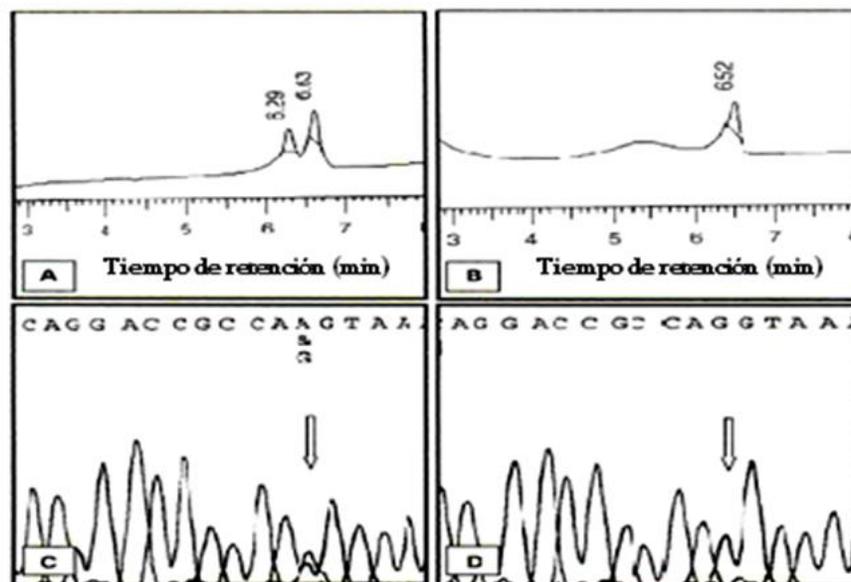
Lemink y colaboradores (1994) estudiaron a 22 pacientes no relacionados entre sí y diagnosticados con SA. Utilizando la técnica de polimorfismos de conformación de cadenas sencillas (SSCP) y posterior secuenciación de los últimos cinco exones del extremo 3' del gen *COL4A3*, fueron reveladas tres variantes: dos en distintos pacientes -hetero y homocigoto, respectivamente- que consistió en una delección de cinco nucleótidos a partir del codón 1,473 (recuérdese de los 1,670 aa habituales); el tercer paciente -heterocigoto- presentó dos diferentes mutaciones sin sentido, es decir introdujeron un paro prematuro en la secuencia proteica de la cadena  $\alpha 3$ (IV), teniendo como consecuencia una proteína truncada.

Ding y colaboradores (1995) describieron en una niña quien alcanzó IRCT con supuesta forma de SAAD, una mutación heterocigota en el gen *COL4A3*. En la determinación de dicha mutación, en primera instancia se amplificaron -mediante reacción en cadena de la polimerasa (PCR)- los últimos cinco exones en el gen de *COL4A3* y posteriormente fueron secuenciados. Los resultados revelaron una delección de siete pb a partir del codón 1448, ocasionando un corrimiento en el marco de lectura que finalizó en un codón de paro prematuro; la pérdida fue de 222 aa en el dominio NC1, por lo tanto puede que la cadena mutante no sea capaz de ensamblarse con las cadenas  $\alpha 4$ - $\alpha 5$ (IV) para conformar el protómero en triple hélice.

Knebelman y colaboradores (1995) encontraron en un paciente diagnosticado con SA una inserción de 74 pb en el transcrito del gen *COL4A3*, derivado de un elemento Alu en el intron cinco, debido a una transversión G>T en el sitio de empalme, dando como resultado una proteína truncada en estructura y función.

Van der Loop y colaboradores (2000) estudiaron a seis miembros de una familia - diagnosticada mediante microscopía electrónica de la MBG- con SA. Por medio de la

técnica retrotranscriptasa inversa (RT-PCR) se estudio la expresión de la cadena  $\alpha 3(\text{IV})$ , cromatografía líquida semi-desnaturalizante de alto rendimiento (DHPLC) y correspondiente secuenciación; se encontró en cuatro pacientes una mutación en el sitio de empalme del exón 21 (p.G493S: cambio de un aa Gly por una Ser en el residuo aa número 493 de la proteína codificada; de aquí en adelante se utilizará esta nomenclatura, esto de acuerdo a Antonarakis y Nomenclature Working Group, 1998; y Dunnen y Antonarakis, 2001) en el dominio colagenoso, debido a la substitución de una G por A en el último nucleótido de dicho exón (fig. 10). El cambio repercute el nulo reconocimiento del empalme con el exón 22, generando un salto en la secuencia de 165 pb (confirmado por la ausencia en la secuencia del ARN; ácido ribonucleico mensajero), por lo tanto suprime a 55 aminoácidos. La delección resultante también se ha identificado en *locus* de COLIII, VI y VII. Por medio del historial familiar, fue establecida la forma dominante de la enfermedad.



**Figura 10.** Análisis de la mutación en el sitio de empalme encontrada en el exón 21 del gen *COL4A3*. Resultado por DHPLC de (A) un paciente mostrando la variante mutada de forma indirecta en picos homo y heteroduplex (de derecha a izquierda, respectivamente), y (B) un caso control mostrando un solo pico homoduplex. La secuencia de esta región muestra el cambio puntual (C), en relación a un caso control (D) (Tomada de Van der Loop *et al*, 2000).

Heidet y colaboradores (2001) diagnosticaron a 41 familias con SA. La forma de SAAR fue establecida de acuerdo a los siguientes criterios: severidad de la enfermedad en mujeres jóvenes, consanguinidad familiar, hematuria microscópica en el padre de un hombre afectado, cambios típicos inmunohistoquímicos para SAAR, así como la exclusión de cualquier vínculo con el gen *COL4A5*. Solamente 60 pacientes fueron afectados con dicha forma recesiva, donde 44 alcanzaron IRCT y dos de estos desarrollaron anti-MBG después de recibir trasplante renal; 27 padecían de sordera y 16 lesiones oculares. En el estudio inmunohistoquímico de la MBG, ocho pacientes presentaron coausencia del trímero conformado por las cadenas  $\alpha3$ - $\alpha4$ - $\alpha5$ (IV), mientras que otro paciente tenía reducida la expresión de dichas cadenas. Pruebas moleculares de ligación mediada por PCR a lo largo de todo el gen de *COL4A3*, y por otro lado, mediante retrotranscriptasa inversa RT-PCR para establecer la expresión su expresión y por correspondientes secuenciaciones; se identificaron 21 mutaciones responsables de SAAR; nueve resultaron en cambios sin sentido y los restantes en cambios en el marco de lectura, considerándose ambos cambios como mutaciones nulas. En el primer caso, las variantes fueron: p.R406X -“X” refiere a un aa no determinado- (c.1216C>T: cambio de cadena sencilla de una citosina por timina en el nucleótido número 1216 del gen *COL4A3*, de aquí en adelante se utilizará esta nomenclatura, esto de acuerdo a Antonarakis y Nomenclature Working Group, 1998; y Dunnen y Antonarakis, 2001) en el exón 21, p.Q1050X (c.3148C>T) en el exón 37 y p.L1503X (c.4508T>G) en el exón 49; en las mutaciones del segundo caso, las variantes fueron: c.389insT (“ins” denota inserción de una timina en el nucleótido número 389 del gen en mención; esto de acuerdo a Antonarakis y Nomenclature Working Group, 1998; y Dunnen y Antonarakis, 2001) en el exón 7, c.2417insC en el exón 31, c.2621delGA + insT (“del” indica delección; Antonarakis y Nomenclature Working Group, 1998; y Dunnen y Antonarakis, 2001) en el exón 32, c.3533delC en el exón 41, c.4279delG en el exón 48, c.4530delATGT en el exón 49; cuatro mutaciones más fueron establecidas en el sitio de

empalme, afectando la región consenso de empalme que involucra a cuatro intrones diferentes: c.IVS7-1G>T (indica un cambio de guanina por timina, un nucleótido antes de comenzar la región codificante del exón 7; de aquí en adelante se utilizará esta nomenclatura, esto de acuerdo a Antonarakis y Nomenclature Working Group, 1998; y Dunnen y Antonarakis, 2001) en el intrón 37, c.3565+1delG en el intrón 41, c.3752-2A>T en el intrón 42, y c.3956-1G>A en el intrón 44; ocho son mutaciones de sentido erróneo, siendo una de estas mutaciones una substitución de Cys por Arg en el dominio NC1: p.R1661C (c.4981C>T) en el exón 52, este cambio no fue encontrado en 36 controles y supone que el residuo Arg esta conservado en todas las cadenas  $\alpha$ (IV), lo que probablemente altera la conformación del dominio NC1 cuando el aminoácido Cys forma puentes disulfuro intramoleculares; las otras 7 substituciones potencialmente patógenas son cambios del residuo Gly en el dominio colagenoso: p.G297E (c.890G>A) en el exón 16, p.G407R (c.1219G>C) en el exón 21, p.G640R (c.1918G>A) y p.G1167R (c.3499G>A) en el exón 40, p.G1207E (c.3620G>A) en el exón 42, p.G1277S (c.3829G>A) en el exón 43, y p.G1334E (c.4001G>A) en el exón 45. En adición a las variantes mencionadas, tres variantes raras y ocho polimorfismos también fueron detectados en el mismo gen; las variantes raras sugieren cambios de otros residuos diferentes de Gly en el dominio colagenoso, el cual no está conservado en las otras cadenas  $\alpha$ (IV) por no ser parte del triplete Gly-X-Y, mientras que las otras 8 variantes polimórficas sí sustituyen el residuo Gly en dicho triplete del dominio colagenoso.

Badenas y colaboradores (2002) reportaron a 11 familias (visualizadas por un adelgazamiento de la MBG mediante biopsia renal con inmunofluorescencia y microscopia electrónica) con HFB. Por medio de análisis conformacional de una sola cadena (SSCA) y posterior secuenciación del gen *COL4A3*, se detectaron dos mutaciones de sentido erróneo en el dominio colagenoso: p.G1015E en el exón 36 y p.G985V en el exón 35, afectando los residuos Gly, las variantes alteran la estabilidad de ensamble de la cadena  $\alpha$ 3(IV) con las cadenas  $\alpha$ 4- $\alpha$ 5(IV) para formar la estructura colagenosa en triple

hélice. Por otra parte fueron identificados otros 13 cambios alelicos, los cuales corresponden a polimorfismos y variantes raras, dichos cambios representaron el 81.2% de todas las variantes detectadas en este estudio.

Longo y colaboradores (2002) encontraron en 36 pacientes con SA mediante análisis de SSCP y posterior secuenciación, seis mutaciones de *novo* en el gen *COL4A3* dispersas en cinco familias distintas (denominadas: PAR, NCL, COZ, GST y DIT) (Fig. 11). En PAR la mutación fue sin sentido (p.R1037X), mientras que la otra involucro el codón de inicio (c.ATG>TTG), el cual probablemente se opone de inmediato a la formación de la cadena  $\alpha 3(\text{IV})$ ; en la familia NCL se observó una deleción dentro del marco de lectura de 27pb en el exón 39 (c.3386del27), causando la pérdida de nueve aa en el dominio colagenoso, por otra parte también se identificó una inserción (c.3209insA) en el exón 37, que conlleva a un cambio en el marco de lectura; mientras que en la familia COZ, una deleción (c.40del24) en el exón 1 dentro del marco de lectura, causó la pérdida de ocho aa en el péptido señal; en la familia GST, se encontró una mutación heterocigota de sentido erróneo (p.G1198S) en el exón 42; y por último, en la familia DIT, identificaron una deleción de nueve nucleótidos, causando la pérdida del tripéptido Ser-Pro-Gly en la interrupción no colagenosa del dominio colagenoso. Las tres deleciones mencionadas dentro del marco de lectura, ocurrieron en los sitios de repetidos directos de ADN en la secuencia normal del gen *COL4A3*, atribuido a un error de deslizamiento de la polimerasa en la duplicación del ADN. Además, 3 cambios de sentido erróneo fueron identificados, aunque sin vinculación patogénica significativa. Por último, fueron encontrados ocho polimorfismos, dos de ellos de *novo*.

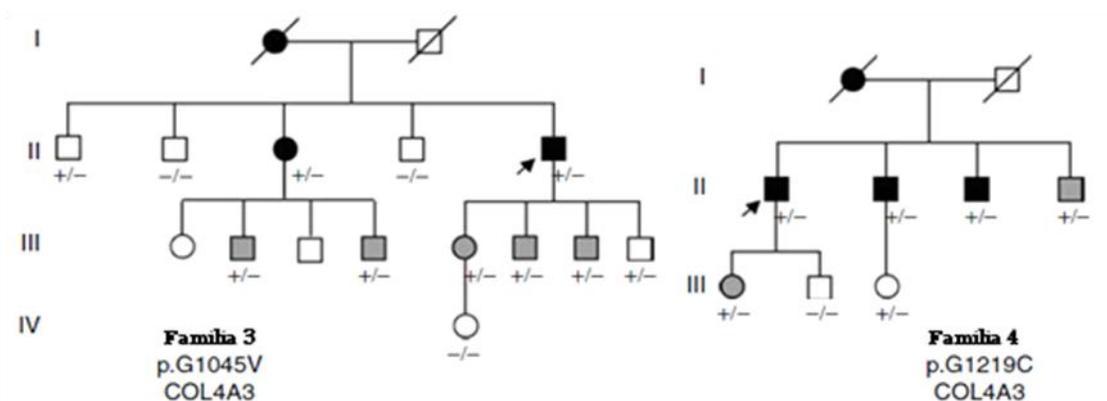


**Figura 11.** Distribución de mutaciones patológicas a lo largo del gen *COL4A3*. Las cajas grises denotan las mutaciones encontradas en fenotipos recesivos y las cajas rayadas son de fenotipos dominantes. Las llaves indican agrupan repercusiones fenotípicas similares (Tomada de Longo *et al*, 2002).

Tazón y colaboradores (2003) identificaron por secuenciación 13 mutaciones en el gen *COL4A3* con patrón de herencia recesivo, ocho de ellas de *novο*. De estas mutaciones, cinco fueron de sentido erróneo: tres involucran residuos altamente conservados de Gly, indicando que es único y suficientemente pequeño para ajustarse en el centro de la cadena  $\alpha 3(IV)$  cuando se forma la estructura en triple hélice junto a las cadenas  $\alpha 4-\alpha 5(IV)$ ; las variantes fueron: p.G1412V (c.4235G>T) en el exón 47, p.G780E (c.2339G>A) en el exón 30, p.G464V (c.1391G>T) en dos pacientes en el exón 22; mientras que en las otras dos mutaciones: p.C1548Y (c.4642G>A) en el exón 50, substituye a un residuo conservado de Cys en el dominio NC1, complicando la unión entre intracadenas  $\alpha(IV)$  mediante puentes disulfuro, y p.F1475S (c.4424T>C) en el exón 48. Además detectaron dos mutaciones sin sentido que ocasionan un corrimiento en el marco de lectura: p.A1208-stop1220 (c.3622delG) en el exón 42 y p.A1439-stop1528 (c.1528delC) en el exón 29. Finalmente, también en el exón 42 se identificó la mutación p.R1215X (c.3643C>T) en dos pacientes. En este mismo gen también fueron caracterizadas 14 variantes polimórficas, cuatro de ellas de *novο*: 3'UTR-10C>T en la

región UTR 3', p.P74P (c.222G>T) en el exón 3, p.P141L (c.422C>T) en el exón 7 y p.K834R (c.25501C>A) en el exón 32. Las variantes polimórficas en dicho estudio representaron el 66% de los cambios totales.

Pescucci y colaboradores (2004) investigaron a dos familias sospechosas de poseer SAAD en el gen *COL4A3*. La evaluación clínica de los miembros que conforman a cada familia, reveló que los individuos afectados tenían un amplio rango de fenotipos, los cuales iban desde IRCT hacia la quinta década de la vida hasta microhematuria aislada no progresiva. La identificación de mutaciones se realizó mediante cribado por DHPLC y posterior secuenciación; identificando en el exón 37 (familia tres) y 42 (familia cuatro), mutaciones de sentido erróneo en cada familia, lo cual causa la substitución del residuo Gly en el dominio colagenoso; dichos cambios fueron: p.G1045V (c.3134G>T) y p.G1219C (c.3655G>T), respectivamente. Estas mutaciones se presentaron en todos los miembros de las dos familias afectadas, entre ellos resalta un miembro de cada familia: un hijo afectado de 11 años totalmente asintomático y otro que lo fue hasta los 22 años (Fig. 12).



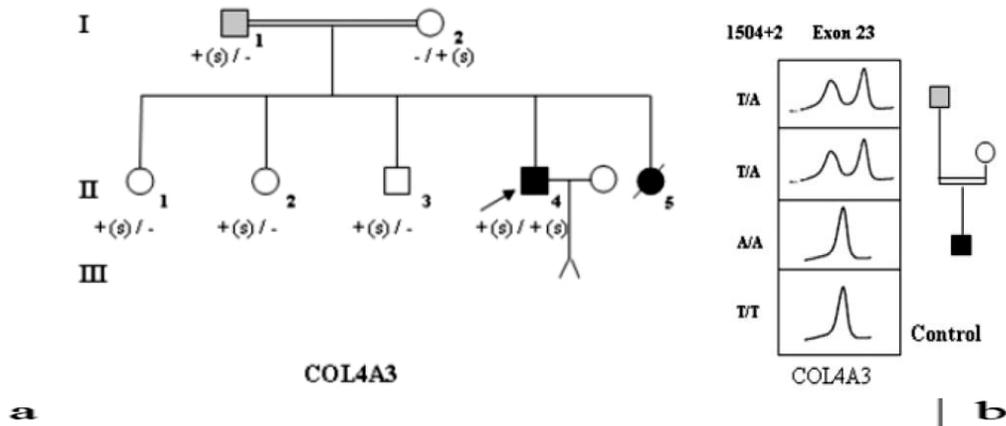
**Figura 12.** Genealogía del SAAD detectada en el gen *COL4A3*. Los cuadros representan a los miembros varones y los círculos a las mujeres. Las figuras en gris indican microhematuria y las negras micro-macrohematuria y falla renal; las figuras blancas presentan sin signos aparentes de la enfermedad. Las flechas indican los pacientes que resaltan de cada familia. La simbología de las mutaciones se muestran debajo de cada figura: -, alelo normal; +, alelo mutante. El tipo y gen mutado son indicados en la genealogía (Tomada de Pescucci *et al*, 2004).

Wang y colaboradores (2004) encontraron siete mutaciones de *novo* (representando el 11% de la población de estudio) en 62 pacientes adultos femeninos no relacionados con NMBD; fueron diagnosticadas mediante biopsia renal con microscopía de contraste de fases. Las variantes fueron detectadas por SSCP y posterior secuenciación. Las mutaciones de sentido erróneo que resultaron en sustituciones de Gly dentro del dominio colagenoso fueron: p.G532C y p.G584C (c.1594G>T y c.1750G>T, respectivamente) en el exón 25, p.G596R (c.1786G>C) en el exón 26, p.G695R (c.2083G>A) en el exón 28; mientras que la mutaciones en el sitio de empalme fueron: c.IVS2224-1C>T en el exón 30, c.IVS2980+1G>A en el exón 35 e c.IVS3518-7C>G en el exón 41. También fueron identificados cuatro polimorfismos de *novo*: p.P116T en el exón 6, p.P690P en el exón 27 y, p.G895G y p.A899A en el exón 33; y designan como polimorfismos a p.P1109S y p.Q1495R, por encontrar estas variantes heterocigotas tanto en pacientes con NMBD como en individuos sanos; dichos polimorfismos fueron independientes del signo hematuria.

Nagel y colaboradores (2005) describieron diez mutaciones de *novo* en el gen *COL4A3*, todas ellas mutaciones de sentido erróneo: p.T117X (c.351C>A) en el exón 6, p.Q371X (c.1111C>A) en el exón 19, p.R1496X (c.4486C>T) y p.R1516X (c.4546C>T) en el exón 49. Una mutación de cambio en el marco de lectura c.3547\_354insG en el exón 41. Mutaciones de sustitución en el residuo de Gly p.G149E (c.446G>A) en el exón 7, p.G532D (c.1595G>A) en el exón 25, p.G739R (c.2215G>A) en el exón 29 y p.G853R (c.2557G>A) en el exón 32. Por último, una mutación dentro de la secuencia intrónica IVS34-5A>T en el exón 34.

Longo y colaboradores (2006) investigaron a cinco familias diagnosticadas con SA, encontrando en dos de ellas la forma recesiva de la enfermedad. El análisis genético mediante cribado por DHPLC en el gen *COL4A3*, demostró en una de las cinco familias en el exón 23 un pico heteroduplex; y en posterior análisis de secuencia, se identificó

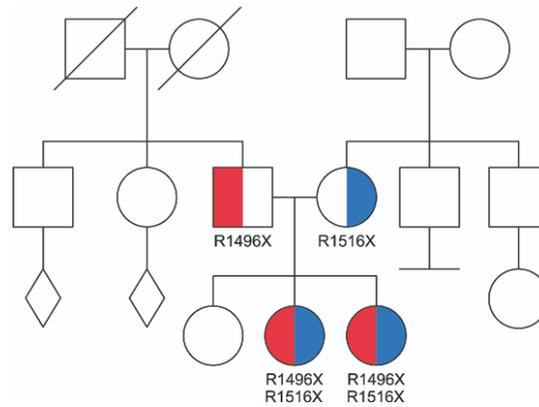
una mutación en el sitio de empalme (c.1504+2T>A), lo cual conlleva a una prematura proteína truncada. Esta mutación estuvo en estado heterocigoto en la madre, hermano y dos hermanas, y en homocigoto en el caso detectado (fig. 13).



**Figura 13.** Representación de la variable mutada en el exón 23 en el gen *COL4A3*. (a) Genealogía de la familia afectada con dicha mutación. El símbolo gris muestra hematuria aislada; los negros, micro-macrohematuria, o hipocausia o IRCT; los blancos, sin signos de la enfermedad; la flecha apunta al paciente estudiado; y las variables debajo de los símbolos (donde s=splicing). (b), muestra el análisis mutacional heteroduplex por DHPLC, del paciente en mención.

Slajpah y colaboradores (2007) evaluaron a 17 familias diagnosticadas con HFB. La identificación de las variantes alélicas se realizó mediante SSCP y posterior secuenciación del gen *COL4A3*. Los resultados mostraron en seis pacientes no relacionados entre sí y con hematuria aislada, tres mutaciones heterocigotas, dos mutaciones de sentido erróneo: p.G487C (c.1459G>T) en el exón 23 y p.G1015E (c.3044G>A) en el exón 36; y una mutación que cambia el marco de lectura (inserción de tres nucleótidos) 3547-3548insGGA en el exón 41, lo cual ocasiona la inserción de otro residuo de Gly en la primera posición de los repetidos Gly-X-Y en el dominio colagenoso. La frecuencia de mutaciones en estas familias fue de 3/40 (7.5%). Por otro lado, también encontraron ocho variantes polimórficas a partir de un grupo control conformado por 127 individuos clínicamente sanos.

Cook y colaboradores (2008) determinaron dos mutaciones de *nov* en el gen *COL4A3* en dos hermanos diagnosticados con SAAR, ambos partieron de mutaciones heterocigotas por parte de los padres (diagnosticados con SA), conduciendo a dos secuencias anormales de ADN encontradas en el exón 49: p.R1496X (c.CGA>TGA) y p.R1516X (c.CGA>TGA), estas son mutaciones sin sentido localizadas en el dominio NC1 y dan como resultado una proteína truncada (fig. 14).



**Figura 14.** Genealogía con SAAR. Representación de los dos hermanos afectados, debido a que los padres también afectados fueron heterocigotos para dicho cambio (Tomada de Cook *et al*, 2008).

### 3 Justificación

El empleo de información disponible de ciertas mutaciones y polimorfismos de manera confiada y detallada, genera la pauta para esclarecer la etiología y el comportamiento de algunas enfermedades de difícil diagnóstico. Tal es el caso del SA, patología estimada con una prevalencia general de 1:50,000 nacidos vivos, donde los genes *COL4A3* y *COL4A4* que involucran las formas SAAR y SAAD contraen aproximadamente el 20% de los casos; dichas formas presentan heterogeneidad fenotípica en sus trastornos como resultado proporcional de su heterogeneidad genotípica, esta última no suele ser constante y principalmente acumula apariciones de *novο*, las cuales generalmente suelen ser disímiles entre poblaciones, no así los polimorfismos asociados a dichos genes, ya que son comunes y generalmente están compartidos entre poblaciones, sin embargo no son excluyentes de la enfermedad, por lo tanto representa un problema que en poblaciones altamente polimórficas. Es por ello que mediante la población de estudio (pacientes y controles), se pretende tener una aproximación del perfil de las variantes alélicas que rigen a la población mexicana, sobre todo aquellas mutaciones que tienen en estado de SA a los niños mexicanos, así como de polimorfismos de sospechosa contribución de la enfermedad; ya que en México no existe la casuística que configura el estado de salud y enfermedad, en base a un análisis molecular enriquecido por un entorno referencial bioquímico, celular y fisiológico de la MBG. La razón de la obtención de este conocimiento establece en primera instancia su utilidad precautoria en el asesoramiento genético, y en segunda su idealización en la terapia génica como un modelo fidedigno de esperanza y calidad de vida para los pacientes afectados.

## 4 Hipótesis

Considerando la heterogeneidad de las mutaciones que inciden en la forma dominante y/o recesiva en niños mexicanos afectados con SA, se espera una baja correlación entre el fenotipo y el genotipo obtenido a partir de las secuencias del gen *COL4A3*.

## 5 Objetivos

### 5.1 General

Identificar y describir el perfil de las variaciones alélicas del gen *COL4A3* en pacientes diagnosticados con síndrome de Alport (SA) y controles clínicamente sanos.

### 5.2 Particulares

- Establecer agrupaciones de los patrones de similitud de los picos heteroduplex obtenidos mediante cromatografía líquida semi-desnaturalizante de alto rendimiento (DHPLC) para homogeneizar los polimorfismos de un solo nucleótido (SNPs) potencialmente presentes en el gen *COL4A3* en pacientes y controles, clarificando de este modo el grado polimórfico, así como su distinción en los estados de salud y enfermedad por SA.
- Relacionar el motivo de los picos heteroduplex visualizados en los cromatogramas del DHPLC con las variantes alélicas puntualizadas mediante secuenciación, y conjeturar mediante análisis molecular y bioquímico el comportamiento de las mutaciones y polimorfismos del gen *COL4A3* en la membrana basal glomerular (MBG), para deducir las implicaciones sobre la etiología de la enfermedad autosómica de SA.
- Apoyar, mediante herramientas bioinformáticas, la predicción del efecto fenotípico del análisis molecular y bioquímico de las variantes alélicas de pacientes y controles en la región codificante del gen *COL4A3*, valorando el

efecto fenotípico de la sustitución de aminoácidos sobre el funcionamiento de la cadena proteica  $\alpha 3(\text{IV})$  en la MBG.

