

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ciencias y Humanidades



Estudio comparativo y propuesta de un método molecular de bajo costo para el pronóstico de la respuesta al tratamiento de Interferón alfa y Ribavirina en pacientes con hepatitis C -Efecto de la homología IL28B-IL28A en la genotipificación del polimorfismo de un solo nucleótido rs12979860

Trabajo de graduación en modalidad de tesis presentado por  
Marie Andre Cosenza Barnéond  
para optar al grado académico de Licenciada en Bioquímica y  
Microbiología

Guatemala

2015





Estudio comparativo y propuesta de un método molecular de bajo costo para el pronóstico de la respuesta al tratamiento de Interferón alfa y Ribavirina en pacientes con hepatitis C -Efecto de la homología IL28B-IL28A en la genotipificación del polimorfismo de un solo nucleótido rs12979860

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ciencias y Humanidades




Estudio comparativo y propuesta de un método molecular de bajo costo para el pronóstico de la respuesta al tratamiento de Interferón alfa y Ribavirina en pacientes con hepatitis C -Efecto de la homología IL28B-IL28A en la genotipificación del polimorfismo de un solo nucleótido rs12979860

Trabajo de graduación en modalidad de tesis presentado por  
Marie Andre Cosenza Barnéond  
para optar al grado académico de Licenciada en Bioquímica y  
Microbiología


Guatemala


2015


Vo.Bo.:

(f)   
M.Sc. Claudia Paiz

Tribunal Examinador:

(f)   
M.Sc. Claudia Paiz

(f)   
Ph.D. Dalia Lau

(f)   
Lic. Andrés Ávalos

Fecha de Aprobación: Guatemala 13 de febrero de 2015 ✓

## PREFACIO

Al iniciar en la universidad pasé cuatro años estudiando la licenciatura en Química Farmacéutica antes de descubrir que mi verdadera pasión se encontraba en la bioquímica, microbiología y biología molecular. Por esta razón es entendible que la farmacogenética sea una rama de alto interés para mí.

Durante mis prácticas profesionales en el Instituto de Investigaciones Químicas, Biológicas, Biomédicas y Biofísicas de la Universidad Mariano Gálvez de Guatemala trabajé por primera vez con un marcador farmacogenético para el diagnóstico de las mutaciones en K-RAS, por lo que poco a poco fui entendiendo cómo funcionan estos marcadores y el impacto que pueden llegar a tener en los pacientes.

Al finalizar mis prácticas profesionales en el instituto me ofrecieron la oportunidad de realizar mi trabajo de graduación en el marcador farmacogenético IL28-B de pronóstico de respuesta viral sostenida en pacientes con Hepatitis C; sin embargo fue el impacto que puede tener la región homóloga no solo en el diagnóstico del marcador sino en el desarrollo de la enfermedad en el paciente lo que cambió el enfoque de mi trabajo a la evaluación de la literatura publicada respecto a esto para diseñar y seleccionar el método que asegure resultados confiables para implementarse en el laboratorio.

# ÍNDICE

PREFACIO.....	i
LISTA DE CUADROS .....	iii
LISTA DE FIGURAS .....	iv
RESUMEN.....	iii
I. INTRODUCCIÓN.....	1
II. OBJETIVOS.....	2
III. JUSTIFICACIÓN .....	3
IV. MARCO TEÓRICO.....	4
V. MATERIALES Y MÉTODOS.....	17
VI. RESULTADOS.....	20
VII. DISCUSIÓN .....	42
VIII. CONCLUSIONES .....	49
IX. RECOMENDACIONES.....	50
X. BIBLIOGRAFÍA.....	51
XI. ANEXOS .....	54

## LISTA DE CUADROS

	Página
1. Parámetros termodinámicos de los iniciadores diseñados	24
2. Información de las estructuras de Hairpin producidas por los iniciadores	27
3. Posibles iniciadores internos para el gPCR	30
4. Iniciadores seleccionados para el asPCR	32
5. Información de las estructuras de Hairpin de los iniciadores del asPCR	33
6. Evaluación de los costos de implementar del SNP por distintos métodos	40
7. Evaluación de los costos de implementar la detección del rs12979860 por Secuenciación.	41

## LISTA DE FIGURAS

	Página
1. Ciclo de replicación del Virus de Hepatitis C	5
2. Esquema de organización del Genoma del virus	6
3. Distribución mundial del VHC	8
4. Posibles respuestas al tratamiento convencional de VHC	9
5. Vía de activación de la IL28B	11
6. Porcentaje de SVR alcanzado en distintas poblaciones del mundo	13
7. Ilustración del mecanismo de acción de los distintos tipos de endonucleasas	15
8. Ilustración de la incorporación de los ddNTPs en la secuenciación	16
9. Esquema del cromosoma 19 y visualización de la región 19q13.13	20
10. Alineación de la 'Región SNP' contra subsecuencias de la 'Región IFN'	22
11. Alineación de la 'Región Interés' y la 'Región homóloga'	23
12. Esquema de la colocación de los iniciadores diseñados sobre la homología	24
13. Estructuras de homodímeros producidas por el Indicador Foward diseñado	25
14. Estructuras de homodímeros producidas por el Indicador Reverse diseñado	25
15. Estructuras de heterodímeros producidas por los indicadores diseñados	26
16. Resultado de asPCR, FRLP-PCR y secuenciación de los indicadores diseñados	28
17. Visualización de la colocación de los indicadores publicados	29
18. Resultados de la amplificación hipotética por asPCR de los indicadores publicados	31
19. Esquema gráfico de la colocación de los iniciadores externos de asPCR	32
20. Estructuras de homodímeros de los iniciadores del asPCR	34
21. Estructuras de heterodímeros producidas por los indicadores del asPCR	35
22. Electroforesis de los resultados del asPCR	35
23. Resultados de la restricción enzimática de los iniciadores publicados	37
24. Esquema de los amplicones producidos por los iniciadores publicados	38
25. Esquema de las regiones amplificadas por distintos iniciadores	39

## RESUMEN

La farmacogenética de la Hepatitis C es un paso clave en el tratamiento efectivo de la enfermedad, ya que permite orientar al médico respecto al régimen y fármaco apropiados para el paciente según su disposición genética y la del virus que le infecte. Por esta razón en los últimos años se han realizado varias publicaciones implementando la genotipificación del SNP rs12979860 -ligado al porcentaje de éxito del tratamiento con Interferón alfa pegilado y Ribavirina en pacientes con hepatitis C- por distintos métodos, pero no es hasta el año 2012 donde se determina por primera vez la existencia de una región homóloga a la que contiene el SNP rs12979860 y su impacto en el resultado de dicha genotipificación. El objetivo de este estudio fue realizar una evaluación de esta homología y de las publicaciones que se han producido respecto a la genotipificación de dicho SNP mediante herramientas bioinformáticas para diseñar y elegir un método de genotipificación de alta especificidad que sea conveniente para la población guatemalteca. Los resultados obtenidos evidencian que el método óptimo para la implementación de la genotipificación del SNP rs12979860 es el PCR multiplex alelo específico con los iniciadores 5'- GGAAGGAGCAGTTGCGCTG-3', 5'- CTTCACGCTCCGAGCATTGC-3', 5'- AGCTCCCCGAAGGCGTG-3' y 5'- GTGCAATTCAACCCTGGTTCGC-3. Sin embargo, es necesario evaluar este método empíricamente para establecer su eficacia como diagnóstico de la farmacogenética de la hepatitis C.



## I. INTRODUCCIÓN

En la actualidad se cree que aproximadamente el 3% de la población mundial se encuentra infectada con el virus de hepatitis C (VHC). Esto representa una gran preocupación para las instituciones responsables de velar por la salud alrededor de todo mundo ya que aproximadamente el 80% de estos pacientes desarrollan hepatitis C crónica, lo cual conlleva a otros problemas de salud como cirrosis (en 20% de los pacientes) o cáncer hepático (en 5% de los pacientes). El tratamiento tradicional para pacientes con VHC consiste en la coadministración de Interferón alfa pegilado y Ribavirina (pegIFN/RBV). Éste es efectivo en aproximadamente un 50% de los pacientes infectados con el genotipo 1 del virus y varía dependiendo a razas y etnias. Este hecho ha impulsado la investigación molecular con la finalidad de identificar las causas de esta diferencia en la respuesta al tratamiento convencional en distintos grupos étnicos. Estos estudios dieron como resultado el desarrollo de un marcador farmacogenético (rs12979860) que permite predecir cuál será la posible respuesta del paciente al tratamiento convencional.

Desde el descubrimiento de la relación del genotipo CC de rs12979860 con la respuesta favorable del paciente al tratamiento con pegIFN/RBV, diversas instituciones a nivel mundial han introducido la genotipificación de dicho marcador como prueba rutinaria en la elección del tratamiento de VHC. Sin embargo, en estudios recientes se ha evidenciado una región altamente homóloga a aquella que contiene el SNP rs12979860, por lo cual es necesario reevaluar las herramientas moleculares que se utilizan para asegurar una detección específica del marcador y evitar así falsos positivos en el diagnóstico de pacientes infectados con el virus de la hepatitis C.

## II. OBJETIVOS

### A. General

Proponer una metodología de bajo costo que permita la detección específica del SNP rs12979860 del IL28-B .

### B. Específicos

1. Delimitar la región del gen IL28-A, homóloga al gen IL28-B, utilizando herramientas bioinformáticas
2. Identificar la posición de SNP rs12979860 en la región del gen IL28-B y la región homóloga en el gen IL28-A.
3. Evaluar la especificidad de amplificación de los iniciadores publicados para la amplificación del SNP rs12979860 considerando la homología entre los genes IL28-A e IL28-B
4. Proponer las herramientas moleculares necesarias que permitan la detección del SNP rs12979860 utilizando secuenciación sanger, qPCR alelo específico, RFLP-PCR y PCR alelo específico de punto final.
5. Realizar un análisis de costos de las diferentes metodologías para establecer la más accesible para los pacientes guatemaltecos.

### III. JUSTIFICACIÓN

La genotipificación del marcador rs12979860 para predecir la respuesta favorable al tratamiento convencional de VHC es importante, ya que podría incrementar los porcentajes de éxito en el tratamiento de VHC mediante la elección del tratamiento adecuado para el paciente. Adicionalmente, la genotipificación puede aumentar los casos de éxito a largo plazo en los trasplantes de hígado, ya que permitirá seleccionar los donantes cuyos hígados tienen menor probabilidad a desarrollar cirrosis producida por hepatitis C. Finalmente, en países extranjeros, la implementación de esta prueba ha demostrado disminuir los costos anuales del tratamiento de hepatitis C en 2,916,000 euros para un único hospital.

Por estas razones la evaluación de la especificidad y el desarrollo de pruebas moleculares específicas que consideren la región homóloga es una prioridad en entidades donde se realiza el tratamiento de hepatitis C alrededor del mundo. Sin embargo, para la implementación de dicha prueba en Guatemala es necesario considerar adicionalmente que se debe establecer una metodología de bajo costo, que permita que la selección de la terapia adecuada para cada paciente se encuentre económicamente accesible a la mayoría de la población; por lo que a continuación se pretende optimizar un PCR que permita la amplificación específica del marcador y que pueda aplicarse en distintas metodologías, según se considere apropiado.

## IV. MARCO TEÓRICO

### A. VIRUS DE LA HEPATITIS C

1. Descripción y genotipos. El virus de la hepatitis C (VHC) pertenece al género Hepacivirus, de la familia Flaviridae. Este virus fue clonado por primer vez en 1989 (Choo et al., 1989), sin embargo no fue hasta 2005 cuando fue posible cultivarlo en el laboratorio mediante un cultivo celular (Wakita et al., 2005). Actualmente se estima que aproximadamente un 3% de la población mundial se encuentra infectado con dicho virus (Sharma, 2010), por lo que su estudio y control es de suma importancia para los servicios de salud a nivel mundial.