- Deducir el posible comportamiento de las frecuencias genotípicas parcialmente obtenidas en la población de estudio y compararlas con el conjunto de poblaciones promedios establecidas alrededor del mundo, para obtener una aproximación de la tendencia de las variantes alélicas encontradas del gen *COL4A3* en la población mexicana.

## 6 Metodología

El presente estudio se llevó a cabo en el laboratorio de investigación en nefrología del Hospital Infantil de México Federico Gómez (HIMFG), como parte del protocolo HIM/2008/019 SSA. 805, titulado “Estudio descriptivo del espectro de mutaciones en colagena tipo IV en niños mexicanos con síndrome de Alport”, del cual es titular la Bióloga Ana María Hernández Sánchez.

### 6.1 Métodos

#### 6.1.1 Población de estudio

Fueron analizados 13 pacientes niños (promedio de 12 años de edad), diagnosticados en consulta externa de nefrología del HIMFG con SA bajo la premisa de dos de los siguientes criterios: i) hematuria familiar, ii) lenticono, iii) sordera neurosensorial y/o iv) alteraciones características de la MBG. Por otro lado, también se analizaron 50 individuos clínicamente sanos para establecer el grupo control, esto con el fin de considerar el número de alélos requeridos para concretar la tasa del 1% de SNPs en el estudio de variantes polimórficas (Inieta *et al*, 2005). En cada uno de los pacientes e individuos control, se obtuvo previa autorización para ser parte de esta investigación.

#### 6.1.2 Toma de muestra

En cada paciente e individuo control fueron tomados 6 ml de sangre periférica con vacutainer/anticoagulante (EDTA). En seguida, cada muestra extraída se almacenó a 4°C para su conservación y posterior utilización en la extracción del ADN.

#### 6.1.3 Extracción de ADN

La extracción de ADN se obtuvo a partir de 200 µl de muestra de sangre de cada paciente e individuo control por medio del kit comercial QIAamp DNA Mini spin de

QIAGEN. Este método se fundamenta por la interacción iónica y solubilidad de las distintas biomoléculas que conforman la muestra con agentes caotrópicos como factores desnaturalizantes y diferentes solventes orgánicos como purificadores, respectivamente, de tal forma que el ADN queda libre de proteínas, nucleasas y otros contaminantes o inhibidores. El procedimiento de la técnica consiste y sustenta los siguientes puntos:

1. Fueron agregados los 200  $\mu$ l de sangre total en un tubo con 20  $\mu$ l de proteasa QIAGEN para digerir las proteínas de los componentes intracelulares.
2. Posteriormente, fueron adicionados 200  $\mu$ l de buffer AL a la muestra -como solución hipotónica- para efectuar lisis celular; en seguida se incubó a 56 °C por 10 min (minutos) en el baño modular Barnstead Lab-line 2052, esto con el fin de aumentar la fuerza iónica del medio y así precipitar proteínas; después se centrifugo de manera breve.
3. Luego, se añadieron 200  $\mu$ l de etanol (96%) a la muestra, para interaccionar con las sales previamente añadidas y de este modo precipitar los ácidos nucleicos; dicha muestra posteriormente se mezcló dando vortex durante 15 seg (segundos); después se centrifugo brevemente.
4. La mezcla del paso anterior se transfirió a la columna membranosa de sílica QIAamp Mini spin, se centrifugó a 8,000 rpm (revoluciones por minuto) durante un min y posteriormente se colocó en un tubo colector para desechar el filtrado; esto con el fin de adsorber específicamente el ADN por medio de su tamaño e interacción iónica en dicha columna, permitiendo el paso de moléculas y sales que acompañaron la reacción de lisis celular.
5. Posteriormente, agregando 500  $\mu$ l de buffer AW1, centrifugando a 8,000 rpm por un min y filtrando el sobrante en un tubo colector; y en seguida el mismo procedimiento pero con buffer AW2 y 14,000 rpm durante 3 min; se efectuaron

los lavados del ADN de moléculas de solventes orgánicos remanentes de los pasos anteriores, obteniendo de esta forma mayor pureza de ADN.

6. Por último, el ADN fue absorbido de la columna con 200  $\mu\text{l}$  de buffer AE, y eluido en un nuevo tubo colector, después se centrifugo brevemente a 8,000 por un min. Posteriormente, el extracto de ADN fue almacenado a  $-20\text{ }^{\circ}\text{C}$  para su conservación.

#### 6.1.4 Cuantificación del ADN

La cuantificación del ADN se determinó por espectrofotometría en NanoDrop®, auxiliado con el programa del equipo (ND-100 V3.2.1). El NanoDrop fue calibrado con 2 $\mu\text{l}$  de agua Milli-Q (Millipore) para conseguir en primera instancia la muestra blanco, posteriormente con 2 $\mu\text{l}$  de cada muestra de ADN sin diluir, se obtuvo la lectura de su concentración (en ng/ $\mu\text{l}$ ) y pureza (intervalo de  $\lambda$  a 260/280 nm).

Después, cada muestra fue diluida para obtener una concentración final de 20 ng/ $\mu\text{l}$ , debido a que dicha concentración se utilizó como punto de partida para efectuar la técnica de PCR.

#### 6.1.5 Integridad del ADN

La verificación de la integridad de cada ADN extraído se realizó para conocer si la extracción de ADN no sufrió degradación, conjeturando de esta forma mayor posibilidad, en que las regiones de las secuencias diana dentro del gen *COL4A3* se encontrarán completas para el interés de la técnica PCR. Lo anterior se efectuó mediante electroforesis en gel de agarosa al 2%. El procedimiento es descrito más adelante en la sección 6.1.7.

### 6.1.6 Amplificación de los fragmentos del gen *COL43* mediante reacción en cadena de la polimerasa (PCR)

En todos los pacientes e individuos control, la amplificación del gen *COL4A3* fue parcial, debido a que solo se efectuó con 17 diferentes fragmentos, los cuales incluyeron: 8 de los 52 exones que componen a este gen y la región UTR 5', los demás en forma sub-dividida correspondieron a toda la región UTR 3' (Anexo 11). Cada fragmento fue flanqueado por dos pares de cebadores 5' y 3' (los cuales fueron diseñados en <http://www.frodo.wi.mit.edu/primer3/> y fabricados por Applied Biosystems by Life Technologies, 2009). Desde la perspectiva funcional del gen *COL4A3*, los fragmentos de este estudio incluyen: al exón del dominio 7S, que contiene el péptido señal, así como parte de su región adyacente UTR 5' (exón 1); algunos exones del dominio colagenoso, flanqueados por las correspondientes regiones intrónicas adyacentes (exones 2, 3, 7, 9 y 21-23); y el último exón del dominio NC1 junto a toda la región adyacente UTR 3' (exón 52 sub-dividido en 52A-I, debido a su gran longitud).

La siguiente lista de reactivos de Invitrogen, fueron utilizados para lograr las condiciones de reacción necesarias para la amplificación -mediante PCR- de los productos de los fragmentos en mención:

- ✓ Buffer para PCR (10X).- Compuesto de 200mM Tris-HCl (pH 8.4). Amortiguador salino implicado en la posesividad de la enzima Taq Platinum Polimerasa. Estabiliza la interacción de la hibridación cebador-ADN molde, neutralizando las cargas de los grupos fosfato.
- ✓ MgCl<sub>2</sub> (50mM).- Conformado por cationes divalentes que funcionan como cofactores de la Taq. Platinum Polimerasa, ayuda a catalizar la cadena creciente mediante interacciones de unión con la cadena molde, cebadores 5' y 3' y nucleótidos incorporados.

- ✓ dNTP's (10mM).- Es una mezcla de desoxinucleótidos trifosfato de adenina, guanina, citosina y timina (dATP, dGTP, dCTP y dTTP; respectivamente); forman parte de la cadena en síntesis.
- ✓ Polimerasa Taq Platinum (~80,000 U/mg).- Es una enzima recombinante termoestable modificada de la especie *Thermus aquaticus*. Contiene una subunidad que la mantiene inactiva a temperatura ambiente, sin embargo la recupera a 94°C. Cataliza la síntesis de ADN en dependencia con el ADN molde.
- ✓ Cebadores (100U/μL).- Son oligonucleótidos de ~20 pb que contienen las secuencias complementarias de los segmentos templados del gen *COL4A3*. Delimitan la posición 5' y 3'.
- ✓ ADN molde.- Corresponde al ADN genómico que contiene los *loci* de interés del gen *COL4A3*.

La mezcla de los reactivos descritos requirió de previa estandarización, (ajustando sobre todo la concentración del MgCl<sub>2</sub>). Para obtener dicha cantidad de producto específico de cada fragmento, se ajustaron las siguientes concentraciones de reacción:

**Cuadro 1.** Condiciones de PCR para la amplificación de los fragmentos del gen *COL4A3*.

Reactivos	Concentración inicial	Concentración final	Volumen final a 25 $\mu$ l
Buffer PCR	10 X	1 X	2.5 $\mu$ l
MgCl <sub>2</sub>	50 mM	2.5 mM (1.5 mM *)	1.25 $\mu$ L (0.75 $\mu$ l*)
dNTP's	10 mM	0.2 mM	0.5 $\mu$ l
Taq. Platinum polimerasa	5 u/ $\mu$ l	1 U/ $\mu$ l	0.2 $\mu$ l
Cebador Sentido 5'	10 pmol/ $\mu$ l	4 pmol/ $\mu$ l	0.4 $\mu$ l
Cebador Antisentido 3'	10 pmol/ $\mu$ l	4 pmol/ $\mu$ l	0.4 $\mu$ l
ADN	-	20 ng/ $\mu$ l	1 $\mu$ l
H <sub>2</sub> O	-	-	Aforar a 25 $\mu$ l

\* Condición única para el exón 1

Las reacciones se llevaron a cabo en el termociclador Techne (TC-412), aplicando 30 y 35 ciclos de acuerdo al tamaño del producto, es decir <500 pb y >500 pb, respectivamente; esto para cada uno de los fragmentos. La condición de reacción fue determinada por la procesividad de la Taq. Platinum Polimerasa (1kb/min) y el punto de fusión o temperatura semi-desnaturalizante de los cebadores ( $T_m$ ) (Anexo 10). Este último parámetro se concreto en base al %G/C de los cebadores sentido y antisentido, así como de las secuencias de ADN molde (Sambrook and Russell, 2001).

La temperatura óptima de ciclado en el termociclador para las fases de desnaturalización, alineamiento y extensión, se mantuvieron durante 40 segundos. Dichas condiciones determinaron la longitud del producto:

**Cuadro 2.** Fases de la PCR en el termociclador para cada uno de los fragmentos.

Fase	Temperatura (°C)	Tiempo
Desnaturalización inicial	96	4 min
Desnaturalización	96	40 seg
Alineamiento	T <sub>m</sub> *	40 seg
Extensión	72	40 seg
Extensión final	72	40 seg
Mantenimiento	4	∞

\*La temperatura de alineamiento se obtuvo del promedio de la T<sub>m</sub> del cebador sentido 5' y antisentido 3' de cada fragmento.

Una vez concluido el ciclado de todos los fragmentos del gen *COL4A3* en la población de estudio, se efectuó la técnica de electroforesis para verificar el tamaño e integridad de cada amplicón.

### 6.1.7 Electroforesis en gel de agarosa

La comprobación de la integridad y tamaño del producto deseado a partir de cada fragmento amplificado en pacientes e individuos control, fue determinado mediante electroforesis en gel de agarosa -de Omega- al 2%.

El corrimiento electroforético se efectuó a partir de 5 µl de cada amplicón (aproximadamente 1.5 µg) con 1 µl de buffer de carga (Fermentas), además de un control negativo (sin ADN en la mezcla de reacción para verificar que no hubiera contaminación por ADN exógeno), y una marcador de peso molecular (Fermentas) de 50 y/o 100 pb (la utilización de uno u otro dependió del tamaño del amplicón esperado). El procedimiento fue el siguiente:

1. Se disolvieron 2 g de agarosa en 100 ml de buffer Tris-boratos-EDTA (TBE) 1X.
2. En seguida, se agregó 1  $\mu$ l de bromuro de etidio (0.5 mg) a la mezcla anterior y posteriormente se vertió dicha mezcla en una placa molde electroforética con peines de 1.5 mm. La solución dejó reposarse hasta la formación del polímero.
3. El gel resultante se sumergió en una cámara electroforética con contenido TBE 1X. Cada una de las muestras fue cargada en un respectivo pozo, mezclando los 5  $\mu$ l de ADN con 1  $\mu$ l de buffer de carga, así como un marcador de peso molecular y un control negativo.
4. La migración del ADN se efectuó aplicando 100 Volts durante 50 min con una fuente de poder eléctrica de Apparatus Corporation EC 105, para que el ADN migrara hacia el ánodo.
5. Una vez concluido el tiempo de migración, el gel fue extraído de la cámara electroforética y fue revelado con luz UV en el foto documentador Kodak. Gel Logic 1500. Finalmente la fotografía de las bandas de ADN fueron visualizadas y analizadas con el programa de ese equipo.

#### 6.1.8 Secuenciación automatizada mediante el principio de terminación de la cadena “didesoxi”

Antes de identificar las variantes alélicas de la población de estudio, en primera instancia se obtuvo una secuencia de cada amplicón, de tal forma que se tomó de manera aleatoria uno de los 50 controles por fragmento, en primera lugar para estandarizar dicha técnica y en segundo para utilizar estos mismos controles como moldes en la subsecuente técnica de hibridación DHPLC (ver sección 6.1.9). De este modo logró conocer y conjeturarse el estado de las secuencias control-molde, facilitando de esta forma el alineamiento y comparación entre secuencias con resultados posteriores

de electroferogramas, y por lo tanto la resolución de los SNPs presentes en la población de estudio. Para confirmar la validez de las secuencias de los controles-molde, se realizaron alineamientos comparativos con las secuencias curadas del gen *COL4A3* en el sitio web *Ensembl* (<http://www.ensembl.org/index.html>, 2011).

Las secuencias fueron preparadas a partir de los productos de PCR. Primero, se utilizó la exonucleasa ExoSAP-IT para que los remanentes de los reactivos suministrados en las reacciones de PCR fueran degradados y no interfirieran con las reacciones de secuencia.

Las secuencias fueron preparadas a partir de 5 µl de cada amplicón (1.5 µg de ADN) y 2 µl de ExoSAP-IT, de acuerdo al protocolo de este último, el cual puede ser consultado en línea (<http://www.iwai-chem.co.jp/products/usb/exosap-it>, 2009). El procedimiento contempla las siguientes condiciones de incubación en el termociclador:

1. A 37°C por 15 min. Temperatura de activación enzimática.
2. A 80°C por 20 min. Temperatura de inactivación enzimática.

Posteriormente fueron cuantificados cada uno de los productos de PCR-ExoSAP-IT en NanoDrop, para conocer y ajustar su concentración mediante diluciones, concretando cada muestra en 40 ng/µl, dicha concentración fue utilizada para llevar a cabo las reacciones de secuencia de todos los fragmentos. Las reacciones óptimas de secuenciación fueron determinadas a partir de las especificaciones del protocolo del manual de BigDye Terminator v3.1 (Cycle Sequencing Kit, 2002).

Las preparaciones de las reacciones de secuencia incluyeron la mezcla de los siguientes de reactivos:

**Cuadro 3.** Reactivos utilizados para las reacciones de secuenciación.

Reactivos	Volumen ( $\mu$ l)
BigDye Terminator <sup>1</sup>	4 $\mu$ l
Buffer (5X)	2 $\mu$ L (1X)
Cebadores (5' ó 3') <sup>2</sup> (2 pmoles)	1 $\mu$ l
ADN (40 ng/ $\mu$ l)	1 $\mu$ l
H <sub>2</sub> O	Aforar a 10 $\mu$ l

<sup>1</sup> Kit comercial de de Applied Biosystems by Life Technologies; el cual incluye dNTP's, ddNTP's marcados fluorescentemente, DNA polimerasa y MgCl<sub>2</sub>.

<sup>2</sup> El criterio de elección del cebador se basó en la lejanía de las bases del exón diana a secuenciar, ya que la lectura inicial de la secuencia es de difícil determinación.

Las reacciones se llevaron a cabo en el termociclador mediante copiado de las secuencias de interés (no es una amplificación de la secuencia como en el caso de la PCR, de allí que la concentración del ADN se llevara a 40 ng/ $\mu$ l). Tomando en cuenta la T<sub>m</sub> de cada uno de los cebadores, se establecieron las siguientes condiciones de programación en el termociclador a 50 ciclos (también conforme al manual BigDye Terminator v3.1):

**Cuadro 4.** Condiciones de secuenciación en el termociclador.

Fase	Temperatura (°C)	Tiempo
Desnaturalización inicial	95	2 min
Desnaturalización	95	15 seg
Alineamiento	T <sub>m</sub> del cebador 5' ó 3'	10 seg
Extensión	60	10 min
Extensión final	4	5 min
Mantenimiento	4	∞

Una vez concluidos los ciclos de secuenciación, el producto de cada reacción fue purificado con columnas de sílica gel CENTRI-SEP de Applied Biosystem, las cuales remueven el ADN monocatenario fuera de la columna, dejando dentro de la misma los reactivos que no fueron incorporados en la reacción de secuencia, principalmente los marcadores terminadores fluorescentes (ddNTP's). Posteriormente el ADN monocatenario fue estabilizado con formamida para que no interaccionará consigo mismo ni con las otras moléculas de ADN, anulando de esta forma cualquier interacción no covalente. El procedimiento de purificación y estabilización de las reacciones de secuenciación se llevaron a cabo de la siguiente forma:

1. La columna CENTRI-SEP fue hidratada con 0.8 ml de agua Milli-Q. Posteriormente se dejó reposarse al menos 30 min.
2. Posteriormente, fue colocado un tubo colector de 1.5 ml en la columna y enseguida se transfirió la reacción de secuencia, después se centrifugó a 3,000 rpm por 2 min.

3. El producto obtenido en el tubo colector fue conservado y colocado en un desecador al vacío (Thermo Electron Corporation ISS110-115) durante 15 min, para evaporar el agua incorporada en la reacción.
4. Una vez seco el ADN, se resuspendió en 12 µl de formamida. Posteriormente la mezcla fue homogeneizada dando vortex brevemente.
5. Por último, la muestra fue colocada en un baño seco modular a 96°C durante 2-3 min y posteriormente fue almacenada a 4°C.

Las reacciones de secuencia se corrieron en el secuenciador ABI Prism 310® (Applied Biosystem). Este equipo consiste en un sistema de electroforesis capilar que utiliza el polímero acrilamida para correr la muestra de secuencia. El ingreso de las reacciones al capilar se llevó a cabo mediante inyección electrocinética, desde el ánodo hacia el cátodo, el capilar y el cátodo fueron inmersos en la muestra, estableciendo de esta forma un corto período de electroforesis. Posteriormente, el extremo del capilar cercano al cátodo fue sumergido en solución buffer y en seguida fue aplicada una elevada diferencia de voltaje, ocasionando que los distintos fragmentos de ADN monocatenario se movieran a lo largo del capilar a diferentes tasas de migración (Fuentes *et al*, 1998 y Sambrock and Rusell, 2001), hasta alcanzar una ventana que contiene dicho capilar cerca del cátodo para que el láser excite los distintos colores fluorescentes de los ddNTPs. La fluorescencia emitida por los colorantes fue colectada por un detector láser a longitudes de onda particulares y se registraron como señales digitales, las cuales fueron colectadas por una cámara CCD®. Mediante soporte informático con el programa Genetic Analyzer del equipo, lograron interpretarse los resultados, debido a la asignación de las bases para cada intensidad de fluorescencia en forma de picos. La lectura generó datos crudos, partiendo de las secuencias más cortas hasta la secuencia completa. Al final, el

equipo creó gráficos con picos de emisión de luz en la posición de cada uno de los nucleótidos de la secuencia de interés (electroferograma).

Finalmente, los resultados fueron analizados con los programas Finch TV 1.4.0 y DNAMAN 7.0; el primero permitió analizar los picos de los electroferogramas y el segundo alinear los segmentos de interés para establecer la igualdad o desigualdad entre comparación de secuencias.

#### 6.1.9 Detección de variantes alélicas mediante cromatografía líquida desnaturalizante de alto rendimiento (DHPLC)

Para identificar los posibles cambios en las secuencias génicas tanto en pacientes como en individuos control, fue empleada la técnica de cromatografía líquida desnaturalizante de alto rendimiento (DHPLC) de Transgenomic. Dicha técnica únicamente hace evidente la existencia de cambios en los fragmentos sin mostrar la naturaleza y posición de los mismos.

Para la utilización de este equipo debió seguirse una metodología previa, la cual consistió en la hibridación de las secuencias control-molde generadas anteriormente (ver 6.1.7 Secuenciación) con las de los pacientes y las de los individuos control, mediante la realización de mezclas 1:1 (300 ng de control-molde y 300 ng de cada paciente ó individuo control) en un volumen final de aproximadamente 8 µl. Esta mezcla se procesó en el termociclador bajo las condiciones mostradas en el cuadro 5, posterior a lo cual se pudieron introducir al DHPLC.

**Cuadro 5.** Condiciones de hibridación de los amplicones en el termociclador.

<b>Fase</b>	<b>Temperatura (°C)</b>	<b>Tiempo</b>
Desnaturalización inicial	96	4 min
Desnaturalización	96	40 seg
Realineamiento gradual	Inicio de 95, disminuye cada 5, hasta llegar a 5	30 seg
Extensión final	4	40 seg
Mantenimiento	4	∞

El fundamento se basa en la detección de secuencias desalineadas o heteroduplex y secuencias correctamente alineadas u homoduplex resultadas de la elución por diferencia de estabilidad (Oefner and Underhill, 1998).

Los controles-molde fueron inyectados y conducidos hasta la columna de sílica en el horno del equipo, donde se sometieron a una estandarización para encontrar la temperatura óptima semi-desnaturalizante ( $T_m$ ) (cuadro 6), a partir de parámetros que fueron obtenidos con ayuda del programa Navigator del DHPLC en base al %GC de una fracción de la secuencia.

**Cuadro 6.** Características cromatográficas óptimas de cada fragmento

<b>Fragmento</b>	<b>Temperatura del horno (°C).</b>	<b>Perfil de elución del buffer B (%)</b>	<b>Tiempo de retención (min)</b>
UTR 5'-Exón 1	64.5	59.1-68.1	3.71
Exón 2	57.1	52.1-61.1	4.49
Exón 3	60.2	52.3-61.3	3.19
Exón 7	55	52.1-61.1	2.99
Exón 9	55.4	52.2-61.2	4.3
Exón 21	52.9	56.8-65.8	3.77
Exón 22	54.6	55.8-64.8	4
Exón 23	56.7	52.9-61.9	3.82
Exón 52A-UTR 3'	59.2	59.2-68.2	4.01
52B-UTR 3'	57.5	58.3-67.3	3.37
52C-UTR 3'	59.2	59.2-68.2	3.71
52D-UTR 3'	53.1	59.5-68.5	3.65
52E-UTR 3'	53.8	59-68	3.61
52F-UTR 3'	54.3	59.1-68.1	3.96
52G-UTR 3'	55.4	58.6-67.6	4.38
52H-UTR 3'	55.7	58.5-67.5	4.19
52I-UTR 3'	52.7	58.6-67.6	4.2

Esta temperatura es importante pues permite la separación parcial de las cadenas, lo que más adelante facilita la unión a la fase estacionaria de la cromatografía. El proceso inicia con la inyección de los buffers A y B, 0.1 M de acetato de trietilamonio (ATEA) y acetato de trietilamonio (ATEA) más 25% de acetonitrilo (AN) respectivamente, en la columna de estireno-divinilbenceno (DNAsep column de Transgenomic) a una proporción inicial de 50:50% y una tasa de flujo de 0.9 ml/min. Dicha solución alcanza la proporción 75:25 (buffer A: buffer B) para permitir la unión de la muestra con el ATEA. El ATEA es una molécula anfifílica y su parte negativa interacciona con la columna, mientras que la positiva lo hace con el esqueleto fosfatado del ADN, es por ello que la muestra debe estar semi-desnaturalizada, para permitir una mejor unión con esta parte del ATEA. Luego de la inyección de los buffers, la muestra fue introducida, a partir de lo cual el buffer B fue incrementado gradualmente en un perfil de elución de aproximadamente un 2% por minuto (cuadro 6), lo que finalmente hace que fluya la muestra (Xiao and Oefner, 2001).

Las cadenas que presentan cambios, denominadas heteroduplex, son menos estables puesto que tienen bases mal apareadas, por lo que fluyen primero y son detectadas por el sensor UV del equipo antes que las cadenas homoduplex o sin cambios y que tienen mayor estabilidad, por lo que fluyen después (Schollen *et al*, 2005).

El resultado, obtenido en tiempo real es un cromatograma, el cual representa una gráfica que relaciona la ABS contra el tiempo de retención de ADN en la columna; en ésta se observan los hetero y homoduplex en forma de picos.

La sensibilidad y confiabilidad del DHPLC para la identificación de cualquier variante es hasta de 1 pb (Oefner and Underhill, 1998); de tal forma que las muestras detectadas con un patrón de picos anormales en los fragmentos de los pacientes que contienen los exones que expresan la cadena  $\alpha 3(IV)$ , es decir, desde el exón 1 hasta el exón 52A de los pacientes, fueron señaladas como candidatas para secuenciación;

mientras que los demás fragmentos de los pacientes, es decir, los que abarcan toda la región UTR 3' (52B-52I), así como de todos los 17 fragmentos para los controles, fueron reserva para otros proyectos. La finalidad de secuenciar estos cambios detectados es para detectar de manera puntual la localización y la naturaleza, así como para descifrar las implicaciones mecánicas, físicas y químicas de dicha variante.

Los resultados de los cromatogramas de los picos homoduplex fueron analizados con la herramienta Mutation Calling (programa Wave DNA Fragment Analysis System del equipo); por medio de representaciones comparativas que normalizan los datos para reducirlos y representarlos en tres dimensiones (3D). Mutation Calling permite utilizar un valor de significancia específico denominado GAP (nivel de confianza GAP= <0.01), el cual exhibe un árbol de afinidad que desplaza los valores cromatográficos normalizados en una gráfica de dispersión de componentes principales (Colley *et al*, 2005).