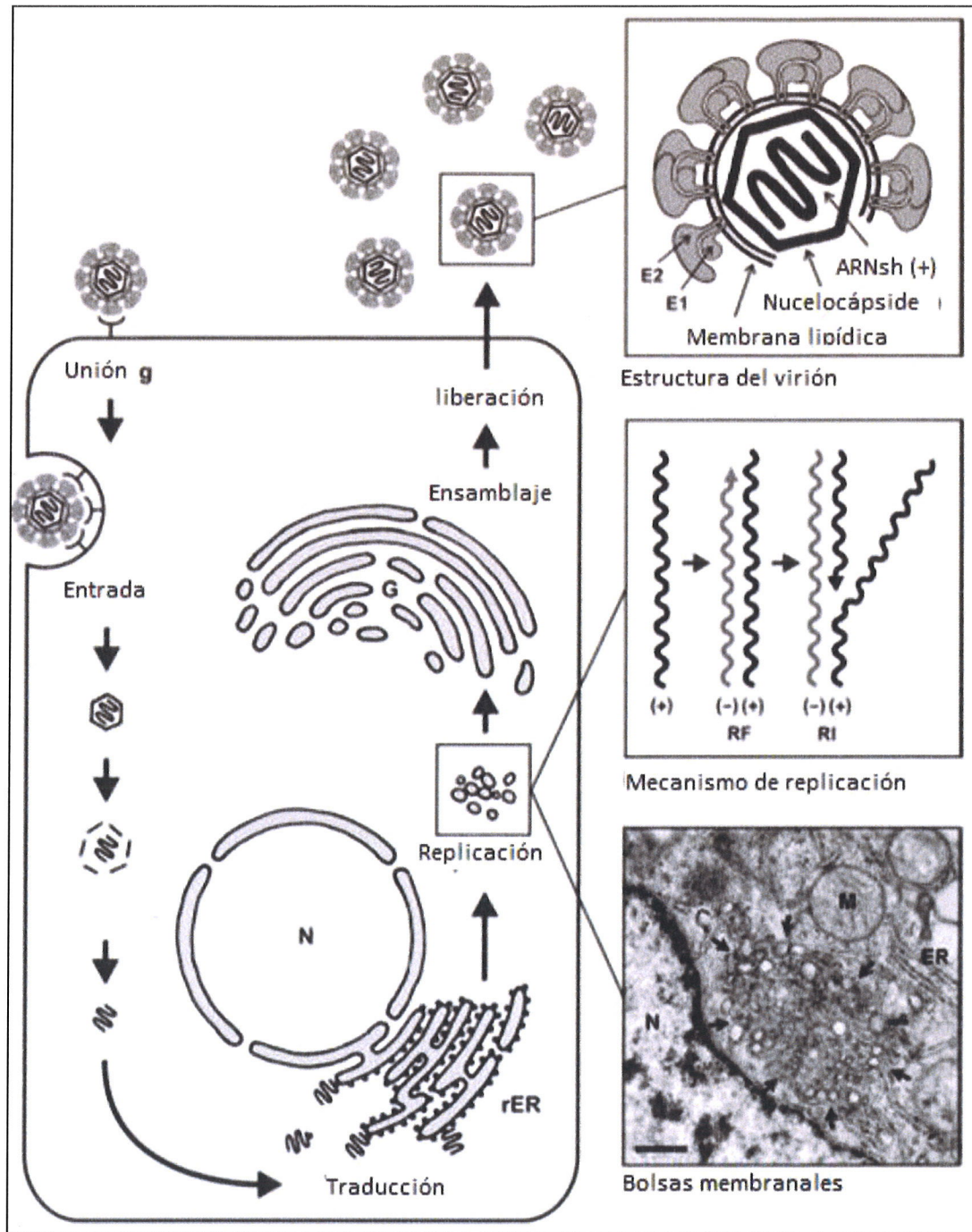
El ser humano es el único hospedero conocido para el virus de hepatitis C; sin embargo, los siete genotipos del virus que existen se encuentran distribuidos de forma no homogénea en la población mundial. Adicionalmente, estos genotipos varían aproximadamente un 30% entre sí, respecto a su material genético, lo cual ocasiona cambios en el desarrollo de la enfermedad y la transmisión de la misma (Sharma, 2010; Bartenschlager y Loham, 2000).

El genoma viral está conformado por una sola hebra simple de ARN en sentido positivo de aproximadamente 9600 nucleótidos. El virión contiene el material genético en una nucleocápside, envuelta por una membrana lipídica que presenta las glicoproteínas virales de envoltura E1 y E2 en la superficie (Figura 1) (Kim y Chang, 2013; Chevaliez *et al.*, 2006).

El genoma del virus de hepatitis C (Figura 2) contiene un único marco de lectura abierto delimitado por dos regiones no codificantes. En el extremo 5' contiene un sitio de entrada interna de ribosoma, el cual es necesario para la transcripción independiente de Cap (Kim y Chang, 2013; Sharma, 2010) y en el extremo 3' presenta regiones poli (U) y poli (U/C), las cuales son indispensables para la replicación del virus (Chevaliez *et al.*, 2006).

La transcripción del marco abierto de lectura en el retículo endoplasmático produce una poliproteína que es procesada con ayuda de proteasas del hospedero y del virus (como la NS3/NS4) para producir 10 proteínas funcionales del virus (Kim y Chang 2013; Sharma, 2010; Bartenschlager y Loham, 2000)

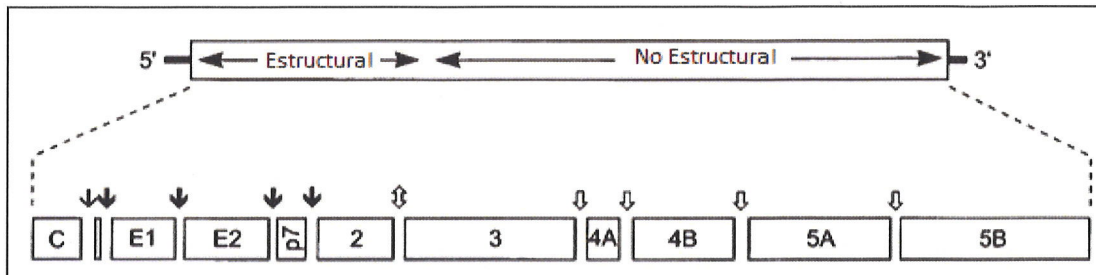
Figura 1: Ciclo de replicación del virus de hepatitis C dentro del hepatocito.



(Adaptado de Chevaliez *et al.*, 2006)



Figura 2: Esquema de la organización del genoma y la poliproteína producida por el único marco de lectura abierto. Las flechas negras representan los sitios de corte producidos por proteasas de hospedero y las blancas simbolizan los cortes producidos por la proteasa viral NS3.



(Adaptado de Bartenschlager y Loham, 2000)

2. Ciclo de replicación del virus. La infección a la célula comienza con la interacción de la glicoproteína E2 con proteoglicanos de la membrana del hepatocito. Los principales receptores que desencadenan la internalización del virus son el receptor CD81, el receptor *scavenger* humano y algunas moléculas que actúan en las uniones estrechas (Sharma, 2010). A pesar de que es posible la internalización mediada por el receptor LDL, ésta internalización promueve la degradación del virus (Kim y Chang, 2013).

El siguiente paso es la fusión de membranas con un endosoma ácido. Esto facilita la liberación del material genético; el cual, por presentarse en forma de ARN en sentido positivo, es capaz de sintetizar la poliproteína en el retículo endoplasmático rugoso sin necesidad de ningún intermediario. En este punto, la proteína NS4B se asocia con la membrana del retículo y promueve la formación de una red de membranas dentro de las cuales se dará la replicación del material genético del virus (Chevaliez *et al.*, 2006).

La replicación del material genético se da por la enzima NS5B, la cual es una polimerasa de ARN dependiente de ARN. Ésta toma la hebra de ARN en sentido positivo y mediante la interacción con la región no codificante del extremo 3' es capaz de formar una hebra de ARN en sentido negativo, la cual sirve de molde para la producción de nuevas hebras en sentido positivo por la enzima NS5B (Kim y Chang, 2013).

3. Diagnóstico y detección. Los síntomas más comunes en pacientes con hepatitis C aguda son: Anorexia, náusea, vómito, dolor abdominal y fatiga (Caruntu y Benea, 2007; WHO, 2003). Debido a que los síntomas son poco específicos, el diagnóstico de hepatitis C únicamente por cuadro clínico resulta poco preciso. Adicionalmente, aproximadamente el 80% de los casos de hepatitis C aguda se presentan sin desarrollo de síntomas (WHO, 2003). Por esta razón las pruebas serológicas y moleculares representan un paso clave en el diagnóstico de la enfermedad.

La primera prueba desarrollada para la detección de hepatitis C consiste en la detección de anticuerpos contra el virus de hepatitis C en el suero del paciente. Sin embargo, para la detección de anticuerpos en el suero es necesario esperar entre 4 y 8 semanas después de la infección. Actualmente es posible detectar la

proteína C de la nucleocápside del virus por métodos inmunoquímicos, lo que permite la detección entre 1 y 3 semanas después de la infección (Caruntu y Benea, 2007; Freeman *et al.*, 2001; WHO, 2003).

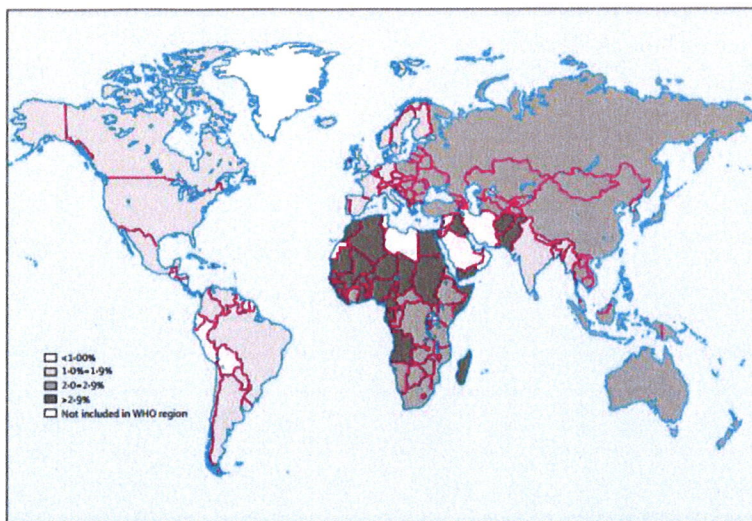
Entre las pruebas moleculares se encuentra la detección de ARN viral por medio de PCR de transcripción reversa. Esta técnica permite detectar la carga viral entre 1 y 2 semanas después de la infección y es de suma importancia para registrar la progresión y evaluación post-tratamiento del paciente (Caruntu y Benea, 2007; Freeman *et al.*, 2001). Adicionalmente, es posible determinar el genotipo del virus, lo que resulta importante en el pronóstico de la enfermedad, ya que los genotipos 1 y 4 son menos propensos a responder a la terapia antiviral clásica (WHO, 2003).

4. **Epidemiología.** Uno de los puntos críticos en la transmisión de la enfermedad ha sido la transfusión de sangre infectada a pacientes sanos. Sin embargo, los índices de infección por esta vía han disminuido gracias a las pruebas inmunoquímicas que se realizan en la donación de sangre. En la actualidad la mayor vía de infección es la utilización de agujas contaminadas para la inyección de drogas, seguido de la exposición a agujas no sanitizadas correctamente en la elaboración de tatuajes, acupuntura y otras actividades relacionadas. Adicionalmente, es importante resaltar que la transmisión vertical es casi exclusivamente en el momento del parto y que es posible la transmisión sexual, a pesar de ser muy inusual (Caruntu y Benea, 2007; Freeman *et al.*, 2001; Shepard *et al.*, 2005).

El virus de hepatitis C produce una de las infecciones más comunes, con una incidencia de aproximadamente 184 millones de personas en el 2005, lo cual representa un 2.8% de la población mundial (Figura 3) (Shepard *et al.*, 2005; Hanafiah *et al.*, 2013).

Según la Organización Mundial de la Salud (2002), entre 70% y 80% de los casos de infección con hepatitis C se vuelven crónicos; de 5% a 20% desarrollan cirrosis y entre 1% y 5% muere de cirrosis o cáncer hepático ocasionado por la infección de VHC (WHO, 2003, CDC, 2010).

Figura 3: Distribución mundial de VHC.



(tomado de Shepard *et al.*, 2005)

5. Tratamiento. El tratamiento de la infección consiste en la administración de Interferón alfa. Sin embargo, se ha observado que la adición de un polietilenglicol al interferón (pegIFN) mejora sus características farmacodinámicas, lo cual incrementa la efectividad del tratamiento. Adicionalmente, la coadministración de Ribavirina aumenta la respuesta al tratamiento, por lo que las guías mundiales de VHC recomiendan una terapia multidroga de pegIFN y ribavirina (Shepard *et al.*, 2005; Kato, 2001).