## 6.1.10 Pruebas estadísticas

### 6.1.10.1 Predicción de la substitución de aminoácidos

El impacto potencial en el fenotipo de cierto SNP que ocasiona la substitución de un aa, fue evaluado mediante soporte informático con dos herramientas analíticas: SIFT ([http://www.sift.jcvi.org/www/SIFT\\_chr\\_coords\\_submit.html](http://www.sift.jcvi.org/www/SIFT_chr_coords_submit.html)) y Polyphen (<http://www.genetics.bwh.harvard.edu/pph/>).

#### 6.1.10.1.1 SIFT

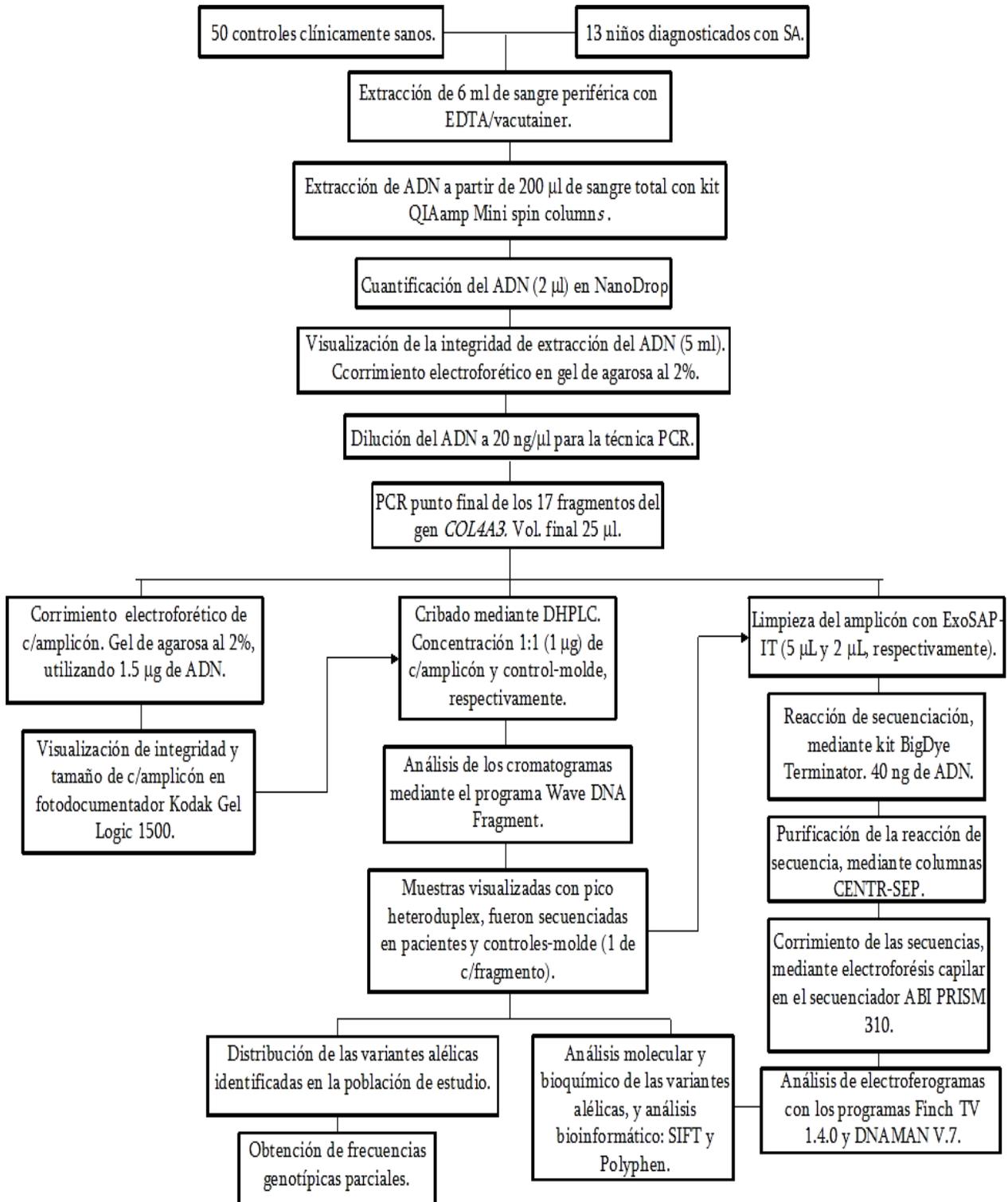
SIFT está basado en la homología que presenta la secuencia tolerante de un aa común de aquella que no lo es, y predice si la substitución de un aa en una proteína dada tendrá un impacto en el fenotipo. La premisa de la que parte esta herramienta considera que la evolución de la proteína esta correlacionada con su función. Importantes posiciones

para la función tienen que estar conservadas en el alineamiento de la familia de la proteína, por el contrario serán diversos en el alineamiento. SIF calcula una cifra estadísticamente significativa en la sustitución del aa, y un puntaje menor de 0.05 es considerado potencialmente dañino.

#### 6.1.10.1.2 PolyPhen

Por otro lado, Polyphen predice el posible impacto de sustitución de un aa en la estructura y función de una proteína humana. Esta predicción está basada en reglas empíricas comunes, las cuales son aplicadas a la secuencia, filogenia e información estructural que caracteriza la sustitución. Polyphen utiliza un programa denominado: cuenta independiente de posición específica, para calcular el perfil de puntuación obtenida desde la probabilidad de que un determinado aa ocupe la misma posición de interés comparada a la secuencia referente.

## 6.2 Diseño metodológico



## 7 Resultados y discusión

Fueron detectadas las variantes alélicas del *locus COL4A3* del COL(IV) de 13 pacientes diagnosticados con SA y 50 controles clínicamente sanos mediante DHPLC; las muestras que presentaron picos heteroduplex fueron secuenciadas para obtener de manera puntual la localización y naturaleza de los polimorfismos y mutaciones. Los resultados que se presentan a continuación van del análisis de la calidad de extracción del ADN hasta la evaluación detallada de cada uno de los 17 fragmentos considerados en el presente estudio.

### 7.1 ADN genómico de la población de estudio

#### 7.1.1 Cuantificación y pureza de la extracción de ADN

En el cuadro 7 se muestra la concentración y pureza del ADN de pacientes y controles (población de estudio). Este material genético fue el punto de partida para ayudar a definir la apreciación de los subsecuentes análisis moleculares, por lo tanto fue importante obtener extracciones de ADN óptimas. La valoración de la pureza ( $\lambda$  260/280 nm) se mantuvo dentro del rango óptimo en todos los casos (1.7-2.0), por lo que el ADN fue considerado sin contaminación por otras moléculas que pudiesen afectar sobre todo la reacción de amplificación por PCR.

**Cuadro 7.** Cuantificación y pureza de ADN de cada muestra.

Paciente	[ng/ $\mu$ l]	$\lambda$ 260/280
1	169.5	1.82
2	69.24	1.94
3	69.60	1.71
4	97.53	1.88
5	216.91	1.85
6	56.53	1.69
7	40.30	1.93
8	100.33	1.91
9	93.70	1.75
10	24.13	1.90

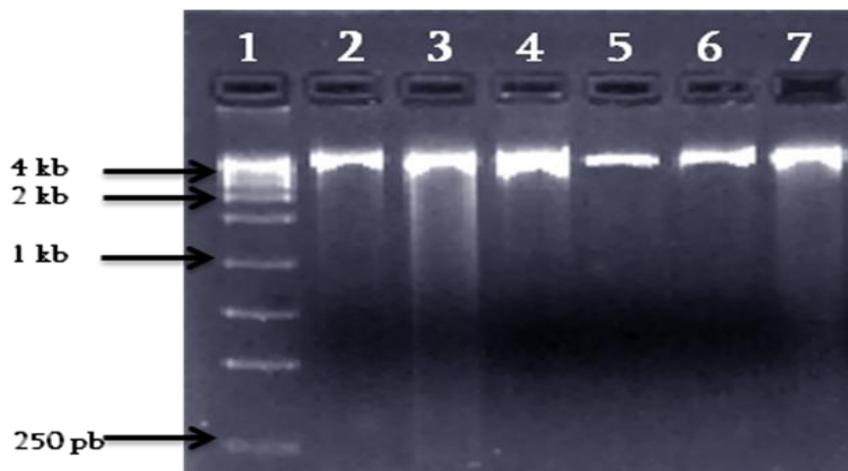
11	67.17	1.95
12	43.81	1.88
13	68.25	1.71

Control	[ng/μl]	λ 260/280			
1	71.78	1.80	25	55.63	1.88
2	87.67	1.93	26	85.72	1.84
3	100.36	1.83	27	51.86	1.79
4	97.53	1.87	28	57.18	1.90
5	272.22	1.84	29	53.45	1.75
6	35.72	1.80	30	21.00	1.56
7	216.91	1.88	31	41.47	1.75
8	107.60	1.89	32	69.02	1.90
9	314.89	1.87	33	76.20	1.88
10	69.51	1.90	34	14.08	1.81
11	56.53	1.87	35	38.20	1.98
12	69.89	1.88	36	108.5	1.73
13	106.18	1.84	37	87.60	1.91
14	40.30	2.11	38	80.50	1.87
15	93.70	1.84	39	140.20	1.86
16	122.84	1.93	40	121.91	1.90
17	45.70	1.70	41	146.22	2.04
18	77.37	1.85	42	143.81	1.92
19	24.13	1.74	43	95.32	2.01
20	67.17	1.69	44	167.73	1.94
21	56.83	1.90	45	133.91	1.91
22	51.64	1.78	46	148.92	1.94
23	48.71	1.69	47	93.22	1.90
24	46.81	1.87	48	99.28	1.87
			49	98.82	1.86
			50	102.16	1.77

### 7.1.2 Integridad de la extracción de ADN

En la figura 15 se visualiza la integridad de las extracciones de ADN de la población de estudio mediante movilidad electroforética de tres pacientes (muestras 2-4) y tres controles (muestras 5-7), en representación de la población de estudio. La nula degradación del ADN puede apreciarse por la integración de bandas definidas con similitud de pesos moleculares, pero diferentes concentraciones, las cuales fueron esperadas de acuerdo al resultado espectrofotométrico. La importancia de evaluar esta cuestión, radicó principalmente para establecer que los cebadores 5' y 3' que delimitan

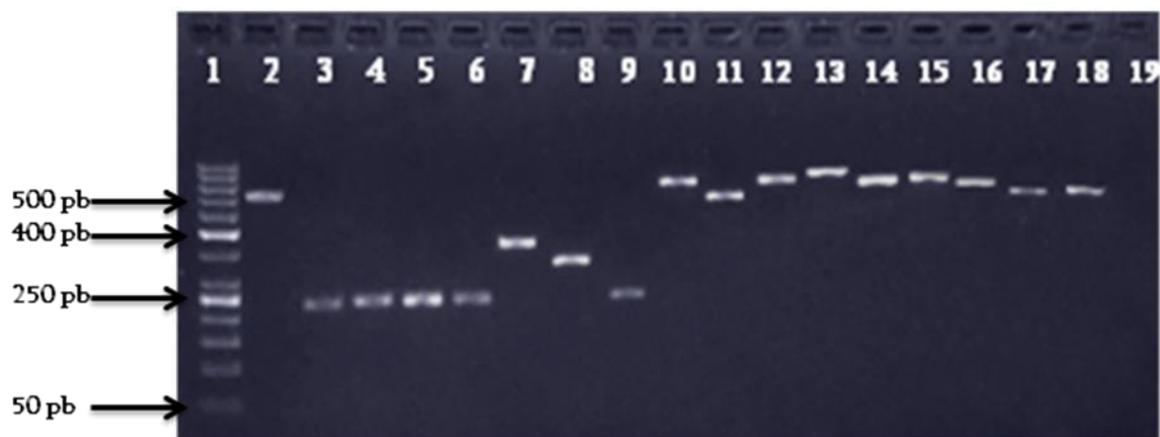
los 17 fragmentos evaluados, encontraran e hibridaran con sus secuencias complementarias dentro del genoma, y de este modo se generara el entorno para una reacción eficaz de la técnica PCR, es decir que la polimerasa extienda los cebadores y catalice eficazmente las regiones de interés.



**Fig. 15.** Fotografía digital tomada con luz UV que muestra la integridad del ADN extraído. **1:**marcador de peso molecular de 1 kb, **2:**paciente 2, **3:**paciente 1, **4:**paciente 5; **5:**control 11, **6:**control 20, **7:**control 44. Se utilizaron 5  $\mu$ l de ADN en el corrimiento electroforético en gel de agarosa al 1% (110 Volts durante 180 min).

## 7.2 Validación de los 17 fragmentos amplificados mediante PCR

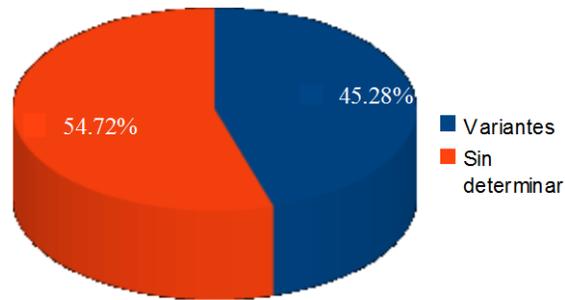
En la figura 16 se observa una muestra de cada uno de los 17 fragmentos del gen *COL4A3*; cada una corresponde a un paciente elegido al azar. Cada fragmento fue amplificado satisfactoriamente, debido a que la prueba electroforética reveló integridad de bandas, además el tamaño del producto esperado coincidió con la movilidad electroforética de la escalera referencial (50 pb); dichas observaciones descartaron la posibilidad de productos inespecíficos. La pauta marcada por las bandas de los amplicones, establecieron las especificaciones requeridas para la incorporación de este ADN en las metodologías de secuenciación y DHPLC.



**Figura 16.** Fotografía digital tomada con luz UV que verifica integridad y tamaño de cada fragmento amplificado por PCR. 1:marcador de peso molecular de 50 pb. 2:UTR 5'-exón 1. 3:exón 2. 4:exón 3. 5:exón 7. 6:exón 9. 7:exón 21. 8:exón 22. 9:exón 23. 10:exón52A-UTR 3'. 11:52B UTR 3'. 12:52C UTR 3'. 13:52D UTR 3'. 14:52E UTR 3'. 15:52F UTR 3'. 16:52G UTR 3'. 17:52H UTR 3'. 18:52I UTR 3'. 19:control negativo (reacción llevada a cabo sin ADN). Se utilizaron 5  $\mu$ l de amplicón (aproximadamente 350 ng/ $\mu$ l) en el corrimiento electroforético en gel de agarosa al 2% (100 Volts durante 60 min).

### 7.3 Detección de las variantes alélicas mediante cribado por DHPLC

Al correr los 17 fragmentos del gen *COL4A3* en la población de estudio, los resultados del DHPLC arrojaron 1071 cromatogramas. Dentro de cada fragmento se detectaron muestras con picos heteroduplex definidos, excepto en los fragmentos UTR 5'-exón 1 y exón 23, debido a que fueron observados con picos homoduplex en la totalidad de la población de estudio; cabe señalar que algunos controles mostraron ligera discrepancia en sus picos homoduplex, sin embargo se les consideró con secuencias íntegras, es decir, sin portación de variante alélica alguna. En general, lograron identificarse 640 cromatogramas con picos heteroduplex, de los cuales 485 han sido depurados hasta la determinación de su(s) variante alélica (fig. 17). Finalmente se estima que el 60% de la población de estudio contiene al menos alguna variante alélica en cualquiera de los fragmentos evaluados, lo cual pone de manifiesto que el gen *COL4A3* es una región altamente asociada a SNPs (Longo *et al*, 2002 y Wang *et al*, 2004).



**Figura 17.** Distribución de los cambios detectados hasta el momento en el gen *COL4A3* de la población de estudio.

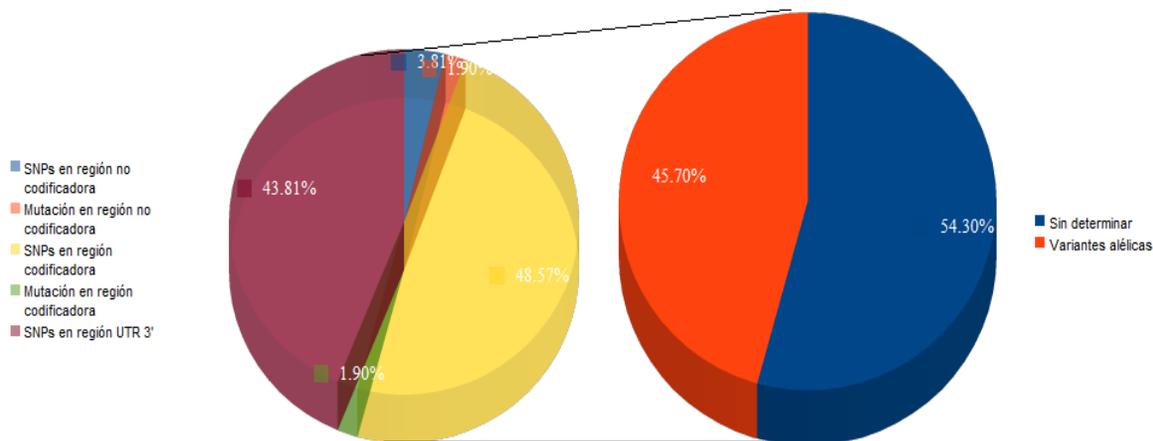
En los cromatogramas analizados en este estudio, lograron conjuntarse 20 patrones de picos heteroduplex que correspondieron a 20 variables alélicas (de acuerdo a lo visualizado en los electroferogramas); esto indica que la tasa de efectividad del DHPLC para la detección de cambios en los fragmento fue del 100% (Schollen *et al*, 2005).

### 7.3.1 Distribución de las variantes alélicas en la población de estudio

#### 7.3.1.1 Pacientes

En la figura 18 se representan los 221 resultados cromatográficos de los pacientes; 101 de ellos (es decir, el 45.7%) aparecieron con picos heteroduplex, los cuales en conjunto implican 19 variantes alélicas; los 120 cromatogramas restantes (54.3%), incluyen tanto a picos homoduplex como aquellos heteroduplex que restan por definir su variante causal. Para los casos visualizados con picos homoduplex, se espera que la etiología de SA recaiga en otros exónes de *COL4A3* o en los exónes que conforman los genes *COL4A4* y/o *COL4A5*. Respecto a la gráfica que muestra la sustracción del 45.7% de los picos heteroduplex, pueden distinguirse cinco distribuciones: por un lado 49 cromatogramas con SNPs en la región exónica (48.57%), 46 en la región UTR 3' (43.81%) y 4 en la región intrónica (3.81%); y por otro, 2 mutaciones idénticas en 2 pacientes: una en la región intrónica y la otra consecutiva en la exónica (1.9%); dichas

mutaciones son registros de *novo*, y fueron consideradas de tal forma debido en parte a que no fueron encontradas en controles.



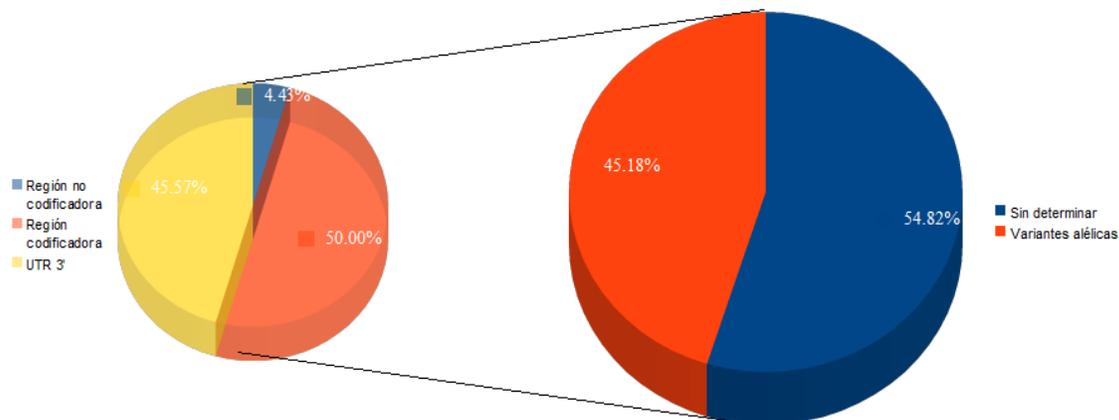
**Figura 18.** Representación de la distribución de los cambios detectados mediante DHPLC en el gen *COL4A3* de los pacientes. La gráfica de la derecha representa el total de pacientes y la de la izquierda el desglose de los resultados cromatográficos heteroduplex y su distribución en distintos segmentos con función definida a lo largo de dicho gen.

Las mutaciones en mención encontradas en 2 pacientes, representa el 15.38% de los casos, por lo tanto se aproxima con la estadística de incidencia para las formas SAAR y SAAD, que en conjunto establecen el 20% de los casos (Thorner, 2007 y Medeiros *et al*, 2008). Para conocer cual de las dos formas de SA es predominante, es necesario consultar el historial familiar (Kashtan, 2010).

### 7.3.1.2 Controles

En la gráfica 19 se representan los 850 resultados cromatográficos de los controles, 384 de ellos (45.18%) fueron visualizados con picos heteroduplex y en conjunto integran 18 variantes alélicas; el resto de los cromatogramas que constituyen el 54.82% son también parte de picos homo y heteroduplex con variantes por determinar. La figura también conjetura la distribución de los subcomponentes de los picos heteroduplex, donde pueden distinguirse tres regiones: 192 cromatogramas con SNPs en la región exónica (50%), 17 en la región intrónica (4.43%) y 175 en la región UTR 3' (45.57%); cabe

mencionar que en esta última región se obtuvieron tres variantes alélicas de *nov*, una de ellas faltante en los pacientes.



**Figura 19.** Representación de la distribución de los cambios detectados mediante DHPLC en el gen *COL4A3* de los controles. La gráfica de la derecha representa el total de pacientes y la de la izquierda el desglose de los resultados cromatográficos heteroduplex y su distribución en distintos segmentos con función definida a lo largo de dicho gen.

Los resultados cromatográficos muestran que la distribución de las variantes alélicas en controles es ligeramente proporcional a la de pacientes, lo cual establece homogeneidad entre la mayoría de dichas variantes en la población mexicana y compatibilidad con otras poblaciones evaluadas alrededor del mundo (Heidet *et al*, 2001; Badenas *et al*, 2002; Stabuc *et al*, 2009).

Los numerosos polimorfismos estimados en el gen *COL4A3* pueden comprometer en conjunto la estructura y en consecuencia la eficiente función de la cadena  $\alpha 3(\text{IV})$  codificante; incluso en los considerados controles clínicamente sanos, sin embargo resulta difícil notar esta cuestión a corto plazo, ya que los deterioros suelen ser lentamente progresivos, sobre todo tomando en cuenta que la cadena  $\alpha 3(\text{IV})$  es parte de un componente de numerosas biomoléculas que participan para formar la MBG (Rana *et al*, 2006). En el caso de los pacientes, la cantidad y tipo de polimorfismos adicionales a las 2 mutaciones detectadas, resaltan de manera más o menos obvia la disfunción de la

estructura glomerular, progresando incluso hasta IRCT (Tryggvason and Patrakka, 2006).

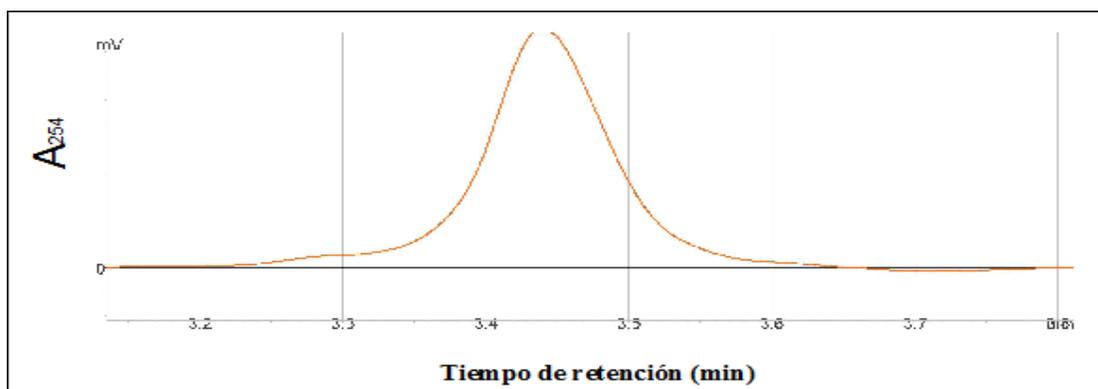
#### 7.4 Análisis molecular y bioquímico

El análisis molecular y bioquímico detalla la naturaleza y posición de las variantes alélicas detectadas y secuenciadas que fueron encontradas en el perfil de la población de estudio. Es importante resaltar que corresponde al primer informe referente al *locus COL4A3* para la población mexicana. Dicho análisis establece el comportamiento polimorfo y mutacional enfocado a la fisiología de la MBG, resaltando las modificaciones que surgen en su entorno, diferenciando sobre todo el estado de salud y enfermedad por SA.

##### 7.4.1 Fragmento UTR 5'-Exón 1

En los cromatogramas de la población de estudio del fragmento UTR 5'-exón 1 no fueron encontrados picos heteroduplex que definieran de manera certera -como lo indica el fundamento del DHPLC- alguna variante alélica; solamente 5 controles discreparon en la forma de su pico homoduplex; es importante aclarar que la secuenciación de estos 5 controles no fue parte de los objetivos de este estudio.