El régimen de tratamiento clásico contra hepatitis C consiste en una inyección semanal de interferón (1.5µg/Kg del paciente) conjunto con la administración oral diaria de Ribavirina (800 mg). El tiempo del tratamiento se encuentra relacionado con el genotipo del virus. Si el paciente presenta los genotipos 1, 5 o 6 debe consumir el tratamiento por 48 semanas, mientras que si presenta el genotipo 4 el tratamiento recomendado es de 36 semanas y los genotipos 2 o 3 únicamente requieren 24 semanas de tratamiento (WHO, 2003; Ghany *et al.*, 2009).

Actualmente existen dos inhibidores de la proteasa NS3/NS4 aprobados por la FDA (telaprevir y boceprevir). Se ha demostrado un incremento en la respuesta viral sostenida para el genotipo 1 del virus al utilizar la coadministración de dichos inhibidores con respecto al tratamiento clásico ( 66%-75% contra 42-46%, respectivamente). La administración de estos antivirales aún sigue restringida debido a la selección de cepas resistentes. Por lo que es clave determinar cuál es la población adecuada para dicho tratamiento (Pearlman y Traub, 2010; Booth *et al.*, 2012).

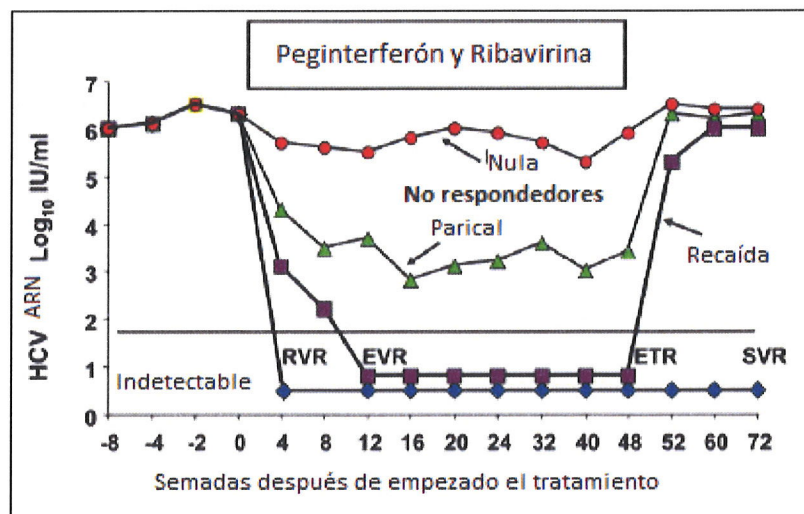


## B. RESPUESTA VIRAL SOSTENIDA

Los pacientes que presentan infección con hepatitis C tienden a desarrollar una amplia gama de respuestas al tratamiento con interferón y Ribavirina (Figura 4). Si en las primeras 4 semanas de tratamiento ya no se puede detectar presencia de ARN viral en el plasma del paciente con un método que posea un límite de detección de 50 IU/mL o menor se conoce como respuesta viral rápida (RVR), si es posible detectar ARN en el plasma del paciente pero la concentración ha disminuido al menos dos logaritmos se conoce como respuesta viral temprana (EVR, por sus siglas en inglés). En el caso en el que no es posible detectar una diferencia en la concentración de ARN en el plasma luego de iniciar el tratamiento se categoriza como respuesta nula (WHO, 2003; Ghany *et al.*, 2009).

Al terminar el tratamiento es necesario cuantificar la concentración de ARN viral en el plasma del paciente; sin embargo, una carga viral indetectable al final del tratamiento no es sinónimo de éxito del tratamiento o que el paciente logró depurar por completo el virus. Por esta razón es necesario un seguimiento de la carga viral del paciente durante las siguientes 24 semanas luego de concluido el tratamiento. Si durante este tiempo el paciente presenta nuevamente ARN viral detectable se considera una recaída (WHO, 2003; Ghany *et al.*, 2009; Pearlman y Traub, 2010).

Figura 4: Posibles respuestas de los pacientes al tratamiento de la infección con virus de hepatitis C: respuesta nula (círculos rojos), respuesta parcial (triángulos verdes), recaída (aumento del ARN observado al final del ejemplo de los cuadrados morados) y respuesta viral sostenida (rombos azules).



(Adaptado de Ghany *et al.*, 2009).

En el caso en el que un paciente presente carga viral indetectable al finalizar el tratamiento y que durante las siguientes 24 semanas de seguimiento su carga viral siga siendo indetectable por métodos con un límite de detección de 50 UI/mL se considera como Respuesta Viral Sostenida (SVR) (WHO, 2003; Ghany *et al.*, 2009; Pearlman y Traub, 2010).

La Respuesta Viral Sostenida es el único indicador de éxito en el tratamiento de VHC debido a que el porcentaje de recaídas luego de ese período varían entre 0% y 1%. Adicionalmente, se ha demostrado que los pacientes con SVR pueden recuperarse completamente de la inflamación hepática y la fibrosis desarrollada (Raptopoulou *et al.*, 2006; Pearlman y Traub, 2010).

### C. Farmacogenética

La primera referencia del concepto de farmacogenética proviene del 1898, cuando Sir Archibald Garrod observó que existía un cambio en la intensidad o el efecto que poseía una droga en los pacientes y debido a que Garrod estaba familiarizado con el trabajo de Mendel, él atribuyó esta diferencia a la existencia de la “individualidad química” de cada paciente (Mayer, 2004).

La interacción droga-droga significa la consideración de que existen drogas que pueden alterar el metabolismo de otras drogas, afectando así su efecto en el paciente. La revisión de dicha interacción es un factor determinante en la elección del medicamento a utilizar. Sin embargo se ha demostrado que existen interacciones gen-droga que pueden presentar efectos similares y que es necesario considerar para poder brindar un tratamiento efectivo y seguro al paciente. Un ejemplo de esto es el anticancerígeno Tamoxifen. Éste necesita ser activado por el citocromo P450 2D6 (CYP4502D6) para tener efecto, por lo que cualquier inhibidor del CYP4502D6 (como los inhibidores selectivos de la recaptación de serotonina, SSRI) disminuye el efecto del Tamoxifen. Por esta razón la FDA recomienda que no se coadministre SSRI junto con Tamoxifen, sin embargo existen mutaciones del CYP4502D6 que producen efectos como la disminución y anulación de la activación del Tamoxifen y que deben ser considerados antes de decidir tratar a un paciente con dicho anticancerígeno (Amstutz y Carleton, 2011; Becker, 2007).

La farmacogenética estudia la diferencia del efecto que tiene una droga en un paciente, ocasionada por la composición genética específica del paciente. En sus orígenes se enfocaba en diferencias farmacocinéticas de las drogas producidas por distintos alelos de las enzimas involucradas en el metabolismo de dicha droga; sin embargo, en la actualidad permite predecir el resultado que tendrá un tratamiento en un paciente, la probabilidad desarrollar efectos secundarios graves o la necesidad de ajustar la dosificación. Por todo esto las pruebas farmacogenéticas deben ser utilizadas como una herramienta en la elección del mejor tratamiento para un paciente en específico (Mayer, 2004).

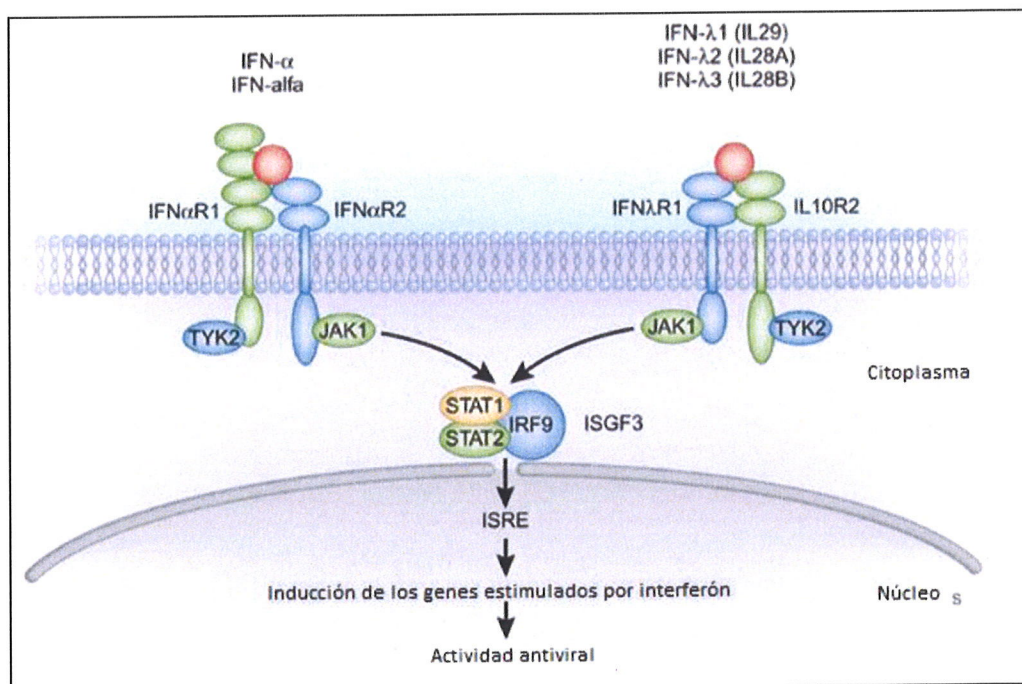
## D. IL-28B

1. Interleucinas (genes, proteínas, receptores). Las interleucinas (IL) son una familia de proteínas inmunomoduladoras que funcionan como mensajeros, interactuando con receptores de membrana de las células en las que presentan efecto. En estas células pueden modular el crecimiento celular, la diferenciación celular o la respuesta inmune (Booth *et al.*, 2012; Hydes y Khakoo, 2010, 24).

La interleucina 28B (IL28B), también conocida como interferón lambda 3 (IFN  $\lambda$ 3), se encuentra codificada en el brazo q del cromosoma 19 (19q13.13) y produce una citosina de la clase II. Esta clase se caracteriza por presentar un motivo de 6 o 7 hélices alfa apiladas en sentido antiparalelo y por actuar sobre receptores heterodímeros de la superficie celular en respuesta a una infección viral o bacteriana (Soriano *et al.*, 2012; Brocker *et al.*, 2010).

2. Ruta de activación. La citosina IL28B es producida en respuesta a la infección viral o bacteriana y tiene efecto mediado por la cascada de fosforilación de la proteína JAK1 (Figura 5), la cual induce la expresión de los genes estimulados por el interferón que producen una actividad antiviral. Sin embargo, se ha reportado que la IL28B presenta actividad antiviral directa uniéndose e inhibiendo el complejo de replicación e inhibiendo el crecimiento celular de la célula infectada (Brocker *et al.*, 2010; Afdhal *et al.*, 2010).

Figura 5: Vía de activación de la Interleucina 28B mediante la cascada de activación de JAK 1.



(Adaptado de Afdhal *et al.*, 2010)



3. IL28B asociado con SVR (sr12979860). En el 2010 se publicaron una serie de artículos que presentan evidencia de que algunos polimorfismos de un sólo nucleótido (SNP por sus siglas en inglés) cercanos al IL28B se encuentran relacionados a la respuesta del paciente al tratamiento de interferón con ribavirina (pegIFN/RBV) (Rauch *et al.*, 2010; Pineda *et al.*, 2010).