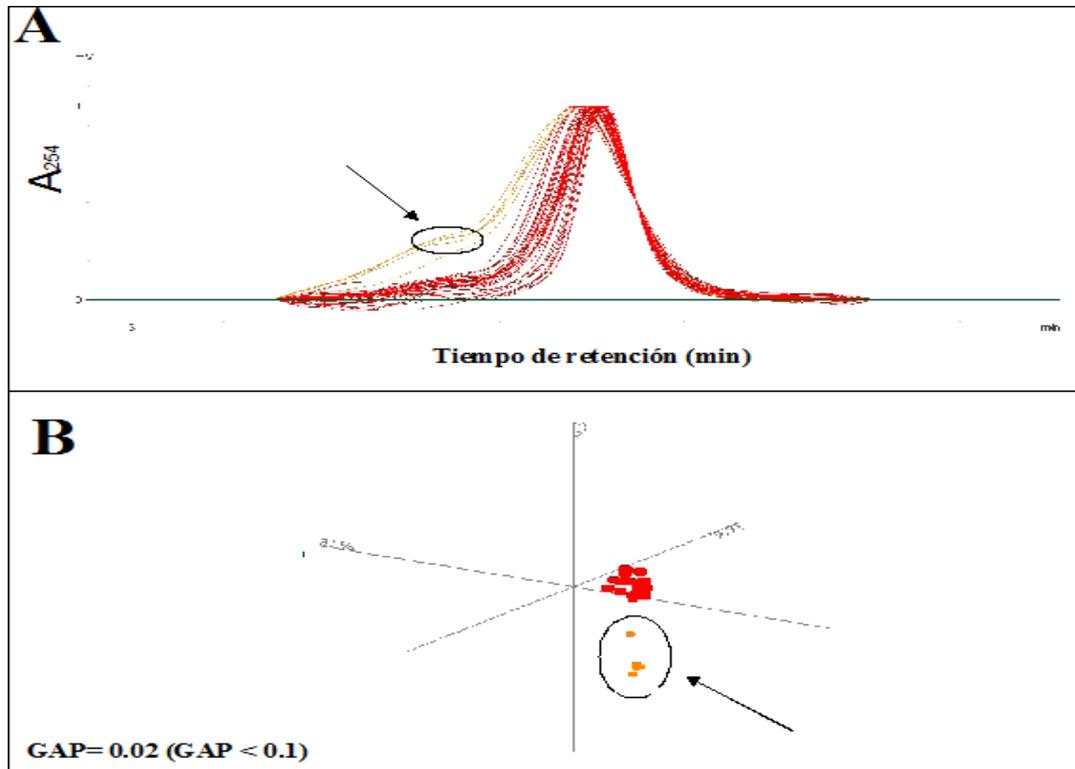
El control-molde utilizado en la hibridación fue confirmado como silvestre por el resultado que arrojó su secuencia (no se muestra), por lo tanto todos los demás picos homoduplex visualizados en los cromatogramas también lo fueron. A partir de la  $T_m$  de 64.5 °C predicha por el DHPLC (cabe notar que es la  $T_m$  más alta de entre todos los fragmentos debido a que posee una secuencia densa de islas CpG), el tiempo de retención promedio de los picos homoduplex fue de 3.4 min (fig. 20), indicando que la tasa de flujo del buffer B para el homoduplex estable (es decir 100% de bases apareadas correctamente) alcanzó un máximo de 65.2% de elución sobre los híbridos de ADN en la columna.



**Figura 20.** Cromatograma con pico homoduplex que de manera similar fue hallado en 45 controles y todos los pacientes.

Los 5 controles disimiles solamente pudieron apreciarse a partir de que la varianza de los cuadrados la población de estudio fue normalizada (fig. 21A). En el cromatograma se observa que los 5 controles no mostraron diferencias de elución ó tiempo de retención en la columna, además sus concentraciones indican similitud a los del resto de la población, estas dos observaciones sugieren la integración de un 100% de los productos de ADN híbridos. Por otra parte y de manera análoga, el análisis por Mutation Calling explica en conjunto el 98% de dicha razón de integridad en las secuencias, debido a que gran parte de la población se ubica entre dos componentes de dispersión principal (fig. 26B); por otra parte la probabilidad que contienen los ejes con el 2 y 81% de dos componentes de dispersión donde se ubican los 5 controles en cuestión, sugieren dos vertientes: la primera señala que pudieran existir diferencias verdaderas en sus secuencias, sin embargo esta cuestión resulta muy poco probable porque como se menciona el 81% de la dispersión se atribuye a secuencias idénticas e integras; la segunda apunta hacia fluctuaciones térmicas del DHPLC. Esta última es mejor aceptada, deduciendo que el comportamiento de los 5 controles es mejor explicado por tales oscilaciones térmicas surgidas en el horno del DHPLC y no porque realmente pudiera existir un SNP relacionado con sus secuencias (Xiao and Oefner, 2001); en adición a esto, hasta el momento solo se tiene registro de una variante en el

fragmento evaluado y se localiza en el exón 1 (<http://www.ensembl.org/index.html>, 2011); la variante se ubica en el primer nucleótido del marco de lectura abierto y establece una mutación de gravedad, la cual fue encontrada en pacientes con macrohematuria y proteinuria (Longo *et al*, 2002). Por lo tanto, se reducen las posibilidades de dichos controles hacia un cambio verdadero en sus secuencias.



**Figura 21.** Identificación de los controles disimiles en la población de estudio (señalados con una flecha). **A.** Normalización de los cromatogramas que muestra el patrón distinto de los picos homoduplex. **B.** Localización de los mismos controles mediante análisis por Mutation Calling, el cual es estadísticamente significativo con un GAP = 0.02.

Partiendo del supuesto de integridad del fragmento evaluado; es importante resaltar que para los pacientes y controles –a menos que se indique lo contrario– los pasos de síntesis y degradación del gen *COL4A3*, que incluye la transcripción, empalme, transporte del ARNm, estabilidad del ARNm, inicio de la traducción, estabilidad proteica y biosíntesis de la cadena  $\alpha 3(\text{IV})$  podrían efectuarse conforme a distintas interacciones entre ADN-ARN ADN-proteína, ARN-ARN, ARN-proteína y proteína-

proteína dirigidas a partir de secuencias consenso dentro del ADN para desencadenar una serie de funciones bien definidas fundamentalmente críticas (Pesole *et al*, 2001). Las funciones implican en primer lugar a nivel de la regulación de la expresión génica, que las islas CpG podrían metilar y desmetilarse por acción de miRNAs (micro RNA monocatenario de entre 19 y 25 nucleótidos de longitud), en condición de que las proteínas reguladoras puedan actuar en su forma negativa y/o positiva (represor y activador, respectivamente) y que resulta fundamental para el comienzo de la regulación propicia de la expresión génica de *COL4A3* (Wilkie *et al*, 2003); este paso regulativo resulta de vital importancia por el continuo recambio que tiene la cadena  $\alpha 3(\text{IV})$  en la incorporación y remodelación de la estructura de la MBG (no tanto así para las MB de otros órganos y tejidos) (Momota *et al*, 1998 y sitio web: [www.ebi.ac.uk/](http://www.ebi.ac.uk/), 2011); en segundo lugar a nivel de transcripción (llevada cabo por la maquinaria de la RNA polimerasa II), la formación del capuchón 5' en el mRNA protegería al transcrito del ataque de exonucleasas, facilitar el escape de la RNA pol II del promotor para transcribir el resto del gen; por último en las funciones referentes hacia los procesos de traducción del mRNA, implicarían en primera, el transporte del transcrito maduro del núcleo hacia el RER para que se lleven a cabo los procesos de iniciación traduccionales que comienzan con el reclutamiento del complejo heteromultimérico de iniciación eucarionte (eIF) 4F (consiste en una proteína de unión al capuchón eIF4E, la helicasa dependiente de RNA eIF4A y su activador eIF4B, y el factor de andamiaje eIF4G) hacia el capuchón GpppN 5' del mRNA; la unión del complejo de pre-iniciación 43S (contiene la sub-unidad ribosómica 40S, el eIF3, y un complejo ternario de eIF2, GTP y Met-tARN<sub>i</sub>) al eIF4F; escaneo de la región UTR 5' por parte del complejo 48S hasta llegar al codón de inicio (AUG); disociación de los factores IF (hidrólisis del GTP unido a eIF2) y subsecuente unión de la sub-unidad ribosomal mayor 60S hacia la remanente 40S para formar el complejo de traducción ribosomal 80S (Pesole *et al*, 2001 y Strachan and Read, 2006); y por último, modificaciones post-traduccionales biosintéticas en el

extremo 5' del polipéptido que incluyen el reconocimiento y escisión del péptido señal por el complejo partícula de reconocimiento de la señal (SRP)-translocón transmembrana y peptidasa de la señal, respectivamente, en el RE; además de la hidroxilación vía enzimática de los residuos de Lys en el dominio 7S, para que finalmente culmine en la exocitosis de la cadena  $\alpha 3(\text{IV})$ , y así realizar su función como red estructural constituyente de la MBG (Agtmael and Tuderman, 2010).

Tomando en cuenta la vital importancia funcional que sucede alrededor de este fragmento y de que hasta ahora como antecedente solo se cuenta con un registro mutacional, es importante considerar hablando en términos evolutivos que la estructura génica de este segmento necesita estar altamente conservada y distanciada de SNPs ó mutaciones que amenacen con cambiar el rumbo proteico de la cadena  $\alpha 3(\text{IV})$ , porque al parecer cualquiera de sus componentes podrían ser poco tolerantes a cambios mínimos (en comparación con el dominio colagenoso que es más tolerante a estos cambios), ya que podrían resultar letales para el individuo a tal punto que alcanzaría con relativa facilidad IRCT (Jais *et al*, 2003 y Kashtan, 2006 y 2010).

De acuerdo a lo establecido, la resolución progresiva y más benigna de la enfermedad por SA en los niños es en parte considerada siempre y cuando las alteraciones mutacionales no afecten a los dominios no colagenosos (en este caso el dominio 7S) y mucho menos la región UTR 5', aunque en esta última no se ha encontrado asociación con esta enfermedad (Kashtan, 2006).

#### 7.4.2 Fragmento 2

En los cromatogramas de la población de estudio del fragmento 2 fueron identificados tres patrones de picos heteroduplex que corresponden a dos diferentes variantes alélicas para dos SNPs, uno en la región exónica y el otro en la intrónica (fig. 22A-C). El

control-molde utilizado en la hibridación resultó heterocigoto para los dos cambios (fig. 22A).

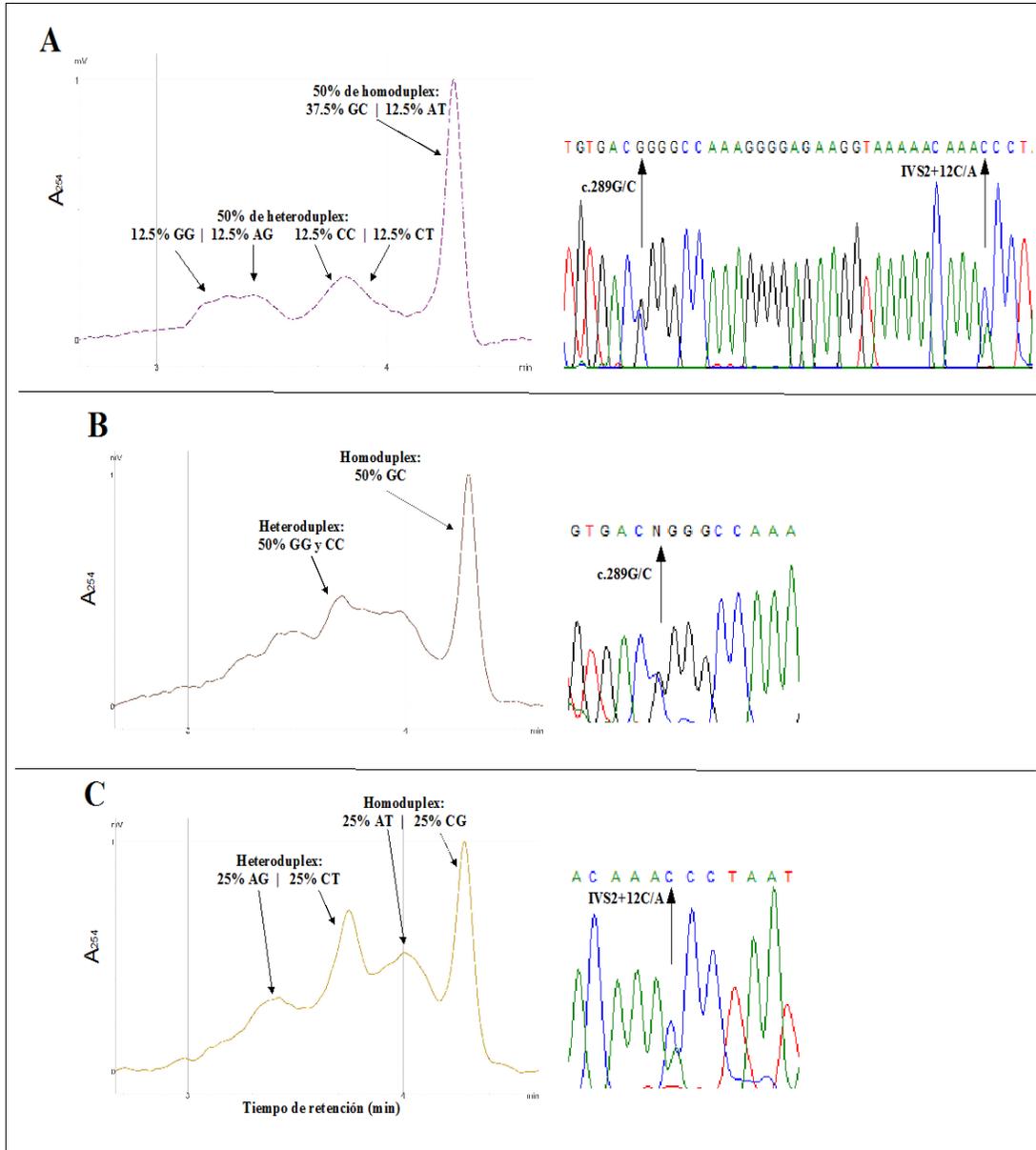
En la figura 22A se representan los cromatogramas con un patrón de heteroduplex definido encontrado en un paciente y 10 controles. De acuerdo a la  $T_m$  de 57.1 °C para este fragmento, el tiempo de retención para el pico homoduplex fue de 4.25 min cuando el buffer B alcanzó una tasa de elución del 59.75%; la relación [tiempo de retención-tasa de elución] de los picos heteroduplex fue directamente proporcional a la inestabilidad de nucleótidos mal alineados en los resultados híbridos de ADN, ya que el tiempo de retención para el primer pico heteroduplex (en seguida del pico homoduplex) fue de 3.78 min a una tasa de elución de buffer B de 58.90%, mientras que el siguiente pico heteroduplex presentó un tiempo de retención de 3.45 min con una elución de buffer B de 58.31%. De acuerdo al resultado de secuenciación se propuso el siguiente orden de estabilidad de los híbridos de ADN: GC > AT > CT y CC > AG y GG, por consiguiente el 50% de los homoduplex está representado por un 37.5% de GC y un 12.5% de AT, mientras que el otro 50% son de heteroduplex repartido en 12.5% para cada apareamiento desigual CT, CC, AG y GG (Oefner and Underhill, 1998 y Xiao and Oefner, 2001); es importante mencionar que estas proporciones son una aproximación que se deduce a partir de los pares de bases que potencialmente se pueden formar, según su existencia en la distribución original en ambos alélos cuando hibridan con el control-molde, de aquí en adelante será resaltada dicha cuestión. Cabe resaltar que este fragmento fue indistinguible la diferenciación de híbridos cuando tuvieron lugar entre bases púricas o pirimidínicas. La baja concentración de los picos heteroduplex respecto al homoduplex indica la heterogeneidad distribuida en dos señales distintas. Los SNPs c.289G/C (clasificación SNP humanos rs13424243) e IVS2+12C/A (rs1882435) localizados en la región exónica e intrónica, respectivamente, son las variantes alélicas que explican el patrón de picos visualizados en el cromatograma; las variantes se encuentran ampliamente reportadas en todos los grupos étnicos

(<http://www.ncbi.nlm.nih.gov/snp>, 2011). Cabe aclarar que a pesar de que no es uno de los objetivos identificar las variantes alélicas en los intrónes puesto que aparentemente no tienen relación alguna con la función génica para el gen *COL4A3*, existen excepciones como el SNP IVS2+12C/A encontrado en este fragmento, debido a la interferencia que presentó el análisis de los cromatogramas porque el control-molde utilizado para la hibridación con la población de estudio resultó portador de esta variable alélica. En cambio, la variable de interés exónica, es decir el SNP c.289G/C, describe una transversión que en el RNAm es localizada en la primera base del codón e implica un cambio del aa Gly a un aa de sentido erróneo que es Arg (p.G43R); la sustitución ocurre en el primer aa del dominio colagenoso. La repercusión de esta variable recaería en la alteración del sitio de torsión del primer pliegue para conformar la triple hélice cuando la cadena  $\alpha 3(\text{IV})$  se une a las cadenas  $\alpha 4-\alpha 5(\text{IV})$ , debido a que el residuo Gly es el único aa sin un átomo  $\text{C}_\beta$ , esto le permite estar mucho menos impedido estéricamente que cualquier otro aa para conformar su extensión en los ángulos diedricos o de torsión  $\phi$  y  $\psi$ , los cuales llevarían a los átomos más cerca que la distancia correspondiente de Van del Waals cuando se genera la rotación y plegamiento de la proteína; por lo tanto esta conformación resultaría prohibida para el aa Arg (Mariyama *et al*, 1994 y Voet *et al*, 2006); por consiguiente el SNP puede implicar un impacto químico, mecánico y físico, sobre la función puntual del sitio identificado con la variante alélica (Srinivasan *et al*, 2009). Debido a la razón expuesta, se deduce que la naturaleza no polar del aa Gly orientado hacia el interior de la cadena  $\alpha 3(\text{IV})$  para generar puentes de hidrógeno inter-cadenas con los residuos de Pro -por triplete Gly-X-Y- (indispensables para la formación del protómero en triple hélice) y establecer una  $\Delta G^\circ = -1.4$  kcal/mol (energía libre) necesaria para la estabilización del sitio afectado (Stryer *et al*, 2003 y Shoulders and Raines, 2009), se vería interrumpido en primer lugar por la sustitución de un aa polar -con cadena lateral cargada positivamente- como lo es la Arg, y desde otro punto de vista, tomando en cuenta el tamaño de dicho aa, se

generaría una protrusión y consecuente distorsión puntual al inicio del dominio colagenoso en la cadena  $\alpha 3(\text{IV})$ , provocando una menor estabilidad de la misma, la cual pudiese influir en la disminución del mantenimiento de la barrera de filtración glomerular a largo plazo (Gubler, 2008). Por lo tanto, es importante establecer que mientras más sustituciones recaigan en los residuos de Gly, existiría una mayor propensión hacia una menor estabilidad de la cadena  $\alpha 3(\text{IV})$  -cuando conforma el protómero- en su papel funcional estructural en la MBG; sin embargo, al parecer si solo existe uno ó muy pocos aa de Gly reemplazados en el dominio colagenoso, la cadena  $\alpha 3(\text{IV})$  por sí misma podría ser poco estable, pero tolerante cuando se incorpora a la estructura del protómero, debido a la cantidad de puentes de hidrógeno potencialmente posibles con residuos Pro, cuando se toma en cuenta un promedio de 1,359 residuos de Gly en todo el protómero (453 para la cadena  $\alpha 3(\text{IV})$ ) (Tryggvason and Patrakka, 2006 y <http://www.ensembl.org/index.html>, 2011).

En la figura 22B son representados los cromatogramas con un patrón de picos heteroduplex encontrados en 4 pacientes y 10 controles. El tiempo de retención del pico homoduplex se mantuvo en 4.25 min, en cambio el pico heteroduplex registro un tiempo de retención de 3.71 min a una tasa de elución promedio de buffer B de 58.77% (fue tomado el pico más alto de un conjunto distorsionado y poco diferenciado del supuesto pico heteroduplex). Conforme al resultado de secuenciación, el orden de estabilidad de los híbridos fue GC > CC > GG; de los cuales el 50% de los homoduplex fue constituido por GC, mientras que el heteroduplex fue distribuido por un 25% de CC y 25% de GG. La concentración de ADN para la relación heteroduplex fue mayor, debido a que establecieron el 50% de los híbridos de ADN (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). El comportamiento de los picos para este cromatogramas se debió a la localización del SNP c.289G/C (descrito en la figura 26A). Por último, es importante mencionar que la cantidad neta de individuos portadores de este SNP (tomando en cuenta a los individuos de la figura 26A) es de 5 pacientes y 20 controles.

En la figura 22C se muestra la representación de un conjunto de cromatogramas idénticos encontrados en 4 pacientes y 17 controles. Aquí pudieron distinguirse con mayor claridad los picos que representaban cierta proporción de los ADN híbridos; en primer lugar y al igual que los otros dos cromatogramas, el tiempo de retención para el pico de mayor estabilidad homoduplex fue de 4.25 min, mientras que el siguiente pico (también homoduplex) registro 4.07 min a una elución de buffer B de 54.42%; por otro lado para los picos heteroduplex los tiempos de retención fueron de 3.75 y 3.38 min cuando alcanzaron la elución del buffer B de 58.85 y 58.18%, respectivamente. Por lo tanto y en concordancia con el resultado de secuencia para este cromatograma, el orden de estabilidad fue  $CG > AT > CT > AG$ , los cuales tuvieron una distribución equitativa del 25% (aunque en la concentración de ADN de cada híbrido fuese desigual) (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). Lo mostrado en el cromatograma fue por la identificación del SNP IVS2+12C/A, que en su resolución neta (incorporando a los individuos de la figura 27C) fue encontrado en 5 pacientes y 27 controles.



**Figura 22.** Identificación de las variantes alélicas polimórficas c.289G/C e IVS2 + 12C/A en el exón e intrón 2, respectivamente, en estado heterocigoto. **A.** Control-molde identificado con los SNPs c.289G/C e IVS2 + 12C/A, los cuales representan la distribución alélica resultante de hibridación mediante picos homo y heteroduplex, y su naturaleza y posición en la secuencia. **B.** Motivo del cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.289G/C, el cual se define en la secuencia. **C.** Motivo del cromatograma con pico heteroduplex con distribución del alélo IVS2 + 12C/A, el cual se define en la secuencia.

### 7.4.3 Fragmento 3

En los cromatogramas de la población de estudio del fragmento 3 fue identificado un patrón de pico anormal no heteroduplex en 2 pacientes, el cual corresponde a cuatro variables alélicas idénticas y continuas (fig. 23C). El control-molde utilizado en la hibridación resultó silvestre (fig. 23B).

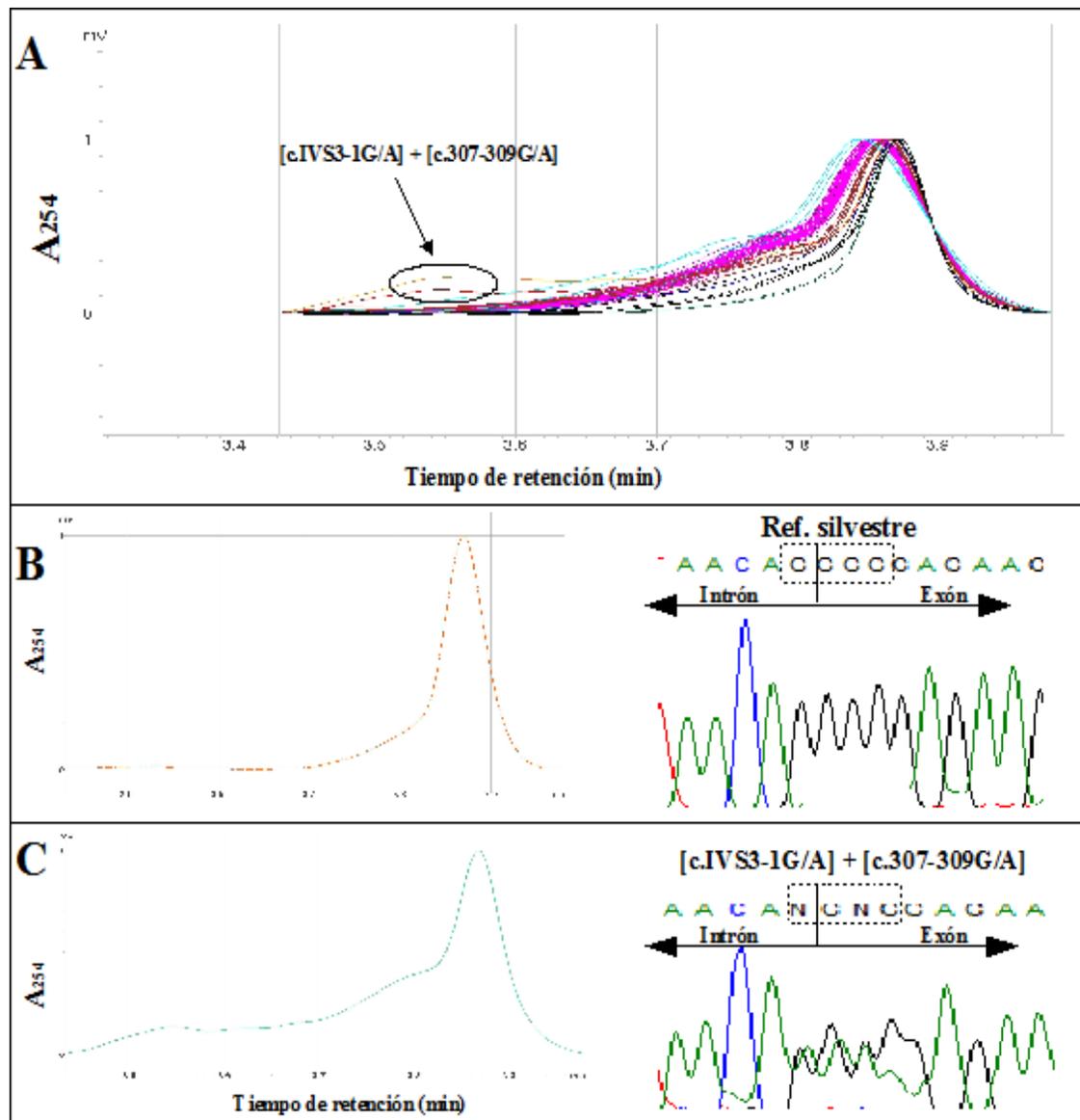
En el cromatograma de la figura 23A se muestran los datos normalizados de la población donde fueron detectados los dos pacientes, los cuales son visualizados con ligera discrepancia hacia el levantamiento inicial de sus picos homoduplex; no mostraron diferencias de elución ó tiempo de retención en la columna, por lo que sus concentraciones de ADN y tiempos de retención -que en promedio son de 3.85 min a una tasa de elución de 59.23% proporcionado por el buffer B- son iguales a los del resto de la población. La pobre resolución cromatográfica de los dos pacientes con semejante multivariante alélica detectada, se atribuye a una razón diferencial no alcanzada por la  $T_m$  predicha por el DHPLC para este fragmento (60.2 °C), debido a los múltiples desalineamientos de los nucleótidos; es decir, tomando en cuenta que la detección diferencial se observó al inicio de los picos normalizados y que en este sitio recae un tiempo de retención aproximado de 3.55 min -que además supondría una elución de buffer B de 58.69%- y que las variantes alélicas abarcan cuatro nucleótidos en un sitio denso dentro de la secuencia, es muy posible que  $\Delta G^\circ$  disminuyera a tal punto que afectara la  $T_m$  del fragmento y como consecuencia requiriera una  $T_m$  menor para que los picos heteroduplex fueran mejor definidos; aunque por otro lado la  $T_m$  predicha por el DHPLC supone que es la óptima para el fragmento íntegro (es decir 100% homoduplex) y en  $T_m$  superiores o inferiores a esta, anularían la detección de los picos. Debido a las razones expuestas, puede decirse que el inicio de la señal para la detección homoduplex y la poca señal de detección para los dos pacientes -que en consecuencia no mostraron picos heteroduplex marcados- es porque ambos sitios recaen sobre un umbral de

detección por el DHPLC y la  $T_m$  de 60.2 °C favoreció el pico homoduplex (Oefner and Underhill, 1998 y Xiao and Oefner, 2001).