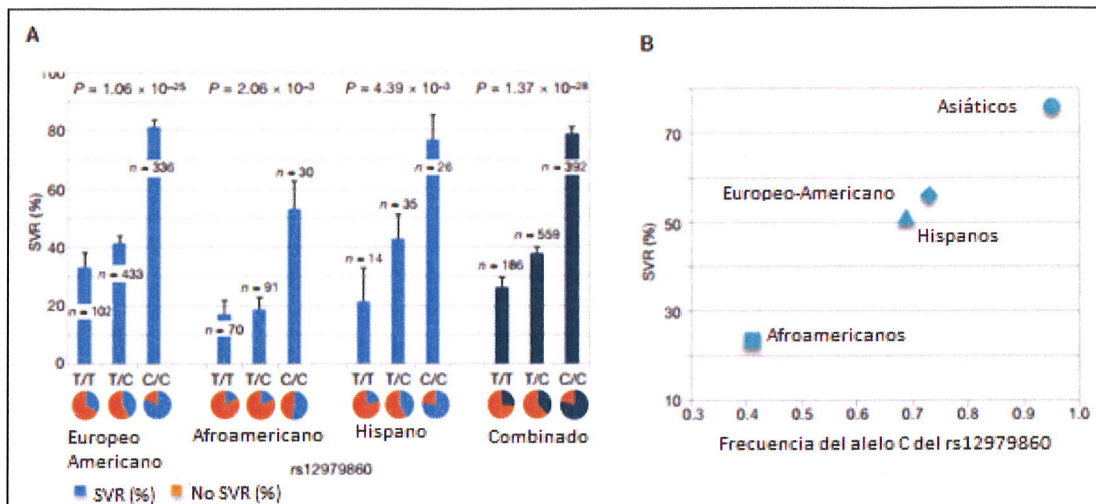
Dentro de los SNPs con potencial de predicción de SVR en un paciente con tratamiento pegIFN/RBV sobresale el sr12979860. Un genotipo homocigoto C (CC) para dicho SNP ha sido relacionado con mayores índices de SVR al completar el tratamiento que el genotipo heterocigoto (CT) o homocigoto T (TT) ( $p \leq 0.001$ ) en distintas poblaciones (Figura 6) (Marco *et al.*, 2012; Lin, 2011; Rauch *et al.*, 2010; Pineda *et al.*, 2010). Adicionalmente se ha encontrado que los genotipos CT y TT de dicho SNP están más asociados a fibrosis severa hepática que el genotipo CC ( $p = 0.005$ ) (Marco *et al.*, 2012).

Un análisis del INF $\lambda$ 3 ha demostrado que luego de empezar el tratamiento con pegIFN/RBV los pacientes con el genotipo CC incrementan sus niveles de INF $\lambda$ 3, mientras que los niveles de INF $\lambda$ 3 de los pacientes con genotipos CT o TT permanecen iguales (Marco *et al.*, 2012; Booth *et al.*, 2012).

La infección con VHC genotipo 1 es la más común a nivel mundial. Igualmente ésta desarrolla, en promedio, menor porcentaje de SVR al tratarse con pegIFN/RBV que los genotipos 2 y 3. El SNP sr12979860 ha demostrado tener el mejor índice de predicción en los pacientes infectados con VHC genotipos 1 y 4, lo cual significa que esta prueba permite predecir la respuesta del paciente al tratamiento con pegIFN/RBV previo a la administración de dicho tratamiento (Marco *et al.*, 2012; Pineda *et al.*, 2010). De este modo el marcador farmacogenético sr12979860 permite clasificar la población de pacientes que requieren tratamiento contra VHC en aquellos que tendrán una respuesta positiva a pegIFN/RBV y aquellos a los que se recomienda tratar con un inhibidor de proteasa NS3/NS4 (Booth *et al.*, 2012; Marco *et al.*, 2012).

Graells (2013) desarrolló un estudio de costos acerca de la implementación de la prueba de IL28B en el tratamiento de hepatitis C para el Hospital General de Alicante de España. Dicho estudio demostró que la implementación de esta prueba farmacogenética representaría un ahorro de 2,916,000 euros al año.

Figura 6: A) Porcentaje de SVR alcanzado por pacientes con distintos genotipos del SNP rs12979860 según su etnicidad y B) Correlación entre porcentaje de SVR alcanzado por los pacientes y frecuencia del alelo CC del SNP rs13979860.



(Adaptado de Rauch *et al.*, 2010)

4. Homología entre la familia IL28. La IL28B se encuentra estrechamente relacionada con la interleucina 28A y la interleucina 29A. Éstas tres interleucinas son evolutivamente muy cercanas, lo cual se representa en la alta homología que existe entre las mismas, no sólo en regiones codificantes sino en regiones no codificantes también (Soriano *et al.*, 2012). Hasta el año 2012 existían diversas publicaciones referentes a la detección del IL28-B y el pronóstico de SVR en pacientes con VHC pero ninguno de ellos consideraba esta homología como parte de su análisis. En el año 2012 se publicó el primer artículo donde se resaltaba la importancia de una región 95% homóloga en el gen IL28-A en el diagnóstico farmacogenético del IL28-B (Reynolds *et al.*, 2012) y desde entonces solamente se ha reportado una publicación adicional (Sharafi *et al.*, 2013) en la cual se considera dicha región en el diseño de los iniciadores que utilizan para la detección y diagnóstico de dicho marcador.

## E. Técnicas de detección

1. Reacción en Cadena de la Polimerasa. La Reacción en Cadena de la Polimerasa (Polymerase Chain Reaction, PCR) es una técnica descrita originalmente en 1971 por Kleppe *et al.* y es una técnica de Biología Molecular ampliamente utilizada ya que permite reproducción de un segmento de material genético *in situ*. Esta técnica se puede aplicar de diversas formas y para fines muy variados pero la reacción básica permanece igual.