En la figura 23B se muestra el cromatograma de un pico homoduplex íntegro que representa a los 11 pacientes y todos los controles -que también son visualizados en la normalización de los picos de la población de la figura 27A-. Mientras que en su respectiva secuencia figura la región de transición exón-intrón en su forma silvestre, la cual sirve como referencia para discernir de la variante mutada que fue hallada en los 2 pacientes (región encerrada por el recuadro punteado); dicha región presenta una posición funcionalmente bien definida para que se efectuó el empalme correcto del exón 2 con el exón 3, debido en parte a que la secuencia de interés establece sitios diana para la maquinaria del ayustosoma (ribonucleoproteínas pequeños nucleares) y de proteínas que potencian el empalme exónico; sitio conservado de AG del intrón 3' y la región de guaninas iniciales del exón 5', respectivamente (Fairbrother *et al*, 2004).

En la figura 23C se representa el cromatograma con pico homoduplex anormal de los dos pacientes; sin embargo aunque no logren diferenciarse picos definidos, de acuerdo al resultado de su secuencia, los híbridos de ADN potencialmente formados supondrían un orden de estabilidad  $CG > TA > CA > TG$ , los cuales estarían conformados en una proporción del 25% para cada híbrido. La razón de este supuesto se debe a la definición de las cuatro variantes alélicas encontradas: [c.IVS3 -1G/A] + [c.307-309G/A], las cuales constituyen un cambio heterocigoto entre bases de transición (bases puricas); la primera se localiza en el último nucleótido de la secuencia consenso GT en el extremo 3' del intrón 2, mientras que la segunda recae en la totalidad de las tres bases que configuran para el primer codón del exón 3. Las variantes alélicas son mutaciones de *novo* potencialmente dañinas, debido al impacto adverso que pudieran presentar sobre la transcripción y función de la cadena  $\alpha 3(IV)$ ; aunque esta condición sería posible siempre y cuando la mutación actuó en su forma dominante (SAAD); en

caso contrario solamente serían portadores latentes con fenotipo estable para ese sitio de la proteína, ya que los pacientes también portan un alélo normal que aleja la condición patológica de la enfermedad (Heidet *et al*, 2001). Si la forma de SAAD es establecida como condición funcional en los pacientes detectados; en primera instancia para el procesamiento del pre-mRNA, la mutación [c.IVS3 -1G/A] conjetura que existe una posibilidad nula o reducida de que se pueda efectuar el ataque nucleófilo hacia el grupo fosforilo en el sitio de empalme 3' durante la segunda reacción de transesterificación para lograr fusionar al exón 2 con el 3 y liberar al intrón 2, debido a que la deficiencia de esta catálisis generaría poca precisión y eficacia hacia el empalme constitutivo, ya que la subunidad 35 de la ribonucleoproteína U2AF (factor auxiliar del RNA pequeño nuclear U2) -reclutada por los potenciadores exónicos de empalme- no podría reconocer un sitio alterado con la secuencia AA en la región de empalme 3', por lo que existe la posibilidad de que el exón 3 sea omitido y que el sitio de empalme 5' del exón 2 se apare con otro sitio 3' más allá del correcto (posiblemente el exón 4) y que finalmente conlleve a que la cadena  $\alpha 3(\text{IV})$  contenga 30 residuos aa menos en el dominio colagenoso (Pesole *et al*, 2001; Longo *et al*, 2006 y Watson *et al*, 2006); las consecuencias de esta última acción consideran que tal vez exista una dificultad simétrica en la configuración espacial de la cadena  $\alpha 3(\text{IV})$  para establecer los puentes de hidrógeno requeridos en la formación de un protómero estable junto con las cadenas  $\alpha 3-\alpha 4(\text{IV})$  (Shoulders and Raines, 2009), aunque si el protómero lograra formarse e incorporarse a la MBG, sería un diana fácil de degradación para las metaloproteasas, entonces la MBG culminaría en ofrecer poca resistencia al filtrado glomerular (Agtmael and Tuderman, 2010). Por otro lado, la mutación que abarca tres variantes alélicas [c.307-309G/A], consiste en un cambio del aa Gly a un aa de sentido erróneo que es Arg (p.G49K), del cual se atribuye que tiene la misma razón funcional e impacto estructural en el dominio colagenoso que el explicado para la variante alélica c.289G/C del exón 2 (véase sección fragmento 2).



**Figura 23.** Identificación de las variantes alélicas mutacionales [c.IVS3 -1G/A] + [c.307-309G/A] en el sitio de empalme y en el primer codón del exón 3, respectivamente. **A.** Normalización de los cromatogramas que muestra el patrón ligeramente distinto de dos picos en dos pacientes (señalados con una flecha), los cuales corresponden a las dos mutaciones [c.IVS3 -1G/A] + [c.307-309G/A]. **B.** Control-molde con resultado homoduplex que en su respectiva secuencia muestra la región de referencial mutacional, íntegra. **C.** Distorsión del pico cromatográfico que en su secuencia resguarda en estado heterocigoto los alélos mutacionales [c.IVS3 -1G/A] + [c.307-309G/A].

La casualidad de incidencia idéntica para una misma mutación -que abarca cuatro posiciones nucleótídicas- en ambos pacientes no emparentados entre sí puede que sea muy poco probable, por lo que esta condición sugiere que los pacientes se

encuentran intrínsecamente relacionados entre sí, de tal forma que pudieran ser hermanos(as); al parecer dicha mutación que configura y/o contribuye a la enfermedad, tuvo su aparición inicial en uno de los padres, donde al parecer es mejor explicada en función de la forma en que aparece el SAAD, ya que es la causa más común de mutación espontánea de entre las tres formas distinguidas de SA (SAAR, SAAD y SALX) y por lo tanto rara vez tiene un historial familiar (Pessucci *et al*, 2004). La mutación en uno de los padres sugiere con mayor posibilidad una etiología que implica deficiencias en el mecanismo de replicación del ADN; la cual en primera instancia considera un “resbalón” en la acción catalítica de la polimerasa por incorporar los nucleótidos AAAA en vez de los correctos GGGG; error conjeturado tal vez por la misma naturaleza química que comparten ambos nucleótidos (bases púricas), y que al parecer la polimerasa no pudo distinguir entre estas bases de manera eficiente; sin embargo en un segundo plano, los mecanismos de reparación del ADN también sugieren que no pudieron detectar y reparar la variante mal apareada, ó simplemente fue detectada pero no pudo ser reparada; por lo tanto y en consecuencia las rondas sucesivas que tuvieron lugar cuando se replicó el ADN parental, ocasionaron la fijación de la mutación que posteriormente fue heredada a los dos supuestos hijos (Wang *et al*, 2004; Watson *et al*, 2006 y Lewin, 2008).

Es importante señalar que la mutación descrita fue la única que se detectó de entre todos los fragmentos que conformaron el análisis de este estudio; lo cual establece que ambos pacientes representan el 15.38% del total de pacientes estudiados y por lo tanto sobrestima el 5% estadísticamente preestablecido para la posible forma de SAAD. Para los 11 pacientes restantes se establece que la etiología de su enfermedad puede que radique en mutaciones en cualquiera de los 45 exones restantes del gen *COL4A3* o en aquellos conformados por los genes *COL4A4* y *COL4A5*; estadísticamente se esperaría que la mayoría de los pacientes tuviera la forma de SALX, ya que es la forma más común

de la enfermedad (aproximadamente el 80% de los casos) (Thorner, 2007 y Medeiros *et al*, 2008).

#### 7.4.4 Exón 7

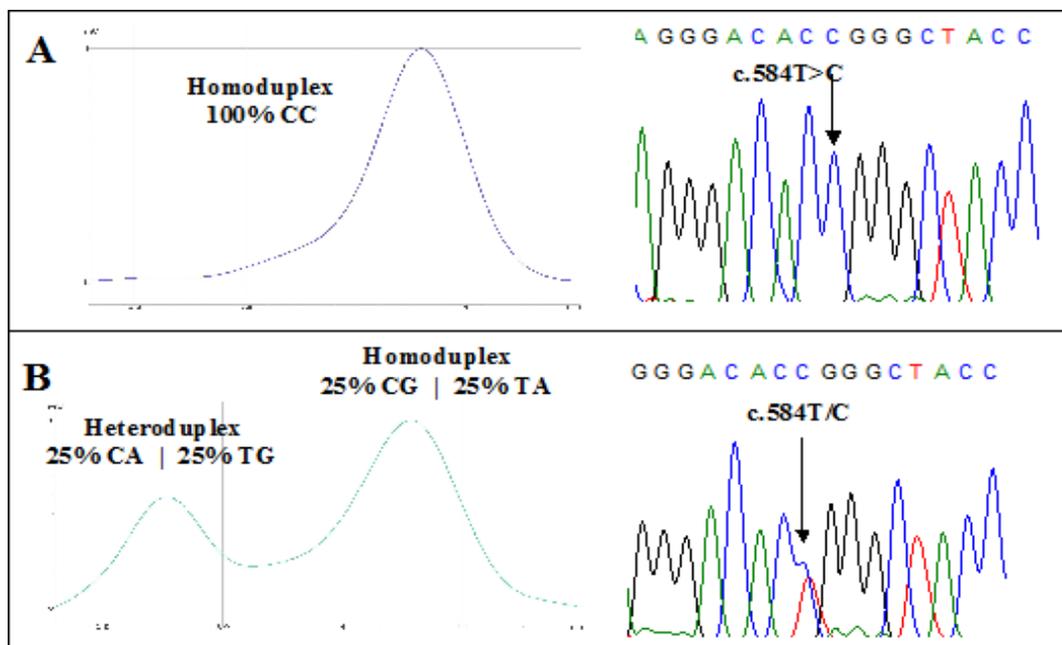
En los cromatogramas de la población de estudio del exón del fragmento 7 fueron identificados dos patrones de picos en su forma homo y heteroduplex, los cuales corresponden a dos variantes alélicas que son alternativas de la silvestre (fig. 24A y B). El control-molde utilizado en la hibridación resultó homocigoto para la variante alternativa de la silvestre (fig. 24A).

En la figura 24A se representan los cromatogramas con un patrón de picos homoduplex encontrados en 11 pacientes y 41 controles. A partir de la  $T_m$  de 55 °C predicha por el DHPLC para este fragmento, se registro un tiempo de retención de 4.06 min cuando el buffer B alcanzó una tasa de elución del 59.4%, indicando que es directamente proporcional a la integridad del 100% de los apareamientos correctos entre las bases de ADN híbridos (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). La variable que explica la razón del pico homoduplex es el SNP c.584T>C (rs10178458), el cual es considerado el alélo ancestral y además se encuentra más ampliamente reportado en todos los grupos étnicos que aquel designado como silvestre (es decir homocigoto T) (<http://www.ensembl.org/html.com>, 2011). El SNP describe una transición que en el RNAm se ubica en la segunda base del codón e implica un cambio del aa Leu a un aa de sentido erróneo que es Pro (p.L141P); de acuerdo con Watson y colaboradores (2006), en percepción de orden en la composición del código genético, el codón con una base pirimidínica en la segunda posición -como en este caso- reemplazo un aa hidrófobo por otro de esta naturaleza. El aa Pro -después del de Gly- es el residuo con mayor abundancia relativa, representando el 18.56% de los aa que componen la cadena  $\alpha 3(IV)$ . La variante Pro se ubica en la posición Y de la combinación Gly-X-Y en

el dominio colagenoso, por consiguiente su función contribuiría más hacia la estabilidad térmica en este sitio de la cadena  $\alpha 3(\text{IV})$ , porque sería diana de la enzima prolil 3-hidroxilasa en el aparato de Golgi para convertir el aa Pro en Hyp; esto conllevaría a la disminución del costo entrópico cuando la cadena  $\alpha 3(\text{IV})$  es plegada para conformar el protómero, debido a que se reduce la interferencia estérica cuando el residuo de Pro se isomeriza en la conformación *cis* (es decir que el  $\text{C}\alpha$  sucesivo esta del mismo lado del enlace péptidico), estableciendo enlaces péptidicos con un ángulo  $\phi$  aproximadamente  $-60^\circ$  de torsión requerida para conformar la triple hélice (Voet *et al*, 2006 y Shoulders and Raines, 2009). Por lo tanto, tomando en cuenta la relación del aa Pro con la mantención de la estabilidad estructural y térmica de la proteína, es difícil concebir un efecto adverso en el producto de este SNP, por lo que finalmente se deduce que la variable alélica que da como resultado la codificación del aa Pro, mejora la integridad y resuelve mayor estabilidad en el sitio colagenoso de la cadena  $\alpha 3(\text{IV})$  donde ocurre el cambio; esta mejora de función sugiere que la fijación de este SNP sea benéfico bajo el contexto evolutivo (Freeman y Herron, 2001).

En la figura 24B se representan los cromatogramas con un patrón de picos heteroduplex hallados en un paciente y 6 controles. El tiempo de retención del pico homoduplex se mantuvo en 4.06 min, mientras que el pico heteroduplex lo hizo en 3.85 min a una tasa de elución por parte del buffer B de 59.03%. De acuerdo al resultado de secuenciación, el orden de estabilidad de los híbridos fue  $\text{CG} > \text{TA} > \text{CA-TG}$ ; las cadenas híbridas de ADN que contienen el apareamiento correcto CG y TA representan al pico homoduplex con una distribución del 25% de cada híbrido, consecuentemente el pico heteroduplex lo representa el conjunto de híbridos CA-TG con la misma distribución del 25% de cada híbrido. La razón de tener un pico heteroduplex homogeneizado y no diferenciado de los híbridos CA-TG se atribuye a que ambas combinaciones pudieron presentar la misma distorsión en la configuración espacial de la cadena de ADN que tuvo lugar en el sitio donde se ubican estos híbridos, en consecuencia presentaron la

misma retención en la columna del DHPLC (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). La variante alélica que explica la definición del pico heteroduplex es el mismo SNP de la figura 24A, pero esta vez en su forma heterocigota, es decir c.584T/C; esta variante también se encuentra reportada en todas las poblaciones, aunque en menor medida (Stabuc *et al*, 2009 y <http://www.ncbi.nlm.nih.gov/snp>, 2011). El SNP genera un producto no sinónimo, ya que tiene la facultad para codificar tanto el aa Pro como Leu, este último sencillamente cumple su función como un componente más en la estructura de la cadena  $\alpha 3(IV)$  (Mariyama *et al*, 1994).



**Figura 24.** Identificación de las variante alélica polimorfica c.584T>C y T/C localizada en el exón 7 en estado homo y heterocigoto, respectivamente. **A.** Control-molde identificado con el SNP c.584T>C, el cual por su estado homocigoto se representa en el cromatograma con pico homoduplex, la naturaleza y posición del cambio se define en la secuencia. **B.** Motivo del cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.584T/C, el cual se define en la secuencia.

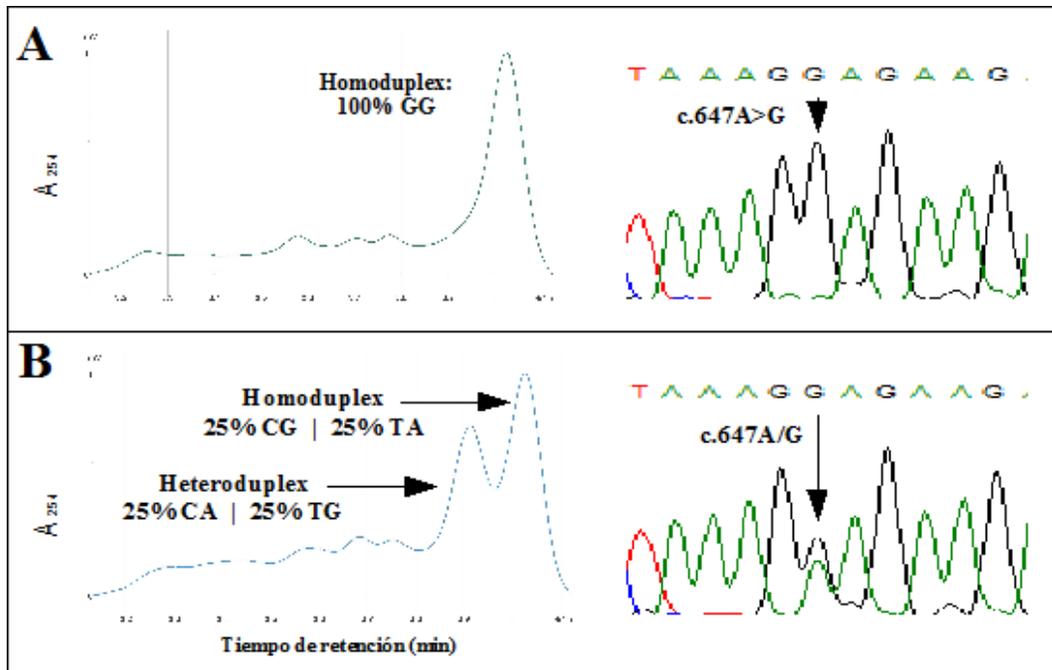
#### 7.4.5 Exón 9

En los cromatogramas de la población de estudio del exón del fragmento 9 fueron identificados dos patrones de picos homo y heteroduplex, las cuales corresponden a las dos variantes alélicas de la silvestre (fig. 25A y B). El control-molde utilizado en la hibridación resultó homocigoto para la variante alternativa de la silvestre (fig. 25A).

En la figura 25A se muestra un cromatograma que de manera similar fue detectado en 8 pacientes y 43 controles. A partir de la  $T_m$  de 55.4 °C predicha para el horno de la columna del DHPLC, logro identificarse el pico homoduplex a un tiempo de retención de 4.09 min cuando el buffer B alcanzo una tasa de elución de 59.56%, indicando que los híbridos de ADN conservaron el 100% de integridad en su estructura cuando tuvo lugar el realineamiento de los ADN duplex. La variante alélica que indica el comportamiento del pico homoduplex es el SNP c.647A>G (rs6436669), el cual se encuentra solamente reportado en algunas poblaciones, estudios filogenéticos lo consideran como alélo ancestral (<http://www.ensembl.org/index.html>, 2011). Este SNP consiste en una transición que ocurre en la segunda base del codón; en el RNAm establece un cambio de sentido erróneo que define para el aa hidrófobo Gly en vez de hacerlo para el residuo Glu (E162G). Es importante resaltar que dicha variable se localiza en el tercer residuo aa (de 13 residuos aa) de la primera interrupción no colagenosa (de 23 registradas) dentro del dominio colagenoso de la cadena  $\alpha 3(IV)$ , en conjunto esta región juega un papel central como uno de los centros de organización supramolecular para integrar estructuralmente la MBG (Mariyama *et al*, 1994 y Voskarides *et al*, 2008); tomando en cuenta el rol individual del residuo Gly dentro de este segmento no colagenoso, puede que contribuya y proporcione una interacción molecular no tan rígida, debido a que este aa es incapaz de proporcionar una carga lateral polar que es necesaria - que si lo hace el aa Glu- para interaccionar con distintas biomoléculas como la laminina-11 ( $\alpha_5-\beta_2-\gamma_1$ ) vía receptores nidogeno/enactina, esto con el fin de establecer un andamiaje estable para formar segundas redes que definen

mayor resistencia mecánica frente al filtrado glomerular (Gubler, 2008). Dichs aspectos sugieren que aunque esta variable en su forma homocigota hasta el momento no se inclina en sí hacia una cierta patología, cabe la posibilidad que durante el transcurso de la vida del individuo, la función de la variante valla deteriorándose en este sitio del dominio colagenoso, por el reflejo que conjetura en menor medida hacia una menor resistencia mecánica frente al filtrado glomerular, debido a fluctuaciones que pudieran presentarse durante las presiones hidrostáticas y coloidosmóticas constantes como demandas homeostáticas de la función renal (Guyton and Hall, 2009).

En la figura 25B se muestra un cromatograma con pico heteroduplex que de manera similar fue encontrado en 5 pacientes y 7 controles. El pico homoduplex sostuvo los 4.09 min de retención y el pico heteroduplex registro 3.91 min a una tasa de elución de buffer B de 56.11%. El orden de estabilidad propuesto por el resultado de secuencia contempla la combinación de híbridos  $CG > TA > CA-TG$ ; el pico homoduplex está compuesto por un 50% del ADN híbrido y está conformado por un 25% de CG y un 25% de TA, en cambio el otro 50% de los híbridos fue heteroduplex y está constituido por un 25% de CA y un 25% de TG (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). La variante alélica que determina el pico heteroduplex es la forma heterocigota del SNP c.647A/G, por lo tanto puede codificar tanto para el aa Glu como Gly; esta variante se encuentra más ampliamente distribuida en todos los grupos étnicos que la de su forma homocigota (<http://www.ncbi.nlm.nih.gov/snp>, 2011).



**Figura 25.** Identificación de la variante alélica polimorfica c.647A>G y A/G localizada en la primera interrupción no colagenosa del dominio colagenoso en el exón 9, en estado homo y heterocigoto, respectivamente. **A.** Control-molde identificado con el SNP c.647A>G, el cual por su estado homocigoto se representa en el cromatograma con pico homoduplex, la naturaleza y posición del cambio se define en la secuencia. **B.** Motivo del cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.647A/G, el cual se define en la secuencia.

#### 7.4.6 Exón 21

En los cromatogramas de la población de estudio del exón del fragmento 21 fueron identificados tres patrones de picos heteroduplex que corresponden a tres diferentes variantes alélicas para dos SNPs distintos (fig. 26A-C). El control-molde utilizado en la hibridación resultó en un alélo homocigoto para la variante alternativa de la silvestre, y para el otro alélo homocigoto silvestre (fig. 26A).

En la figura 26A se muestra un cromatograma con pico homoduplex que de manera similar también fue encontrado en 9 pacientes y 32 controles. De acuerdo a la  $T_m$  de 52.9 °C predicha por el DHPLC, el pico homoduplex registro un tiempo de retención de 3.51 min a una tasa de elución de buffer B de 63.11%. De acuerdo a lo visualizado en el resultado de secuencia, el producto del pico homoduplex interpreta a

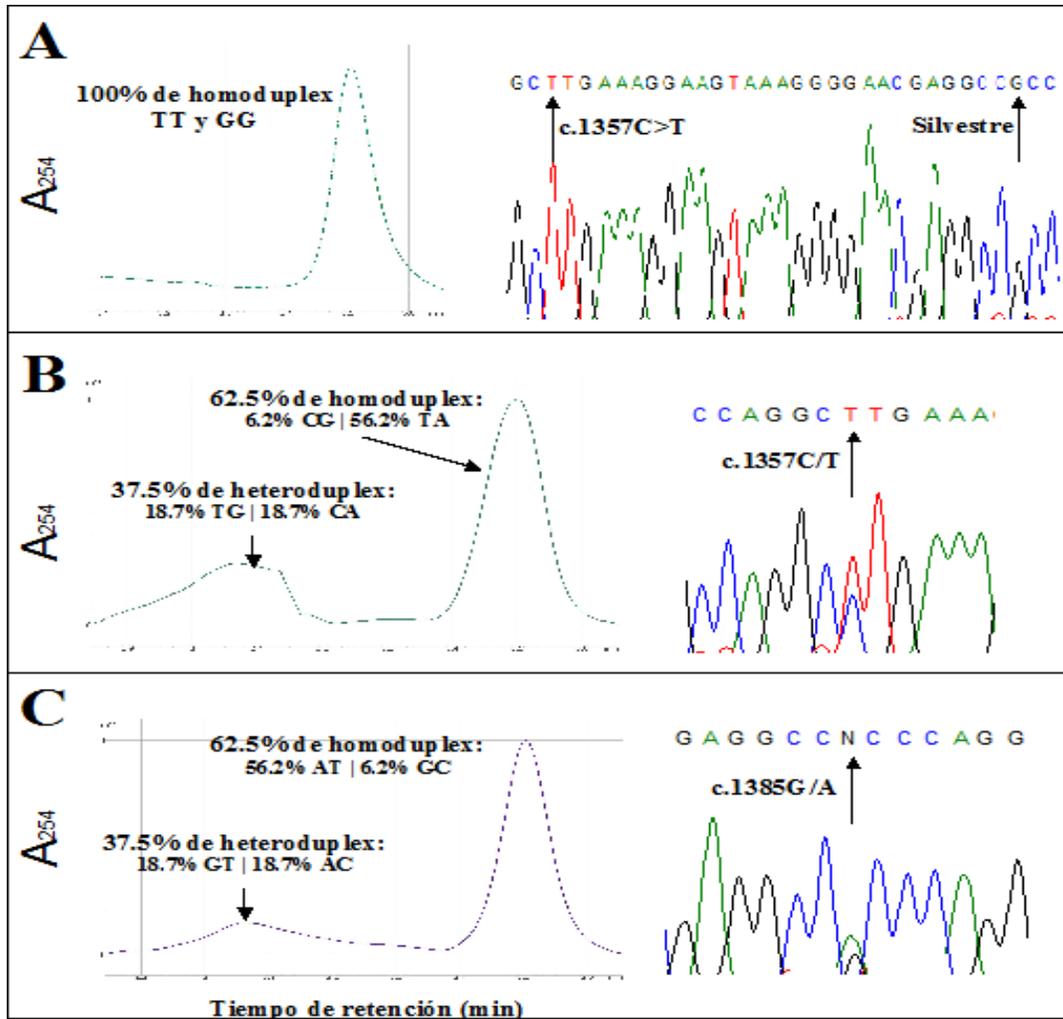
dos nucleótidos que hacen referencia a las dos posiciones que indican las variantes alélicas detectadas en este fragmento (señaladas en la figura 26B y C, respectivamente). La variante en forma homocigota corresponde al SNP c.1357C>T (rs10205042), y consiste en una transición sinónima, que en el RNAm se localiza en la primera base del codón, por lo tanto no altera el producto del aa original que es Leu (p.L399L); de acuerdo con Watson (2006), es común que un cambio en la primera posición del codón resulte en un aa similar ó el mismo, como en este caso (Lewin, 2008). Hasta el momento no existen registros genotípicos en la distribución de este SNP en las poblaciones. La presencia promedio del aa Leu en las proteínas es del 9.1%, siendo el de mayor abundancia relativa de entre los 20 aa (Voet *et al*, 2006), sin embargo en la proteína  $\alpha 3(\text{IV})$  constituye solamente el 5.5% y se encuentra disperso en todo el dominio colagenoso, en la mayoría de los casos después de un residuo Gly (Mariyama *et al*, 1994 y Voet *et al*, 2006); por esta razón, el papel que podría tener el aa Leu para la cadena  $\alpha 3(\text{IV})$  limitaría únicamente hacia una participación estructural en la configuración de menor rigidez para los sitios nacientes del surco mayor de la cadena  $\alpha 3(\text{IV})$  cuando se pliega en triple hélice junto a las cadenas  $\alpha 4$  y  $\alpha 5(\text{IV})$  (Shoulders and Raines, 2009).

En la figura 26B se muestra un cromatograma con pico heteroduplex hallado en 4 pacientes y 18 controles, donde el pico homoduplex se mantuvo con un tiempo de retención de 3.51 min, mientras que el pico heteroduplex tuvo 3.07 min a una tasa de elución de buffer B de 62.32%. Respecto a la secuencia obtenida, la estabilidad resultante de los híbridos deduce el siguiente orden  $\text{CG} > \text{TA} > \text{TG} > \text{CA}$ , distribuido en 62.5% de homoduplex repartido en 6.2% de CG y 56.2% de TA, y por el otro un 37.5% de heteroduplex dividido en 18.7% de TG y 18.7% de CA. La distorsión estructural presentada por los híbridos heteroduplex en las cadenas de ADN, corresponden a una elución secuencial y poco diferenciada entre el apareamiento resultante de TG y CA en la columna del DHPLC a la correspondiente  $T_m$  del fragmento (52.9 °C) (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). La variante alélica que explica la razón de

este cromatograma es el mismo SNP descrito en la figura 26A, pero en su forma heterocigota resuelve una transición con un contenido de bases pirimidínicas: c.1357C/T; la variante alélica también es sinónima en la codificación para el aa Leu, por lo tanto también posee las mismas características estructurales y funcionales que la variante c.1357C>T descrita en la figura 26A (<http://www.ensembl.org/index.html>, 2011).

Por último, en la figura 26C se muestra la definición grupal de un pico heteroduplex encontrado en 3 pacientes y 24 controles, donde el pico homoduplex también registro un tiempo de retención de 3.51 min y el heteroduplex de 3.06 min a una tasa de elución de buffer B de 62.30%. La estabilidad de híbridos resultantes en la conformación de cadenas fue GC > AT > GT > AC, esto de acuerdo a lo establecido por las secuencias. Debido a esto, de manera análoga al pico de la figura 26B, el 62.5% del pico homoduplex está representado por un 6.2% de GC y 56.2% de AT, dejando un 37.5% aproximadamente del heteroduplex repartido de un 18.7% de GT y 18.7% de AG. En lo que respecta la concentración del pico heteroduplex, hace referencia a la igualdad de elución indistinta, otorgada por el horno del DHPLC ante las distorsiones estructurales similares que presentaron las hibridaciones no complementarias GT y AC, debido a que la presencia de la base timina conduce a una  $\Delta G^\circ$  menos negativa, sugiriendo que es relativamente análoga al de un apareamiento incorrecto entre purinas (Schollen *et al*, 2005); por otro lado considerando que en suma estas hibridaciones reflejan el 37.5% del total de híbridos, no es depreciable la baja concentración presentada por este pico. El comportamiento de este cromatograma se manifiesta por la identificación en la secuencia del SNP c.1385G/A (rs34505188), el cual consiste en una transición y que en el RNAm se ubica en la segunda base del codón y posee la facultad para codificar dos aa, generalmente el residuo común es Arg, sin embargo la variante alternativa repercute en un aa de sentido erróneo que es His (p.R408H). Cabe aclarar que hasta el momento no existen registros genotípicos acerca de este SNP en las

poblaciones. En concordancia con Watson y colaboradores (2006), el cambio en la segunda posición del codón, para este SNP fue reemplazado un aa (Arg) por otro muy semejante (His); debido a que los dos aa poseen la misma naturaleza de ser polares y con cadenas laterales cargadas positivamente, la His también puede aparecer en su forma neutra por ser el único aa que tiene un  $pK$  (constante de disociación) en su cadena lateral cercano al pH fisiológico de 7 (Voet *et al*, 2006); por lo tanto también surge la posibilidad de que la His no solo desempeñe un papel estructural en la cadena  $\alpha 3(IV)$ , sino que si el aa es hallado en su forma neutra, puede que sea incapaz de servir como estabilizador electrostático frente a fluctuaciones negativas proporcionadas por el aa Glu (Yurchenco *et al*, 2004). De acuerdo a lo establecido sobre la presencia promedio de los aa Arg e His en las proteínas que es de 5.1 y 2.3% (Voet *et al*, 2006); para la cadena  $\alpha 3(IV)$  se encuentran las proporciones de 3.83 y 1.13%, respectivamente, y suelen ubicarse hacia el lado adyacente inmediato de un aa Gly a lo largo de la cadena proteica, su presencia sugiere configuración hacia una zona de menor tensión estructural para la proteína (Mariyama *et al*, 1994).



**Figura 26.** Identificación de las variantes alélicas polimórficas c.1357C>T y C/T, y c.1385G/A en el exón 21. **A.** Control-molde identificado con el SNP c.1357C>T, donde además también se indica el aléilo silvestre implicado con la otra variante alélica, ambos se muestran íntegros en el cromatograma debido a su pico homoduplex definido, la naturaleza y posición del polimorfismo se muestra en la secuencia correspondiente. **B.** Motivo del cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.1357C/T, el cual se define en estado heterocigoto en la secuencia correspondiente. **C.** Motivo del cromatograma con pico heteroduplex con distribución del polimorfismo c.1385G/A, el cual también se muestra en estado heterocigoto en su secuencia correspondiente.

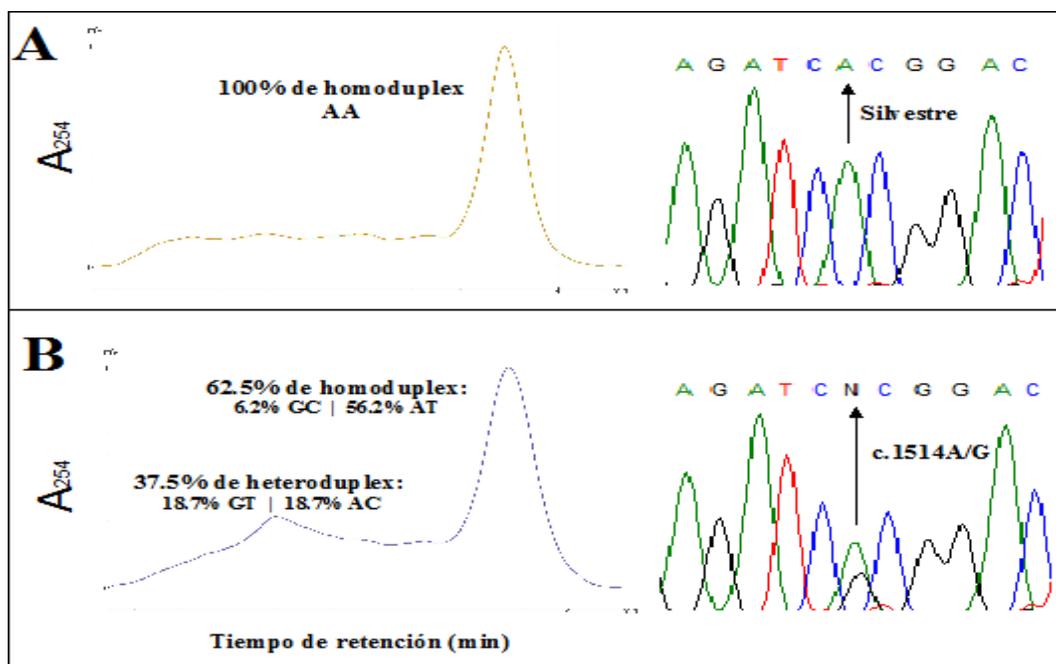
#### 7.4.7 Exón 22

En los cromatogramas de la población de estudio del exón del fragmento 22 fue identificado un patrón de pico heteroduplex que corresponde a una variante alélica para un SNP (fig. 27B). El control-molde utilizado en la hibridación resultó silvestre (fig. 27A).

En la figura 27A se muestra la representación de un patrón de pico homoduplex encontrado en 9 pacientes y 39 controles. A partir de la  $T_m$  de 54.6 °C predicha por el DHPLC, el tiempo de retención para este pico homoduplex fue de 3.95 min a una tasa de elución por parte del buffer B de 61.7%. El resultado de secuencia señala la región de interés donde se localiza la base donde fue encontrado el único cambio para este fragmento (fig. 27B), que además refiere al alélo silvestre (es decir 100% homoduplex AA); sin embargo se considera al nucleótido G -que es la variante alternativa de la considerada silvestre- como el alélo ancestral (<http://www.ensembl.org/index.html>, 2011). Por el momento no se cuenta con registros genotípicos en las poblaciones para el alélo en cuestión. Este alélo silvestre en el RNAm conjetura generalmente el aa His en el dominio colagenoso y su papel funcional estructural dentro de la cadena  $\alpha 3(IV)$  se considera similar al SNP c.1385G/A, el cual fue descrito para el exón 21.

En la figura 27B se representan los cromatogramas con un patrón de picos heteroduplex encontrados en los 4 pacientes y 11 controles restantes. El pico homoduplex mantuvo el tiempo de retención de 3.95 min, sin embargo el pico heteroduplex fue detectado en un tiempo de retención de 3.58 min cuando el buffer B alcanzó una tasa de elución de 62.2%. De acuerdo al resultado de secuenciación se deduce el siguiente orden de estabilidad respecto a la composición de los supuestos híbridos formados: GC > AT > GT > AC; los cuales se atribuyen en la proporción de un 62.5% para el pico homoduplex distribuido en 6.2% de GC y 56.2% de AT, por lado el pico heteroduplex refleja el 37.5% dividido en 18.7% de GT y 18.7% de AC; estas

proporciones definen la concentración de ambos picos, además de que los cuatro posibles híbridos establecen apareamientos purina-pirimidina (dos de estos incorrectos) (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). El SNP que explica la razón del cromatograma es el c.1514A/G (rs11677877), y consiste en una variante alélica de transición, hallada en forma heterocigota; aunque no se tienen registros de este SNP en grupos étnicos, se encuentra ampliamente reportado en la población de Eslovenia (Stabuc *et al*, 2009). Esta variante alélica se ubica en la segunda posición del codón en el mRNA y posee la capacidad para codificar tanto para His como para Arg (H451R); por lo tanto resulta en el mismo motivo funcional de dicha variante que la localizada en el exón 21.

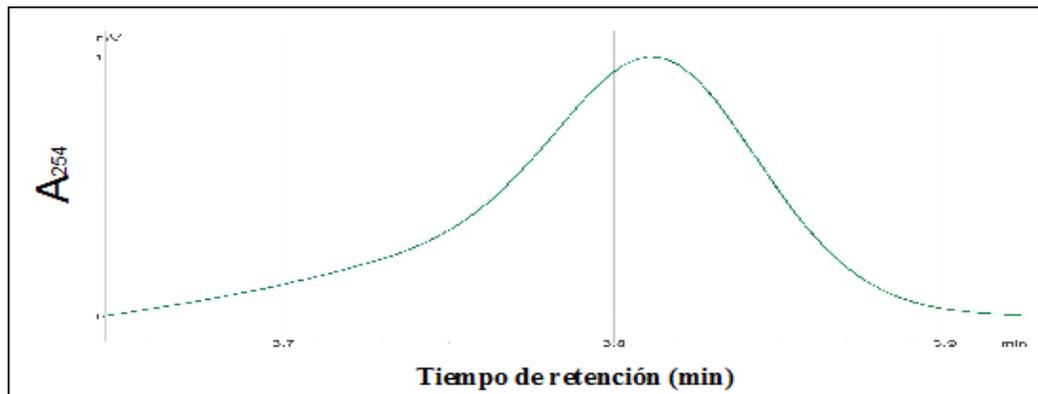


**Figura 27.** Identificación de la variante alélica polimórfica c.1514A/G en el exón 22, en estado heterocigoto. **A.** Control-molde que muestra tanto integridad en el pico homoduplex como en su respectiva secuencia. **B.** Motivo del cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.1514A/G, el cual es puntualizado en su respectiva secuencia.

### 7.4.8 Fragmento 23

En los cromatogramas de la población de estudio del fragmento 23, no se detecto ningún pico heteroduplex que indicara alguna variante alélica, sin embargo únicamente 2 controles fueron observados con ligera distorsión en su pico homoduplex, pero la determinación de sus secuencias no fueron parte de los objetivos de este estudio. El control-molde utilizado en la hibridación fue confirmado como silvestre debido al resultado que arrojó su secuencia (no se muestra).

En la figura 28 se observa un cromatograma que muestra un pico homoduplex que de manera similar también fue hallado en toda la población de estudio. Con una  $T_m$  de 56.7°C, el tiempo de retención fue de 3.81 cuando el buffer B alcanzó una tasa de elución de 59.75%.

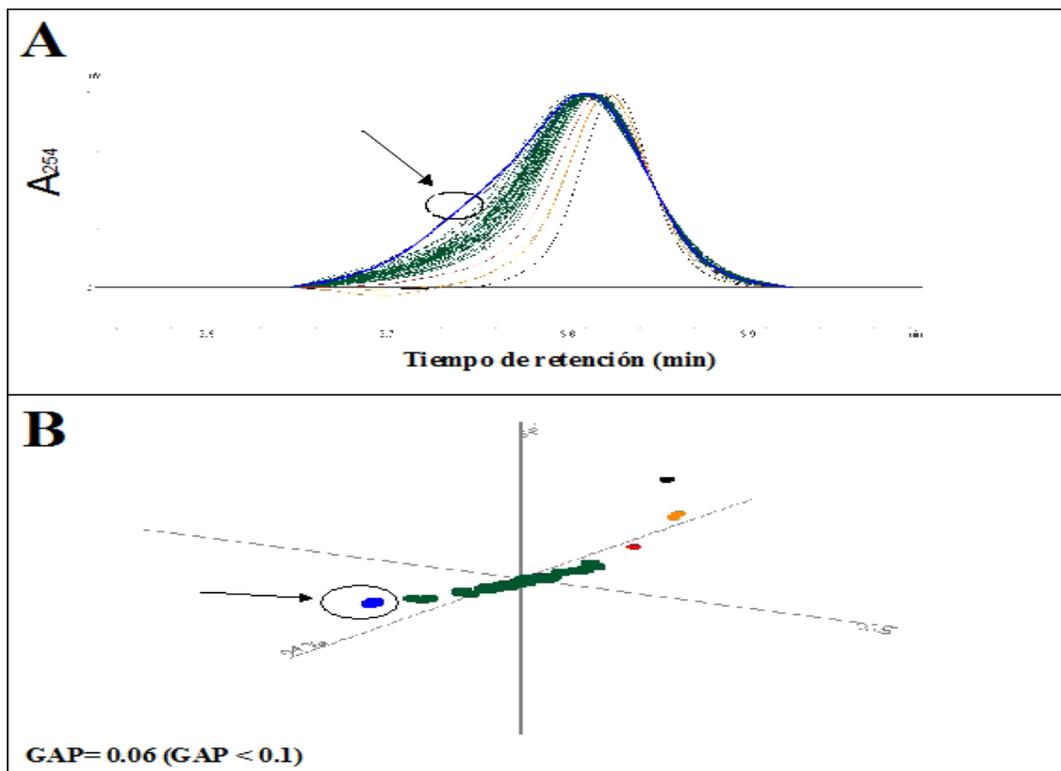


**Figura 28.** Cromatograma con pico homoduplex que de manera similar fue hallado en la totalidad de la población de estudio.

En la figura 28A se observa un cromatograma con la varianza de los cuadrados de la población de estudio normalizada, el cual en promedio conserva 3.8 min de tiempo de retención. En dicho cromatograma puede apreciarse la discrepancia de dos controles (en azul), sin embargo al igual que los cromatogramas del fragmento UTR 5'-exón 1, la distorsión que presentan dichos controles también es atribuida a fluctuaciones térmicas del DHPLC, indicando la integración de un 100% de los productos de ADN híbridos,

por lo tanto la posibilidad de un cambio verdadero en su secuencia es nulo o muy poco probable (Oefner and Underhill, 1998 y Xiao and Oefner, 2001).

En la figura 28B se observa el análisis de dispersión de los resultados cromatográficos mediante Mutation Calling, donde se muestran cinco grupos definidos distribuidos a lo largo del eje que explica el 94% del componente principal. Cuatro de dichos grupos se encuentran estrechamente dispersos de la gran mayoría de la población, los cuales incluyen los 2 controles con pico homoduplex anormal antes mencionados (en azul); como este último se acerca ligeramente al eje que representa el 5% del segundo componente principal, es posible que el DHPLC haya considerado dichos controles con cambios de otra naturaleza, la cual enfatiza a dichas oscilaciones térmicas (Schollen *et al*, 2005). De acuerdo a lo establecido, la herramienta de análisis por Mutation Calling apoya la deducción de similitud de las secuencias en este fragmento; el cual registra un polimorfismo sin registros genotípicos en las poblaciones y una mutación reportada en la población de Eslovenia (Slaipah *et al*, 2007).



**Figura 28.** Identificación de los controles ligeramente disimiles respecto a la población de estudio (señalados con una flecha). **A.** Normalización de los cromatogramas que muestra un patrón ligeramente distinto entre picos homoduplex de los 2 controles. **B.** Localización de dichos controles mediante análisis por Mutation Calling, los cuales se aprecian distribuidos a lo largo del eje del componente principal, el cual es estadísticamente significativo con un GAP = 0.06.

#### 7.4.9 Exón 52 y fragmentos UTR 3' A-I

En los cromatograma y sus respectivas secuencias de la población de estudio del exón 52 no se detecto ninguna variante alélica (no se muestra); sin embargo esto no ocurrió así para la región UTR 3', ya que fueron identificados ocho variantes alélicas (fig. 29A-G), tres establecen nuevos registros, es decir son variantes polimórficas de *novo* (fig. 29A, B y E), debido a que también fueron encontradas en controles. El control-molde utilizado en la hibridación resultó con las ocho variantes en mención (fig. 29A-G), dichos polimorfismos se distribuyen a lo largo de toda la región UTR 3'.

De acuerdo a la integridad de las secuencias del exón 52 en toda la población de estudio y considerando que dicho exón es el segmento final del gen codificante, siendo además uno de los cinco exones que componen el dominio NC1 y que contribuyen principalmente con dos codones de relativa significancia que codifican para el aa Cys y que en conjunto las propiedades de estos aa permiten configurar la estructura globular en el extremo de la cadena  $\alpha 3(\text{IV})$  (Mariyama *et al*, 1994), se establece de forma crítica que el papel funcional de dicho exón permitiría formar puentes disulfuro con otras Cys de las cadenas  $\alpha 4(\text{IV})$  ó  $\alpha 5(\text{IV})$  (mediante la oxidación de los dos grupos tioles) (Voet *et al*, 2006 y Khoshnoodi *et al*, 2008), logrando una mayor  $\Delta G^\circ$  necesaria para aportar mayor estabilidad al protómero cuando los dos residuos Cys del exón 52 y los otros diez aa Cys codificados por los restantes exónes del dominio NC1, interaccionen entre sí a lo largo del dominio NC1 (Shoulders and Raines, 2009 y Agtmael and Tuderman, 2010); sin embargo dicha funcionalidad no puede esclarecerse para la población de estudio, porque hasta el momento no se han evaluado los demás exónes que complementan el

dominio NC1, por lo tanto se desconoce alguna modificación que pudiera estar presente en toda la región NC1, sobre todo en las secuencias que corresponden a los residuos de Cys altamente conservados, para saber con mayor aproximación que tan estable sería el dominio globular NC1 si los residuos de Cys estuviesen alterados (Stevens and Lower, 2006), la posibilidad surge a partir de que se han reportado tres variantes alélicas en los sitios donde se localiza el codón que codifica para dichos aa Cys: dos polimorfismos (sinónimo y no sinónimo) y una mutación (delección localizada en pacientes con SA) en el exónes 48 y 49, respectivamente (<http://www.ensembl.org/index.html>, 2011 y Heidt *et al*, 2001); aunque por otro lado, de no existir alteración alguna de los aa Cys, proporcionarían mayor rigidez y menor flexibilidad cuando el protómero es incorporado a la MBG, debido a que dicho sitio se considera como punto de contacto de uniones covalentes para la formación del dominio globular entre las tres cadenas  $\alpha3-\alpha4-\alpha5(IV)$ , donde mediante segundas interacciones con otros protómeros, establecerían enlaces covalentes con otros protómeros idénticos para formar dímeros colagenosos (Kobayashi and Uchiyama, 2003 y Hudson, 2004); logrando de este modo una estructura más rígida y por lo tanto más resistente contra las constantes presiones hidrostáticas y coloidsmóticas ocurridas en el filtrado glomerular (Guyton and Hall, 2009).

Respecto a los resultados cromatográficos de las ocho variantes detectadas en la región UTR 3', la distribución y proporciones asignadas de las variantes puntualmente identificadas por el resultado de sus secuencias, para los picos heteroduplex visualizados en las figuras 29A-E y G, remontan la misma deducción que la utilizada para los cromatogramas de los exónes anteriormente descritos.

En la figura 29A se muestra un cromatograma con pico heteroduplex que de manera similar fue encontrado en 2 pacientes y 15 controles. El comportamiento de este cromatogramas se relaciona con una variante alélica de *nov*, la cual fue encontrada en

forma heterocigota y consiste en una transición c.5352A/G; por otro lado también se identificó otra variante, la transversión c.5490A>C (rs2070735), este cambio no pudo apreciarse en el cromatograma debido a que toda la población de estudio fue portadora y en forma homocigota, por lo tanto está incluido en el pico homoduplex de dicho cromatograma, hasta el momento no se cuentan con registros genotípicos para dicha variante. En el resultado cromatográfico, de acuerdo a la  $T_m$  de 59.2 °C predicha para el horno de la columna del DHPLC, los picos homo y heteroduplex registraron un tiempo de retención de 3.75 y 3.49 min a una tasa de elución de buffer B de 65.95 y 65.48% respectivamente; el resultado de las secuencias permitieron deducir el siguiente orden de estabilidad -esto de acuerdo a los híbridos de ADN resultantes: GC > AT > GT > AC (Oefner and Underhill, 1998 y Xiao and Oefner, 2001).

En la figura 29B se observa un cromatograma con tres picos heteroduplex bien definidos, que de forma similar también fueron observados en 6 pacientes y 9 controles. La variante alélica que resuelve este patrón de picos visualizados en el cromatograma, es porque en sus secuencias se identificó una doble variante de *novó*, la cual consta tanto de una transición como de una transversión (ambas en su forma heterocigota) c.5436T/A>C; es decir, ambos alélos son alternos del aquel considerado como silvestre. Por otro lado, a partir de una  $T_m$  de 57.5 °C predicha por el DHPLC, el tiempo de retención para el pico homoduplex fue de 3.71 min a una tasa de elución por parte del buffer B de 64.97%, mientras que los picos heteroduplex registraron un tiempo de retención de 3.45, 3.21 y 3.07 min a una tasa de elución de 64.97, 64.51 y 64.07%, respectivamente. Mediante el resultado de sus secuencias se deduce el orden de estabilidad de TA y AT > TT > AA; es importante mencionar que aunque los dos primeros híbridos TA ó AT sean los mismos apareamientos comunes, al momento de que ocurrió la realineación de los amplicones de ADN, la combinación pudo ser TA ó AT con TT ó AA y como resultado define las posiciones de los tres picos heteroduplex

(Oefner and Underhill, 1998 y Xiao and Oefner, 2001). Es importante resaltar que dicha variante fue la única detectada con dos alélos alternativos de *novο* (T y A).

En la figura 29C se presenta un cromatograma con un patrón de pico heteroduplex definido, el cual fue encontrado en un paciente y 8 controles. La razón del pico heteroduplex fue debido a que en las secuencias se encontró el SNP c.5461C/T (rs6436677), este consiste en una transición en forma heterocigota que cuenta con registros genotípicos en todos los grupos étnicos; se considera a la citosina como el alélo ancestral. Tomando en cuenta la  $T_m$  de 59.2 °C predicha por el DHPLC, los picos homo y heteroduplex fueron detectados a los 3.61 y 3.18 min de tiempo de retención a una tasa de elución de 65.69 y 64.92%, respectivamente. Finalmente en el cromatograma, el orden de estabilidad propuesto de acuerdo al resultado de las secuencias fue CG > TA > TG > CA (Oefner and Underhill, 1998 y Xiao and Oefner, 2001).

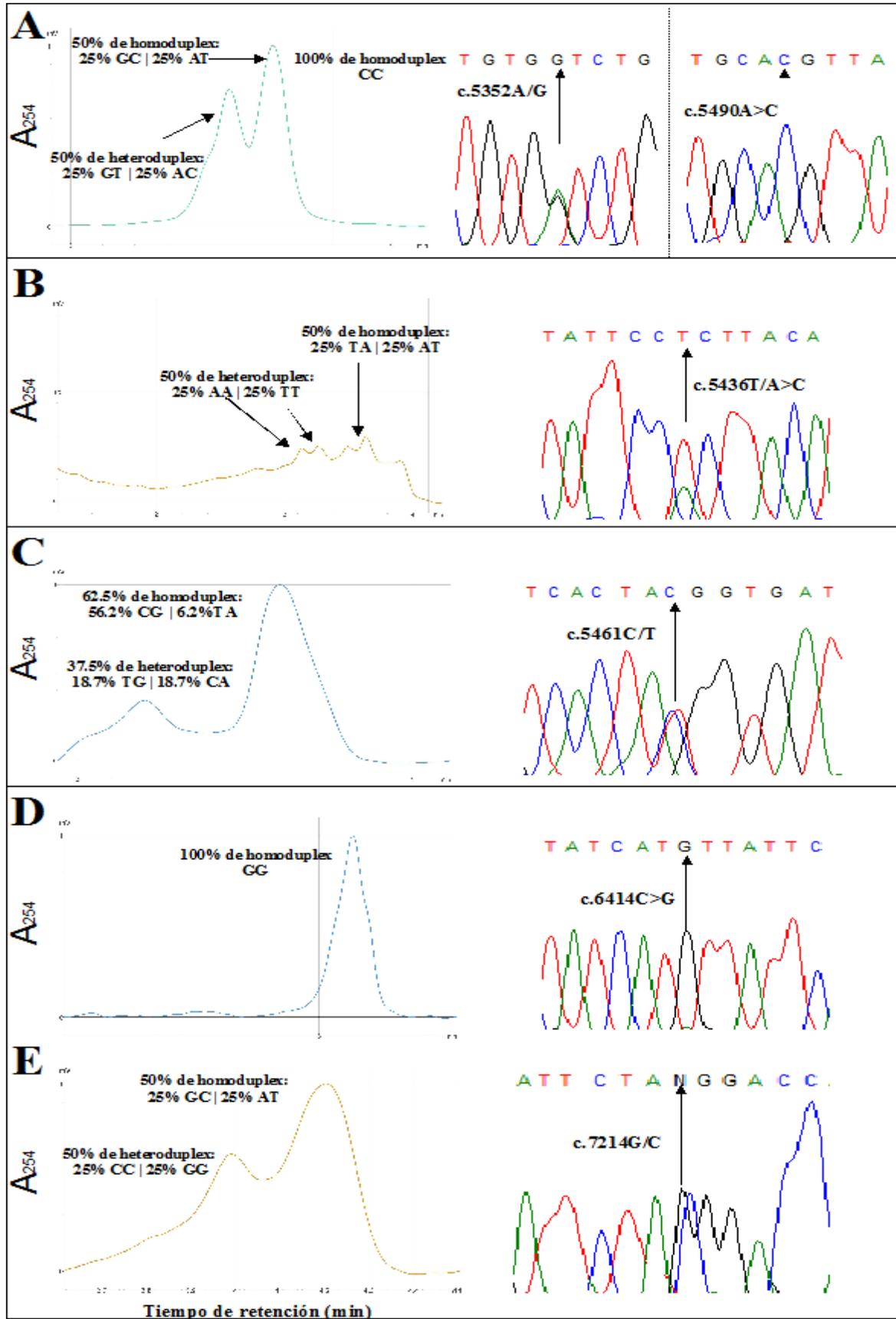
En la figura 29D se observa un pico homoduplex que fue encontrado con el mismo patrón en 8 pacientes y 19 controles. El motivo de tal pico homoduplex corresponde a la identificación del SNP c.6414C>G (rs10188531), el cual consiste en una transversión que está ampliamente reportada en todas las poblaciones, esto debido a que la base guanina se considera el alélo ancestral. La razón del pico homoduplex sostiene que se debe a que el resultado de la secuencia mostró la variante en su forma homocigota (Oefner and Underhill, 1998 y Xiao and Oefner, 2001). Conforme a la  $T_m$  de 53.8 °C predicha por el DHPLC, se obtuvo un tiempo de retención de 3.18 min cuando el buffer B alcanzó una tasa de elución del 64.72%.

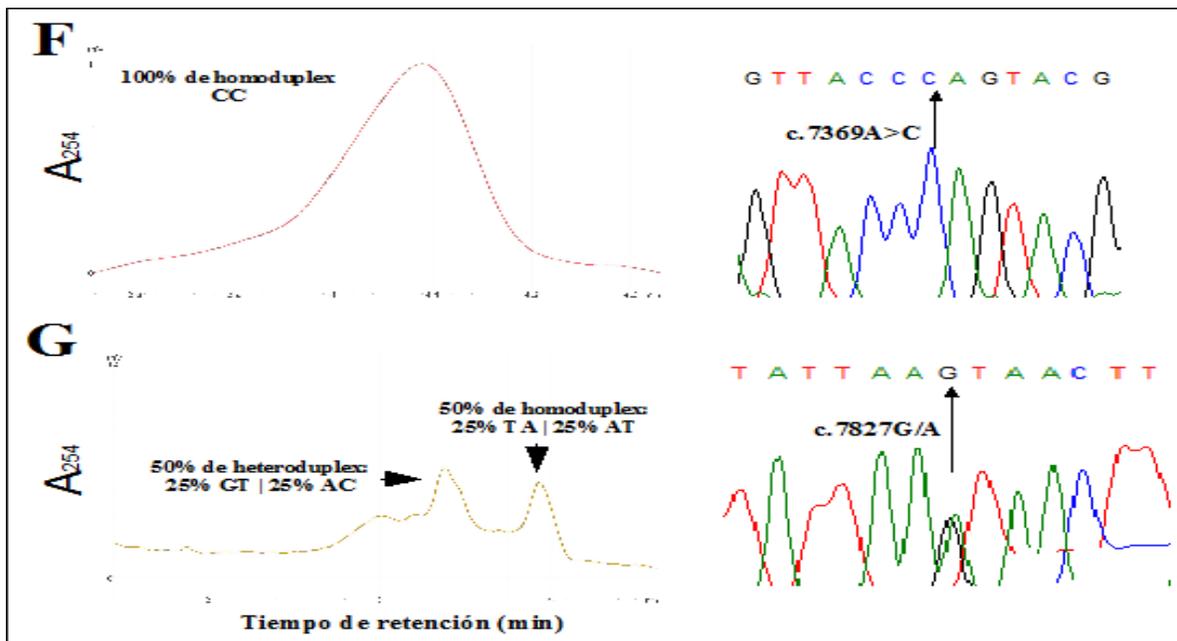
En la figura 29E se muestra un pico heteroduplex que de manera similar fue observado con el mismo patrón solamente en 6 controles. El comportamiento de este cromatograma se debe a que en las secuencias se identificó una variante alélica de *novο* c.7214G/C, esta consiste en una transversión, la cual fue encontrada en su forma heterocigota. De acuerdo a la  $T_m$  de 55.4 °C predicha por el DHPLC, el tiempo de

retención registrado para los picos homo y heteroduplex fueron de 4.2 y 3.97 min a una tasa de elución de 7.56 y 7.14%, respectivamente. El orden de estabilidad propuesto de los híbridos de ADN es GC > AT > GG y CC (Oefner and Underhill, 1998 y Xiao and Oefner, 2001).

En la figura 29F se representa un cromatograma con pico homoduplex que fue observado en toda la población de estudio. Mediante secuenciación se demostró que el pico homoduplex portaba el SNP c.7369A/C (rs7567291), el cual consiste en una transversión hallada en forma heterocigota. Entre todos las variantes alélicas encontradas en este estudio, esta es la más ampliamente reportada en todos los grupos étnicos y en este estudio no fue la excepción, además se ha establecido que la citosina es el alélo ancestral (<http://www.ensembl.org/index.html>, 2011). Respecto a la  $T_m$  de 55.7 °C predicha por el DHPLC, logró registrarse el pico homoduplex en un tiempo de retención de 4.09 min a una tasa de elución de 65.86%.

Por último, en la figura 29G se muestra un pico heteroduplex que fue encontrado de manera similar en 3 pacientes y 18 controles. La razón del pico heteroduplex fue establecido porque el resultado de secuenciación arrojó el SNP c.7827G/A (rs57817160); este consiste en una transición encontrada en forma heterocigota, la cual aún no cuenta con registros genotípicos en las poblaciones, sin embargo se establece la base guanina como el alélo ancestral. A partir de la  $T_m$  de 52.7 °C predicha por el DHPLC, el pico homo y heteroduplex se registraron a los 3.4 y 3.99 min cuando el buffer B alcanzó una tasa de elución del 64.72 y 65.78%, respectivamente. Para el orden de estabilidad de los híbridos se deducen el siguiente orden: TA ó AT > GT > AC (Oefner and Underhill, 1998 y Xiao and Oefner, 2001).





**Figura 35.** Identificación de ocho variantes alélicas polimórficas localizadas a lo largo de toda la región UTR 3', todos los incisos también corresponden al resultado del control-molde. **A.** Cromatograma con pico heteroduplex que muestra la distribución del polimorfismo de *novo* c.5352A/G en estado heterocigoto, así también el polimorfismo 5490A>C en estado homocigoto; ambas son definidas en su secuencia correspondiente. **B.** Cromatograma con picos heteroduplex definidos que muestra la distribución del polimorfismo de *novo* c.5436T/A>C en estado heterocigoto T/A, las cuales son cambios alternos del silvestre (C), dicha variante se muestra definida en su secuencia correspondiente. **C.** Cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.5461C/T en estado heterocigoto, la definición de dicha variante se muestra en su secuencia correspondiente. **D.** Cromatograma con pico homoduplex que corresponde al polimorfismo altamente conservado c.6414C>G en estado homocigoto, el cual es distinguido en su secuencia correspondiente. **E.** Cromatograma con pico heteroduplex que muestra la distribución del polimorfismo de *novo* c.7214G/C en estado heterocigoto, la definición de esta variante se muestra en su secuencia correspondiente. **F.** Cromatograma con pico homoduplex que corresponde al otro polimorfismo altamente conservado c.7369A>C en estado homocigoto, el cual se distingue en su respectiva secuencia. **G.** Cromatograma con pico heteroduplex que muestra la distribución del polimorfismo c.7827G/A en estado heterocigoto, dicha variante se distingue en su secuencia correspondiente.

De acuerdo a la ubicación de estas ocho variantes alélicas en la región UTR 3', en primera instancia puede apreciarse que el fenómeno final de procesamiento del ARNm podría completarse en la población de estudio, debido a que la secuencia consenso de poliadenilación (poli-A) AAUAAA -localizada río abajo- se encontró de manera íntegra;

lo cual indica que factores como los de especificidad de poliadenilación y escisión (CPSF), el estimulante de escisión (CstF), entre otras proteínas menores; son reclutadas a este sitio de señal poli-A para conducir a la escisión del ARN primero (alrededor de 30 nucleótidos corriente abajo) y a la poliadenilación después (alrededor de 200 nucleótidos de adenina añadidos por la enzima poli-A polimerasa, PAP) (Pesole *et al*, 2001; Mazumder *et al*, 2003 y Lewin, 2008); finalmente estas modificaciones en el extremo 3' terminarían por estabilizar y facilitar el transporte del ARNm maduro al RER para promover la traducción (Wilkie *et al*, 2003).

La región UTR 3' representa un área extremadamente rica en diversos mecanismos de control traduccional (Hellen and Sarnow, 2011), los cuales deben estar sujetos a interacciones altamente coordinadas con proteínas que mejoran el reclutamiento del ribosoma 40S, esto con el fin de cumplir con la demanda de recambio constitutivo de la cadena  $\alpha 3(IV)$ , requerida sobre todo en la MBG (Wilkie *et al*, 2003 y Watson *et al*, 2006); por lo que cabe la posibilidad de que los polimorfismos encontrados en esta región comprometan la eficacia traduccional de tal forma que las interacciones físicas entre los extremos 5' y 3' que conllevan a la circularización del ARNm, reduzcan la disponibilidad ribosómica 40S durante numerosas rondas de traducción, debido a posibles interacciones de baja afinidad relativa, principalmente del complejo eIF4G-proteína de unión a la cola poli-A (PABP) (Hellen and Sarnow, 2011). Esta razón antecede el grado de inhibición que pudieran presentar miARN ó ARN de interferencia (ARNi) (Pesole *et al*, 2001 y Stark *et al*, 2005), ya que pudiesen ser complementarios -total o parcialmente- sobre los sitios donde fueron encontradas las variantes alélicas en cuestión, en la región UTR 3'; sin embargo, hasta el momento no existen reportes que vinculen la funcionalidad de dichas variantes de manera específica con miARN en el gen *COL4A3*, ni que vías mecanicistas se adoptan en el control sobre la regulación de la expresión génica. A pesar de esto, los tres nuevos registros polimórficos se dispondrán

para diferenciar de sitios potencialmente implicados en alguna nefropatía o alguna otra enfermedad.

## 7.5 Predicciones PolyPhen y SIFT

En complemento al potencial impacto funcional de cada variante alélica descrita en la sección 7.4.1-9; solamente para la predicción del efecto de sustituciones encontradas en la región codificante, se utilizaron dos herramientas diferentes, Polyphen y SIFT, ambas predijeron de manera similar el posible impacto en la estructura y función de la sustitución de un aa en la cadena  $\alpha 3(\text{IV})$ .

En el cuadro 8 se visualiza que las codificaciones de p.G43R y p.L141P con puntaje superior a 1.5, tienen un efecto cuando fueron analizados con PolyPhen; pero SIFT predijo que únicamente la primera podría ser considerada potencialmente dañina, debido a que tuvo un puntaje menor de 0.05. La concordancia entre estas dos herramientas bioinformáticas sobre la evaluación de p.G43R, coincide a lo expuesto anteriormente en la sección 7.4.2, es decir que cualquier sustitución del aa Gly a lo largo del dominio colagenoso contribuye a reducir la  $\Delta G^\circ$  de la estructura proteica (Stryer *et al*, 2003 y Shoulders and Raines, 2009). Por otro lado, la discrepancia resultante de la alteración p.L141P entre las dos predicciones es esperada, porque las matrices y la naturaleza de la evaluación de los efectos dañinos están basados de manera diferente; PolyPhen predijo el efecto funcional de sustituciones mediante la determinación del nivel de conservación entre genes homólogos a través del tiempo evolutivo, las propiedades del intercambio de residuos aa, y la proximidad de la sustitución de los dominios proteicos y características estructurales dentro de la proteína; por otro lado SIFT predijo la importancia funcional de la sustitución de un aa basado en el alineamiento de secuencias proteicas altamente similares; ambas predicciones establecen sí o no un aa en la posición de interés está conservado en la familia proteica, por consiguiente puede ser indicativo de la importancia de su función

normal ó estructura de la proteína expresada; por lo tanto en este estudio se acepta que la modificación del aa pudiera resultar benéfica más que perjudicial, en primera porque el aa Pro es altamente competente para estabilizar la cadena  $\alpha 3(IV)$ , como lo mencionado en la sección 7.4.4 (Freeman y Herron, 2001), y en segunda -aunque es una aproximación- al motivo que hace referencia a su alta frecuencia encontrada en la población de estudio (sección 7.7).

El efecto alcanzado de la mutación c.307-309G/A solo pudo predecirse con PolyPhen, este considera que el cambio es posiblemente dañino con un puntaje de 1.8. De acuerdo a las consideraciones funcionales de la mutación deducidas en la sección 7.4.3, la evaluación de PolyPhen refuerza la misma complicación surgida por la substitución del residuo Gly en la cadena  $\alpha 3(IV)$ .

No todas las predicciones realizadas afectan la función de la cadena  $\alpha(IV)$ , tal es el caso de la predicción de los polimorfismos p.E162G, p.R408H y p.H451R, los cuales no representan un efecto significativo por PolyPhen y SIFT, ya que la predicción fue benigna con puntajes menores a 1.5 y tolerado con puntajes superiores a 0.05, respectivamente. La dilucidación de dichas variantes establece fluctuaciones menores y no amenazan de manera significativa la función proteica, esto conforme a lo reportado en las secciones 7.4.5-7.

**Cuadro 8.** Predicción del efecto de las sustituciones no sinónimas en la población de estudio.

Variante alélica	Localización	Efecto en la secuencia codificante	Predicción PolyPhen (puntuación)	Predicción SIFT (puntuación)
c.289G/C	Exón 2	p.G43R	Posiblemente dañino (1.800)	Dañino (0.01)
c.307-309G/A*	Exón 3	p.G49K	Posiblemente dañino (1.800)	No determinado
c.584T>C	Exón 7	p.L141P	Probablemente dañino (2.250)	Tolerado (1)
c.647A>G	Exón 9	p.E162G	Benigno (0.024)	Tolerado (0.49)
c.1385G/A	Exón 21	p.R408H	Benigno (1.426)	Tolerado (0.54)
c.1514A/G	Exón 22	p.H451R	Benigno (1.426)	Tolerado (0.62)

\* Única mutación implicada en la secuencia codificante.

## 7.6 Frecuencias genotípicas

A pesar de que aún no se cuentan con datos genotípicos completos de la población de estudio, es posible tener una aproximación de como se comportan las variantes alélicas en esta población, partiendo de una idea comparativa de las frecuencias alélicas de este estudio con las observadas en otras poblaciones.

Todas las alteraciones polimórficas de sentido erróneo (excepto la mutación en la región codificante), la única intrónica (excepto la mutación en el sitio de empalme), la sinónima y cinco de la región UTR 3', han sido publicadas en otros estudios; las restantes variantes alélicas son registros de *novo*: 2 mutaciones (una en la región exónica y la otra) y 3 en la región UTR 3'. Esta resolución podría resaltar dos aspectos en los cambios preliminares encontrados en el gen *COL4A3*; el primero deduce que la población mexicana no solo comparte algunas variantes con otras poblaciones alrededor del mundo, sino que surge la posibilidad de que la población mexicana sea una de las poblaciones con más polimorfismos en este gen; el segundo adiciona la complicación

para establecer el posible papel funcional que proporcionarían los polimorfismos de la región UTR 3' para regular la expresión génica de dicho gen.

En el cuadro 9 se muestran los porcentajes hasta ahorita establecidas de las frecuencias de la población de estudio y el promedio que establecen otras poblaciones (<http://www.ncbi.nlm.nih.gov/snp>, 2011). En primera instancia resalta la baja frecuencia de las dos mutaciones c.IVS3 -1G/A y c.307-309G/A (0.03% cada una), ya que solo se asocian a 2 pacientes, sin embargo este es un aspecto esperado debido a la naturaleza misma de dichas mutaciones, es decir su relación con la patología del SA. Respecto a los tres polimorfismos de *nov*o en la región UTR 3': c.5352A/G, c.5436T/A>C y c.7214G/C; hasta el momento registran frecuencias relativamente bajas (0.26%, 0.23% y 0.09%; respectivamente); si el estado de estos cambios resultasen neutros dentro de su papel funcional, cabe la posibilidad de que estas frecuencias se mantengan (sobre todo la variante c.5436T/A>C), aunque si su estado confiriera ser funcionalmente ventajoso, podría esperarse que las frecuencias aumentarían a través del tiempo (Freeman y Herron, 2001). 2 variantes en la región UTR 3' destacan por su absoluta conservación alélica (1%) en la población de estudio: c.5490A>C y c.7369A>C, de esta última se tiene el mismo registro del 1% de frecuencia en otras poblaciones alrededor del mundo; la alta conservación de dichas variantes podrían dar una pista acerca de si son dianas de algunos elementos regulatorios en la región UTR 3' (mencionados al final de la sección 7.4.9). De entre las frecuencias obtenidas de los polimorfismos en la región codificante, el resultado de la frecuencia de la variante c.584T>C considera un aspecto de importancia evolutiva, esto dentro del marco que sugiere la ventaja de portar dicho alélo por proporcionar mayor estabilidad a la cadena a3(IV) (mencionado en la sección 7.4.4 y 7.5), de acuerdo a la frecuencia obtenida relativamente alta (0.82%) en comparación con la frecuencia alélica de la misma variante pero en su forma heterocigota c.584T/C (0.24%), sugiere que el polimorfismo c.584T>C fue incrementándose con el paso de generaciones hasta establecer su fijación como una posible variante dominante

(Freeman y Herron, 2001); esta misma tendencia -aunque en menor medida- también es observada en la frecuencia de otras poblaciones (0.68%). Por último, las variantes c.289G/C, IVS2 + 12C/A, c.647A>G y c.647A/G, establecen valores de frecuencia cercanos a los determinados en otras poblaciones; no así para dos variantes UTR 3': c.5461C/T y c.6414C>G, ya que sus frecuencias se observan al doble de las establecidas en otras poblaciones, sin embargo como ya se menciono, son variantes que quedan por determinar con mayor precisión sus frecuencias alélicas.

**Cuadro 9.** Frecuencia parcial de las variantes alélicas identificadas en este estudio en comparación con la frecuencia promedio observada en otras poblaciones.

Variante alélica	Base de datos SNP ref. ID	Localización	Observación en pacientes (n=13)	Observación en controles (n=50)	Frecuencia en la población de estudio (%)	Frecuencia promedio en otras poblaciones (%)
c.289G/C	rs13424243	Exón 2	4	10	0.22	0.25
IVS2 + 12C/A	rs1882435	Intrón 2-3	4	17	0.33	0.40
c.IVS3 -1G/A*	En este estudio	Intrón 2-3	2	0	0.03	-
c.307-309G/A*	En este estudio	Exón 3	2	0	0.03	-
c.584T>C	rs10178458	Exón 7	11	41	0.82	0.68
c.584T/C			1	6	0.11	0.24
c.647A>G	rs6436669	Exón 9	8	43	0.8	0.68
c.647A/G			5	7	0.19	0.23
c.1357C>T	rs10205042	Exón 21	8	32	0.63	No determinado
c.1357C/T			4	18	0.34	
c.1385G/A	rs34505188	Exón 21	3	24	0.42	No determinado
c.1514A/G	rs11677877	Exón 22	4	11	0.23	No determinado
c.5352A/G	En este estudio	UTR 3'	2	15	0.26	-
c.5490A>C	rs2070735	UTR 3'	13	50	1.0	No determinado

c.5436T/A>C	En este estudio	UTR 3'	6	9	0.23	-
c.5461C/T	rs6436677	UTR 3'	1	8	0.14	0.33
c.6414C>G	rs10188531	UTR 3'	8	19	0.42	0.88
c.7214G/C	En este estudio	UTR 3'	0	6	0.09	-
c.7369A>C	rs7567291	UTR 3'	13	50	1.0	1.0
c.7827G/A	rs57817160	UTR 3'	3	18	0.33	No determinado

\* Mutaciones relacionadas con SA.

## 8 Conclusión

El presente estudio corresponde al primer análisis molecular enriquecido con referencias bioquímicas y fisiológicas acerca del posible comportamiento que generan las variantes alélicas del gen *COL4A3* en la proteína resultante, la cadena  $\alpha 3(\text{IV})$ , así como en su función en la MBG, principalmente en pacientes con SA. Además de reportar tres polimorfismos encontrados de *novo* en la región UTR 3', el trabajo incluye las posibles implicaciones que tiene esta región en los mecanismos transcripcionales y postranscripcionales.

El estatus de los polimorfismos en controles correspondió con el de los pacientes, excepto uno localizado en la región UTR 3'. Esta variabilidad se estima que puede alcanzar el 60% en la población mexicana, de este modo puede que esta población sea una de las más polimórficas.

Solamente 2 mutaciones idénticas en dos pacientes distintos fueron asociadas a SA, esto concuerda con la estadística de incidencia preestablecida del 20% para el SA autosómico. La mutación que recae sobre el sitio de empalme, al parecer resulta letal para la totalidad del producto génico, por otro lado la mutación implicada en la sustitución de una Gly resulta potencialmente dañina. Sin embargo, se especula que el efecto de ambas mutaciones sobre la función de la MBG podría ocasionar de manera progresiva el debilitamiento de la estructura del tejido de filtración glomerular, culminando así en deterioro sistémico renal.

De los seis polimorfismos identificados en la región codificante, existe la posibilidad de que aquel que sustituye un residuo Gly podría contribuir al estado progresivo de SA, lo cual podría servir de base para posteriores estudios sobre genética molecular de dicha enfermedad. Por el contrario, el cambio que involucra un residuo Pro al parecer podría

ser ventajoso para la proteína. Los restantes cuatro polimorfismos, al parecer, no tienen un impacto fenotípico relevante.

Los ocho polimorfismos encontrados en la región UTR 3 definen una alta tasa polimórfica, de este modo se especula que el gen *COL4A3* podría conducir a una mayor complejidad en el mecanismo de regulación de la expresión génica. Las frecuencias absolutamente conservadas de dos polimorfismos en particular (previamente reportados), podrían ser diana de una función en específica.

Finalmente, cabe resaltar que las herramientas bioinformáticas, como SIFT y PolyPhen, son de gran importancia para predecir posibles efectos de los cambios en las secuencias estudiadas, lo que permite dar un impacto fenotípico más preciso.

## 9 Perspectivas

A pesar de que se tiene una estimación confiable de que la población mexicana presenta alrededor del 60% de variantes en el *locus COL4A3*, la complementación del 20% faltante, establecerá ya sea, más variantes o mayores frecuencias de las mismas; en esta última podría realizarse un análisis de equilibrio de Hardy-Weinberg, con la finalidad de establecer las variantes alélicas que pueden multiplicarse por sus frecuencias genotípicas de las que no lo pueden hacer, esta última determinada con otra prueba más específica, indicando de esta forma la fuerza evolutiva que es responsable de dicha variabilidad; cabe relatar que de acuerdo a las mutaciones presentadas en este estudio se sospecharía de depresión consanguínea. Una vez establecida la condición, podría utilizarse como una herramienta más precisa en el asesoramiento genético de los pacientes afectados.

Las deducciones y predicciones del comportamiento de las variantes alélicas tendrían mayor soporte y serían enriquecidas cuando se disponga de la configuración tridimensional de la cadena  $\alpha 3(\text{IV})$  y del protómero  $\alpha 3\text{--}\alpha 4\text{--}\alpha 5(\text{IV})$ . Así también, como de estudios de expresión, particularmente sobre la región UTR 3'.

La utilización del DHPLC para la búsqueda de SNPs la convierte en una herramienta bastante útil para detectar cambios rápidos a partir de un número grande de muestras. Por lo que alto rendimiento de dicha técnica podría aprovecharse para establecer diagnósticos de rutina de difícil obtención (análisis, costo y tiempo); por otro lado, sería además también de gran ayuda en asesoramiento genético, más en enfermedades hereditarias, incluidos los casos donde existe depresión consanguínea, como presentan algunos casos de SA.

A partir del conocimiento del comportamiento mutacional del SA en las poblaciones, pueden llevarse a cabo distintos estudios en diversas áreas (ingeniería

genética, farmacogenómica, bioquímicos, celulares, etc). La razón, obtener alternativas terapéuticas para una mejor calidad y esperanza de vida de los afectados. Una de las investigaciones en puerta para el SA establece la incorporación de células madre para que sean diferenciadas en células podocíticas y por lo tanto puedan biosintetizar de manera íntegra el protómero  $\alpha 3-\alpha 4-\alpha 5(IV)$ . Otras simplemente sugieren la incorporación de plásmidos para reemplazar solamente las regiones alteradas. Por lo tanto, cualquiera de las formas de SA (SALX, SAAR y SAAD), podría ser indistinta en el caso de la terapia con células madre, incluso con mutaciones de *novο*; sin embargo no así para el uso de plásmidos como alternativa terapéutica.

## 10 Bibliografía

- Abrahamson, D., Hudson, B., Stroganova, L., Borza, D. y John, P. 2009. Cellular Origins of Type IV Collagen Network in Developing Glomeruli. *Journal of the American Society of Nephrology*. **20**:1471-1479.
- Agtmael, V. y Tuderman, B. 2010. Basement membranes and human disease. *Cell and Tissue Research*. **339**:167-188.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. y Walter, P. 2008. Molecular Biology of the Cell. 5a Edition. Garland Science Taylor & Francis Group. 1164-1669 pp.
- Antonarakis, E. y Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. *Human Mutation*. **11**:1-3.
- Applied Biosystems. 2002. BigDye® Terminator v3.1 Cycle Sequencing Kit Protocol. 72 pp.
- Avendaño, H., Aljama, P., Arias, M., Caramelo, C., Egado, J. y Lamas, S. 2006. Nefrología Clínica. 2a Edición. Editorial Médica Panamericana. 534-538 pp.
- Badenas, C., Praga, M., Tazón, B., Heidet, L., Arrondel, C., Armengol, A., Andrés, A., Morales, E., Camacho, J. Lens, X., Dávila, S., Milá, M., Antignac, C., Darnell, A. y Torra, R. 2002. Mutations in the *COL4A4* and *COL4A3* Genes Cause Familial Benign Hematuria. *Journal of the American Society of Nephrology*. **13**:1248-1254.
- Boutaud, A., Bogdan, B., Bondar, O., Gunwar, S., Olaf, N., Singh, N., Ninomiyas, Y., Sado, Y., Noelken, E. y Hudson, B. 2000. Type IV Collagen of the Glomerular Basement Membrane. *The Journal of Biological Chemistry*. **275**(39):30716-30724.
- Buzza, M., Wilson, D. y Savage, J. 2001. Segregation of hematuria in thin basement membrane disease with haplotypes at the loci Alport syndrome. *Kidney International*. **59**:1670-1676.
- Castagnino, J. 2000. Electroforesis capilar. Acta Bioquímica Clínica Latinoamericana. **33**(3):297-329.
- Colley, J., Jones, S., Dallosso, A., Maynard, J., Humphreys, V., Dolwani, S., Sampson, J. y Cheadle, J. 2005. Rapid Recognition of Aberrant dHPLC Elution Profiles Using the Transgenomic Navigator™ Software. *Human Mutation*. #825.
- Colville, D. y Savage, J. 1997. Alport syndrome: a review of the ocular manifestations. *Ophthalmic Genetics*. **18**:161-173.

- Cook, C., Friedrich, C. y Baliga, R. 2008. Novel *COL4A3* Mutations in African American Siblings with Autosomal Recessive Alport Syndrome. *American Journal of Kidney Diseases*. **51**(5):25-28.
- Deltas, C. 2009. Thin basement membrane nephropathy: is there genetic predisposition to more severe disease? *Pedriatic Nephrology*. **242**:877-879.
- Ding, J., Stitzel, J., Berry, P., Hawkins, E, y Kashtan C. 1995. Autosomal recessive Alport syndrome: Mutation in the *COL4A3* gene in a woman with Alport syndrome and posttransplant antiglomerular basement membrane nephritis. *Journal of the American Society of Nephrology*. **5**:1714-1717.
- Ensembl Genome Browse. 2011. *COL4A3* Human Gen. [www.ensembl.org/index.html](http://www.ensembl.org/index.html).
- European Bioinformatics Institute. 2011. Expresión, estructura y literatura del gen *COL4A3*. [www.ebi.ac.uk/](http://www.ebi.ac.uk/).
- ExoSAP-IT, 2009. Protocolo para limpiar los reactivos remanentes desde amplicones. <http://www.iwai-chem.co.jp/products/usb/exosap-it>.
- Fairbrother, W., Yeo, G., Yeh, R., Goldstein, P., Mawson, M., Sharp, P. y Burge, C. 2004. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Research*. **32**:187-190.
- Feingold, J., Bois, E., Chompret, A., Broyer, M., Gubler, C. y Grunfeld, P. 1985. Genetic heterogeneity of Alport syndrome. *Kidney International*. **27**:672-677.
- Feng, L., Xia, Y. y Wilson, C. 1994. Alternative Splicing of the NC1 domain of the human  $\alpha 3$ (IV) collagen gene. *The Journal of Biological Chemistry*. **269**(3):2342-2348.
- Freeman, S. y Herron, J. 2001. Análisis Evolutivo. 2A Ed. *Prentice Hall*. 109-188 pp.
- Fuentes, A., Castiñeiras, L. y Queraltó, C. 1998. Bioquímica Clínica y Patología Molecular. Reverté. S. A. Vol. I. 2a Ed. 177-182 pp.
- Garcia, D. y Gordillo, P. 1981. Investigación epidemiológica en cuatro núcleos familiares con síndrome de Alport. Boletín Médico del Hospital Infantil de México Federico Gómez. **38**:887-902.
- Gordillo, P. 1996. Nefritis hereditaria familiar. Nefrología Pediátrica. Mosby/Doyma. 123-127 pp.

- Gubler, M. 2008. Inherited diseases of the glomerular basement membrane. *Nature Clinical Practice Nephrology*. **4**(1):24-37.
- Guyton, A. y Hall, J. 2009. Tratado de fisiología medica. 11a Ed. *Elsevier Saunders*. 307-325 pp.
- Haas, M. 2009. Alport Syndrome and Thin Glomerular Basement Membrane Nephropathy: A Practical Approach to Diagnosis. *Archives of Pathology & Laboratory Medicine*. **133**:224-232.
- Heidet, L., Arrondel, C., Forestier, L., Solal, C., Mollet, G., Gutiérrez, B., Stavrou, C., Gubler, M. y Antignac, C. 2001. Structure of the Human Type IV Collagen Gene *COL4A3* and Mutations in Autosomal Alport Syndrome. *Journal of the American Society of Nephrology*. **12**:97-106.
- Hellen, C. y Sarnow, P. 2011. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes & Development*. **15**:1593-1612.
- Hudson, B. 2004. The molecular basis of Goodpasture and Alport syndromes: Beacons for the discovery of the collagen IV family. *Journal of the American Society of Nephrology*. **15**:2514-2527.
- Iniesta, R., Guino, E. y Moreno, E. 2005. Análisis estadístico de polimorfismos genéticos en estudios epidemiológicos. *Gaceta Sanitaria del Instituto Catalán de Oncología*. **19**(4):333-341.
- Jais, J., Knebelmann, B., Giatras, I., De Marchi, M., Rizzoni, G., Renieri, A., Weber, M., Gross, O., Netzer, K., Flinter, F., Pirson, Y., Verellen, C., Wieslander, J., Persson, U., Tryggvason, K., Martin, P., Hertz, J., Schroder, C., Sanak, M., Krejcova, S., Carvalho, M., Saus, J., Antignac, C., Smeets, H. y Gubler, M. 2000. X-linked Alport syndrome: natural history in 195 families and genotype-phenotype correlations in males. *Journal of the American Society of Nephrology*. **11**:649-657.
- Jais, J., Knebelmann, B., Giatras, I., De Marchi, M., Rizzoni, G., Renieri, A., Weber, M., Gross, O., Netzer, K., Flinter, F., Pirson, Y., Dahan, K., Wieslander, J., Persson, U., Tryggvason, K., Martin, P., Hertz, J., Schroder, C., Sanak, M., Carvalho, M., Saus, J., Antignac, C., Smeets, H. y Gubler, M. 2003. X-linked Alport syndrome: natural history and genotype-phenotype correlations in girls and women belonging to 195 families: a “European Community Alport Syndrome Concerted Action” study. *Journal of the American Society of Nephrology*. **14**:2603-2610.

- Jefferson, A., Lemmink, H., Hughes, E., Hill, M., Smeets, J., Doherty, C. y Maxwell, P. 1997. Autosomal dominant Alport syndrome linked to the type IV collagen alpha 3 and alpha 4 genes (*COL4A3* and *COL4A4*). *Nephrology Dialysis Transplantation*. **12**:1595-1599.
- Kalluri, R., Van Den Heuvel, L., Smeets, H., Schroder, C., Lemmink, H., Boutaud, A., Neilson, E. y Hudson, B. 1995. A *COL4A3* gene mutation and post-transplant anti-  $\alpha 3(\text{IV})$  collagen alloantibodies in Alport syndrome. *Kidney International*. **47**:1199-1204.
- Kashtan, C. y Michael, A. 1996. Alport syndrome. *Kidney International*. **50**:1445-1463.
- Kashtan, C. 2000. Alport syndrome: abnormalities of type IV collagen genes and proteins. *Renal Failure*. **22**(6):737-749.
- Kashtan, C. 2006. Renal transplantation in patients with Alport syndrome. *Pediatric Transplantation*. **10**:651-657.
- Kashtan, C. 2009. Familial hematuria. *Pediatric Nephrology*. **24**:1951-1958.
- Kashtan, C. 2010. Collagen IV-Related Nephropathies (Alport Syndrome and Thin Basement Membrane Nephropathy. *GeneReviews*. NCBI.
- Khoshnoodi, J., Pedchenko, V. y Hudson, B. 2008. Mammalian collagen IV. *Microscopy Research and Technique*. **71**:357-370.
- Knebelmann, B., Forestier, L., Drouot, L., Quinones, S., Chuet, C., Benessy, F., Saus, J. y Antignac, C. 1995. Splice-mediated insertion of an *Alu* sequence in the *COL4A3* mRNA causing autosomal recessive Alport syndrome. *Human Molecular Genetics*. **4**:675-679.
- Kobayashi, T. y Uchiyama, M. 2003. Characterization of assembly of recombinant type IV collagen  $\alpha 3$ ,  $\alpha 4$  and  $\alpha 5$  chains in transfected cell strains. *Kidney International*. **64**:1986-1996.
- Kriz, W. y Kaissling, B. 2008. Structural Organization of the Mammalian Kidney. *Elsevier*. **20**:479-563.
- Kuhn, K. 1994. Basement Membrane (Type IV) Collagen. *Matrix Biology*. **14**:439-445.
- Lemmink, H., Mochizuki, T., Van Den Heuvel, L., Schroder, C., Barrientos, A., Monnens, L., Van Oost, B., Brunner, H., Reedérs, S. y Smeets, H. 1994. Mutations in the type IV collagen  $\alpha 3$  (*COL4A3*) gene in autosomal recessive Alport syndrome. *Human Molecular Genetics*. **3**:1269-1273.
- Lewin, B. 2008. Genes IX. *Jones and Bartlett Publishers*. 23-700 pp.

- Longo, I., Porcedda, P., Mari, F., Giachino, D., Meloni, I., Deplano, C., Brusco, A., Boiso, M., Massella, L., Lavoratti, G., Roccatello, D., Frascá, G., Mazzucco, G., Muda, O., Conti, M., Fasciolo, F., Arrondel, C., Heidet, L., Renieri, A. y Marchi, M. 2002. *COL4A3/COL4A4* mutations: From familial hematuria to autosomal-dominant or recessive Alport syndrome. *Kidney International*. **61**:1947-1956.
- Longo, I., Scala, E., Mari, F., Caselli, R., Pescucci, C., Mencarelli, M., Speciale, C., Giani, M., Bresin, E., Caringella, D., Borochowitz, Z., Siritwaderna, K., Winship, I., Renieri, A. y Meloni, I. 2006. Autosomal recessive Alpor syndrome: an in-depth clinical and molecular analysis of five families. *Nephrology Dialysis Transplantation*. **21**:665-671.
- Lyons, O., John, E., Morales, J., Chan, Y. y Taylor P. 2007. Ruptured thoracoabdominal aortic aneurysm in a renal transplant patient with Alport's syndrome. *Annals of Vascular Surgery*. **21**:816-818.
- Mariyama, M., Leinonen, A., Mochizuki, T., Tryggvason, K. y Reeders, S. 1994. Complete primary structure of the human alpha-3(IV) collagen chain: cexpression of the alpha-3(IV) and alpha-4(IV) collagen chins in human tissues. *The Journal of Biological Chemistry*. **269**(37):23013-23017.
- Mazumder, B., Seshadri, V. y Fox, P. 2003. Translational control by the 3'-UTR: the ends specify the means. *TRENDS in Biochemical Sciences*. **28**(2):91-98.
- McCarthy, P. y Maino, D. 2000. Alport syndrome: a review. *Clinical Eye and Vision Care*. **12**(3-4):139-150.
- Medeiros, D., Fuentes, Y., García, R., Hernández, A., Morán, B. y Velásquez, J. 2008. Síndrome de Alport. Boletín Médico del Hospital Infantil de México Federico Gómez. **65**:331-340.
- Merchant, S., Burgess, B., Adams, J., Kashtan, C., Gregory, M., Santi, P., Collins, B. y Nadol, J. 2004. Temporal bone hitopathology in alport syndrome. *Laryngoscope*. **114**:1609-1618.
- Miner, J. 1999. Renal basement membrane components. *Kidney International*. **56**:2016-2024.
- Mochizuki, T., Lemmink, H., Mariyama, M., Antignac, C., Gluber, C., Pirson, Y., Doumoulin, C., Chan, B., Schroder, H., Smeets, J. y Reeders, S. 1994. Identification of mutations in the alpha-3(IV) and alpha-4(IV) collagen genes in autosomal recessive Alport syndrome. *Nature Genetics*. **8**:77-81.
- Momota, R., Sugimoto, M., Ohashi, T., Kigasawa, K., Yoshioka, H. y Ninomiya, Y. 1998. Two genes, *COL4A3* and *COL4A4* coding for the human  $\alpha$ 3(IV) and  $\alpha$ 4(IV) collagen chains

are arranged head-to-head on chromosome 2q36. Federation of European Biochemical Societies. **424**(1-2):11-6.

- Nagel, M., Nagorka, S. y Gross, O. 2005. Novel *COL4A5*, *COL4A4*, and *COL4A3* Mutations in Alport Syndrome. *Human Mutation*. **820**:1-5.
- National Center for Biotechnology Information. 2011. Descripción de técnicas actuales de biología molecular. [www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/).
- National Center for Biotechnology Information. 2011. Frecuencias genotípicas de las variantes alélicas. [www.ncbi.nlm.nih.gov/snp](http://www.ncbi.nlm.nih.gov/snp).
- Oefner, P y Underhill, P. 1998. DNA Mutation Detection Using Denaturing High-Performance Liquid Chromatography (DHPLC). *Current Protocols in Human Genetics*. 7.10.1-7.10.12 pp.
- Passwell, H., David, R., Boichis, H. y Herzfeld, S. 1981. Hereditary nephritis with associated defects in proximal renal tubular function. *Journal of Pediatrics*. **98**:85-87.
- Perez, T., Mora, M., Castillo, M. y Garcia, G. 1999. Enfermedad de Alport en niños. *Nefrología Mexicana*. **20**(4):171-174.
- Pescucci, C., Mari, F., Longo, I., Vogiatzi, P., Caselli, R., Scala, E., Abaterusso, C., Gusmano, R., Seri, M., Miglietti, N., Bresin, E. y Renieri, A. 2004. Autosomal-dominant Alport syndrome: Natural history of a disease due to *COL4A3* or *COL4A4* gene. *Kidney International*. **65**:1598-1603.
- Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F. y Liuni, S. 2001. Structural and functional features of eukaryotic mRNA untranslated regions. *Gene*. **276**:73-81.
- PolyPhen. 2011. Predicción del efecto funcional de SNPs no sinónimas en el humano. [www.genetics.bwh.harvard.edu/pph/](http://www.genetics.bwh.harvard.edu/pph/).
- Pöschl, E., Schlötzer-Schrehardt, U., Brachvogel, B., Saito, K., Ninomiya, Y. y Mayer, U. 2003. Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development*. **131**:1619-1628.
- Primer 3. 2009. Diseño de cebadores del gen *COL4A3* para fabricante Applied Biosystems. [www.frodo.wi.mit.edu/primer3/](http://www.frodo.wi.mit.edu/primer3/).
- Ramensky, V., Bork, P. y Sunyaev, S. 2002. Human non-synonymous SNPs: server and survey. *Nucleic Acids Research*. **30**(17):3894-3900.

- Rana, K., Tonna, S., Wang, Y., Sin, L., Lin, T., Shaw, E., Mookerjee, I. and Savige, J. 2007. Nine novel *COL4A3* y *COL4A4* mutations and polymorphisms identified in inherited membrane diseases. *Pediatrics Nephrology*. **22**:652-657.
- Rhys, C., Snyers, B. y Pirson, Y. 1997. Recurrent corneal erosion associated with Alport's syndrome. *Kidney International*. **52**:208-211.
- Sado, Y., Kagawa, M., Naito, I., Ueki, Y., Seki, T., Momota, R., Oohashi, T. y Ninomiya, Y. 1998. Organizatio and Expression of Basement Membrane Collagen IV Genes and Their Roles in Human Disorders. *The Journal of Biochemistry*. **123**:767-776.
- Sambrook, J. y Russell, D. (2001). Molecular Cloning a laboratory manual. Vol 1 and 2. Cold Spring Harbor Laboratory Press. 3a Edition. 5.1-12.114 pp.
- Savige, J., Rana, K., Tonna, S., Buzza, M., Dagher, H. y Wang, Y. 2003. Thin basement membrane nephropathy. *Kidney International*. **64**:1169-1178.
- Schollen, E., Dequeker, E., McQuaid, S., Vankeirbilck, B., Michils, G., Harvey, J., Van den Akker, E., Van Schooten, R., Clark, Z., Schrooten, S. y Matthijs, G. 2005. Diagnostic DHPLC Quality Assurance (DDQA): A Collaborative Approach to the Generation of Validated and Standardized Methods for DHPLC-Based Mutation Screening in Clinical Genetics Laboratories. *Human Mutation*. **25**:583-592.
- Shoulders, M. y Raines, R. 2009. Collagen Structure and Stability. *The Annual review of Biochemistry*. **78**:929-958.
- SIFT. J. Craig Venter Institute. 2011. Predicción del efecto fenotípico en variantes alélicas de sentido erróneo en la región codificadora. [www.genetics.bwh.harvard.edu/pph/](http://www.genetics.bwh.harvard.edu/pph/).
- Slajpah, M., Gorinsek, B., Berginc, Vizjak, A., Ferluga, D., Hvala, A., Meglic, A., Jaksa, I., Furlan, P., Gregoric, A., Pavlovcic, K., Glavac, R. y Glavac, D. 2007. Sixteen novel mutations identified in *COL4A3*, *COL4A4*, and *COL4A5* genes in Slovenian families with Alport syndrome and benign familial hematuria. *International Society of Nephrology*. **71**:1287-1295.
- Srinivasan, M., Uzel, S., Gautieri, A., Keten, S. y Buehler, M. 2009. Alport Syndrome mutations in type IV tropocollagen alter molecular structure and nanomechanical properties. *Elservier*. **168**:503-510.
- Stabuc, S., Ravnik, G., Glavac, D., Hawlina, M. y Strazisar, M. 2009. Polymorphisms in *COL4A3* and *COL4A4* genes associated with keratoconus. *Molecular Vision*. **15**:2848-2860.

- Stark, A., Brennecke, J., Bushati, N., Russell, R. y Cohen, S. 2005. Animal MicroRNAs Confer Robustness to Gene Expression and Have a Significant Impact on 3'UTR Evolution. *Cell*. **123**:1133-1146.
- Stevens, A. y Lowe, J. 2006. Histología Humana. 3a Edición. Elsevier. 300-303 pp.
- Strachan, T. y Read, A. 2006. Genética Humana. 3a de. *McGraw-Hill Interamericana*. 275-298 pp.
- Stryer, L., Berg, J. y Tymoczko, J. 2003. Bioquímica. 5a Ed. *Reverté, S. A.* 189-221 pp.
- Tazón, V., Célia, B., Elisabet, A., Xose, L., Montse, M., Darnell, A. y Torra, L. 2003. Autosomal Recessive Alport's Syndrome and Benign Familial Hematuria Are Collagen Type IV Diseases. *American Journal of Kidney Diseases*. **42**(5):952-959.
- Thorner, P. S. 2007. Alport Syndrome and Thin Basement Membrane Nephropathy. *Nephron Clinical Practice*. **106**:82-88.
- Tortora, G. y Derrickson, B. 2006. Principios de Anatomía y Fisiología. 11a Ed. Editorial Médica Panamericana. 999-1034 pp.
- Tryggvason, K. y Patrakka. 2006. Thin Basement Membrane Nephropathy. *Journal of the American Society of Nephrology*. **17**:813-822.
- Van de Loop, F., Heidet, L., Timmer, E., Van den Bosch B., Leinonen, A., Antignac, C., Jefferson, A., Maxwell, P., Monnens, L., Schroder C. y Smeets, H. 2000. Autosomal dominant Alport syndrome caused by a *COL4A3* splice site mutation. *Kidney International*. **58**:1870-1875.
- Voet, D., Voet, J. y Pratt, C. Fundamentos de Bioquímica. 2a Ed. Editorial Panamericana. 76-174 pp.
- Voskarides, K., Pierides, A. y Deltas, C. 2008. *COL4A3/COL4A4* Mutations Link Familial Hematuria and Focal Segmental Glomerulosclerosis. Glomerular Epithelium Destruction via Basement Membrane Thinning? *Connective Tissue Research*. **49**:283-288.
- Wang, Y., Rana, K., Tonna, S., Lin, T., Sin, L. y Savage, Judy. 2004. *COL4A3* mutations and their clinical consequences in thin basement membrane nephropathy (TBMN). *Kidney International*. **65**:786-790.
- Watson, J., Baker, T., Bell, S., Gann, A., Levine, M. y Losick, R. 2006. Biología Molecular del Gen. *Editorial Medica Panamericana*. 47-613 pp.

- Wein, J., Kavoussi, R., Novick, C., Partin, W. y Peters, A. 2007. Campbell-Walsh. Urology. *Elsevier*. 9rd Edition. 1245-1247 pp.
- Wickens, M., Bernstein, D., Kimble, J. y Parker, R. 2002. PUF family portrait: 3' UTR regulation as a way of life. *TRENDS in Genetics*. **18**(3):150-157.
- Wilkie, G., Dickson, K. y Gray, N. 2003. Regulation of mRNA translation by 5'-and 3'-UTR-binding factors. *TRENDS Biochemical Sciences*. **28**(4):182-188.
- Xiao, W. y Oefner, P. 2001. Denaturing High-Performance Liquid Chromatography: A Review. *Human Mutation*. **17**:439-474.
- Yurchenco, P., Amenta, P. y Patton, B. 2004. Basement membrane assembly, stability and activities observed through a developmental lens. *Matrix Biology*. **22**:521-538.

## 11 Anexo

Características físicas de los cebadores para las condiciones de PCR

Fragmento	Cebadores. Sentido (FW) y antisentido (RV)	Longitud (pb)*	Tm	Secuencia	Producto de amplificación (pb)
UTR 5'-exón 1	COL4A3-X1 FW	20	59.9	GGA CAC TGC CTG GTA AGT TG	546
	COL4A3-X1 RV	18	60.28	CGC ATT CCA CCT GTT CCG	
Exón 2	COL4A3-X2 FW	22	60.16	TGC CGG TTG GGA TTT ATT CAG	227
	COL4A3-X2 RV	21	59.93	AAT GGA ACA ACA TAC TGC GGA T	
Exón 3	COL4A3-X3 FW	21	59.88	ATT GTG GGA TAG TTG CTG CTC	232
	COL4A3-X3 RV	21	60.36	AAT TCT CTC CAG GTT TGC AGC	
Exón 7	COL4A3-X7 FW	19	58.02	GT TGG AAG CGA GAA ACA GG	227
	COL4A3-X7 RV	22	58.16	ATG ACG GAT TTG TGT ATA TGG C	
Exón 9	COL4A3-X9 FW	22	56.44	ACT TTG AAT AAG CAC TTG AAG G	229
	COL4A3-X9 RV	20	57.02	CGA AGT AAA GAG AGG CGA AA	
Exón 21	COL4A3-X21 FW	22	58.37	GTT ATG TAC CTC TCC ATT GTG C	387
	COL4A3-X21 RV	20	58.83	AGA CTG ATG GCT AAG CTG TG	
Exón 22	COL4A3-X22 FW	21	58.95	GTC ACC ATG CTT TCT CAG TTG	339
	COL4A3-X22 RV	22	58.95	TGG GAT CAT TGT TAT CTC AGG G	
Exón 23	COL4A3-X23 FW	23	58.27	TTG TTC TTT CTG AGG ACT CAA TG	246
	COL4A3-X23 RV	21	58.72	GTG ACA TCC ACC AAT CCA ATG	
Exón 52A- UTR 3'	COL4A3-X52A FW	20	59.48	CGC CCG ACC CTG TAA TAA AT	559
	COL4A3-X52A RV	22	59.04	GAA GAC AAA TGT GGG AAT CCA G	
52B UTR 3'	COL4A3-X52B FW	22	56.95	AG CAA TTC TTT CAA GTC AGT TC	482
	COL4A3-X52B RV	21	56.72	TTC CAA GGG TAT TAT GGT TCC	

52C UTR 3'	COL4A3-X52C FW	22	57.68	CTA AAT CAG TGT TTG ATT GCC C	550
	COL4A3-X52C RV	21	57.45	ATG GAT TAT TGA ACT GCT CCC	
52D UTR 3'	COL4A3-X52D FW	24	58.83	AGC TCT AAA ATC TGC TTG TAT	582
	COL4A3-X52D RV	22	58.44	TCC TCT CTG TAG GTT AGT ACA GTG C	
52E UTR 3'	COL4A3-X52E FW	23	57.47	TGT CAT CAC TCT AGG TTT TAC AG	537
	COL4A3-X52E RV	23	57.02	TTC AGA CCT TAT CTT CTC TTA GC	
52F UTR 3'	COL4A3-X52F FW	24	56.60	GAC TTG TTT CTG ATC TAA TAT	543
	COL4A3-X52F RV	22	56.55	CCC AAT GCT GCA CAT CTT AGT TTA G	
52G UTR 3'	COL4A3-X52G FW	20	57.15	GAG TCC AGA CAC AAA GAG AC	523
	COL4A3-X52G RV	22	57.10	AGA GTA CAT CAG AAT CAT GTG G	
52H UTR 3'	COL4A3-X52H FW	21	58.67	CTC TGT AGG CCA GTA GTT CTC	494
	COL4A3-X52H RV	21	58.48	TTA GAA GTG CAA CTG AAC AGC	
52I UTR 3'	COL4A3-X52I FW	20	54.41	ACA GGT GAC TCT AAG GTT AT	505
	COL4A3-X52I RV	20	54.21	AAG TTT AGT CGC AGT AAG TT	

\*Incluyen las regiones intrónicas adyacentes.