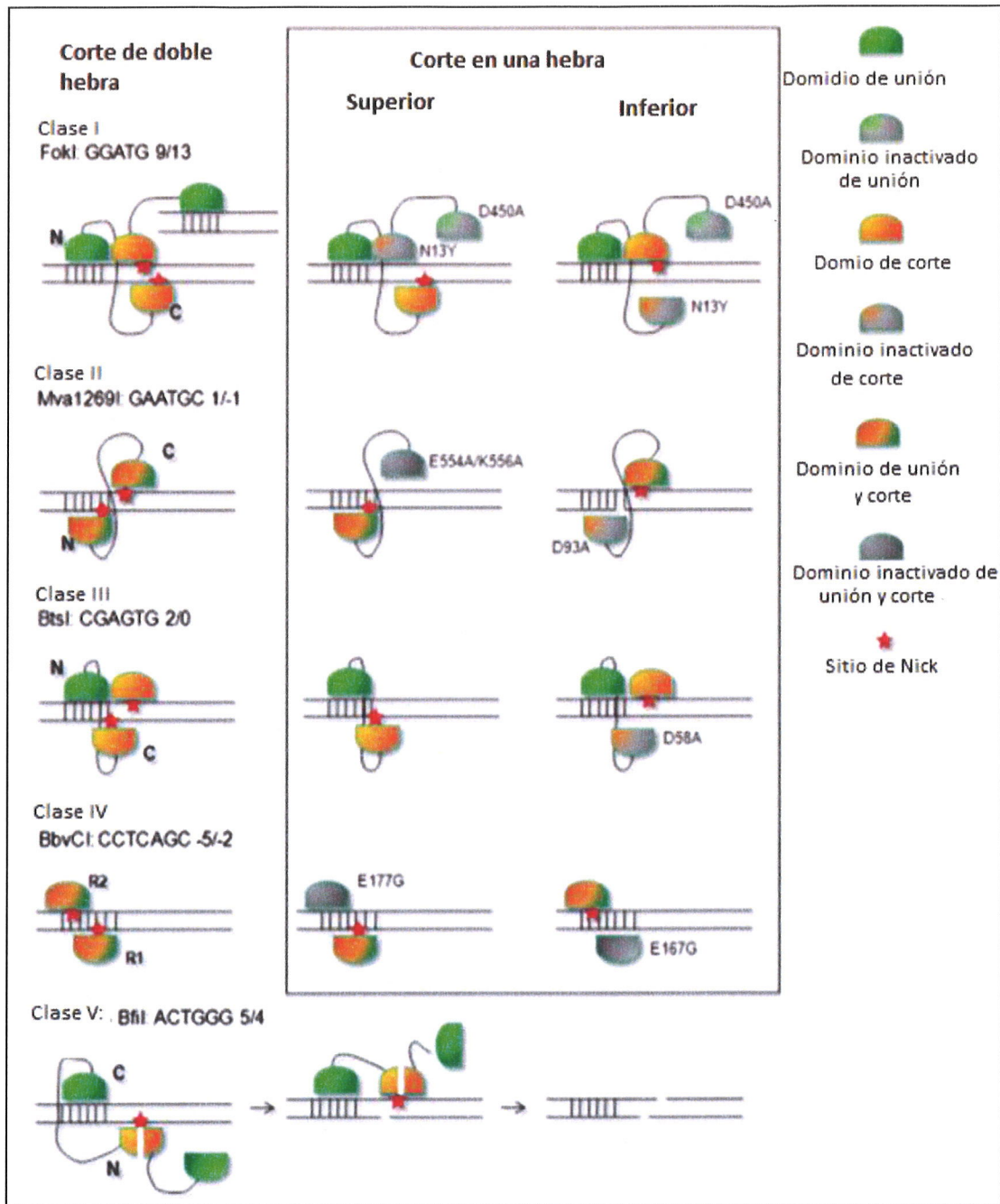
El PCR se basa en utilizar la maquinaria de replicación celular para la amplificación de un segmento de ADN de interés para el investigador. Se utiliza la polimerasa de ADN dependiente de ADN estable a

temperatura ambiente denominada Taq Polimerasa de ADN, la cual es extraída de *Thermus aquaticus*. Esta enzima es colocada en un ambiente electroquímico similar a las condiciones celulares, el cual se obtiene mediante la utilización de un buffer diseñado para esta función y para el buen funcionamiento de la enzima se agrega un ión catiónico, usualmente  $Mg^{++}$  en forma de Cloruro de Magnesio. Al igual que en la célula la polimerasa requiere de fragmentos cortos de ADN con el extremo 3' disponibles para iniciar la replicación de ADN. Por esta razón se introducen iniciadores (primers) los cuales no solo proporcionan el grupo hidroxilo libre 3' necesario para la actividad de la enzima sino que también determinan qué regiones serán amplificadas de todo el material genético que se presente en la muestra.

2. PCR Alelo específico (Allele Specific PCR, asPCR). El PCR alelo específico o asPCR, es un caso específico de la reacción normal de PCR en el cual uno de los iniciadores es colocado sobre uno o más nucleótidos que permiten diferenciar entre dos alelos de un gen. Esta modalidad de PCR es muy utilizada en el diagnóstico de SNPs, especialmente por su bajo costo comparado con otras tecnologías (Gaudet *et al.*, 2009).

3. PCR en tiempo real (qPCR). El PCR en tiempo real es una variación de PCR que permite la cuantificación de la amplificación que se produce, por esta razón se conoce también como PCR cuantitativo (quantitative PCR: qPCR). Esta tecnología puede utilizar una o más sondas, dependiendo de la cantidad de alelos o mutaciones que se desee detectar, las cuales son diseñadas para alinearse en la región que contiene los nucleótidos de interés. Esta tecnología permite detectar el genotipo en poblaciones mixtas, es decir que las células de la muestra presenten más de un genotipo. El principio básico del qPCR depende de la acción exonucleasa 5' de la ADN polimerasa ya que durante el paso de elongación del PCR se hidroliza el nucleótido 5' (que presenta el fluoróforo en la sonda), lo cual activa al fluoróforo, el cual es excitado mediante radiación a cierta longitud de onda y al regresar al estado basal emite radiación que es captada por el equipo y traducida como una señal de amplificación (Biosearch Technologies, 2014).

Figura 7: Ilustración del mecanismo de acción de los distintos tipos de endonucleasas de restricción.



(Adaptado de Sui-Hong *et al.*, 2010)

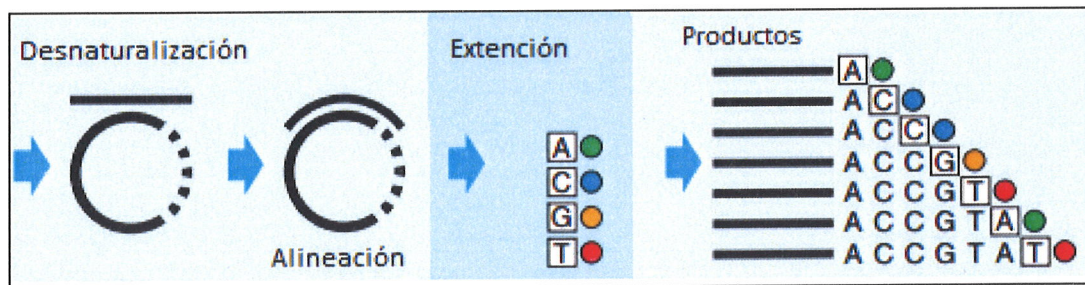


4. Polimorfismo de tamaño de la restricción enzimática de un producto de PCR. Las enzimas de restricción representan una poderosa herramienta en la biología molecular y la ingeniería genética. Estas fueron descubiertas por primera vez al estudiar el factor que ‘restringía’ la infección de un bacteriófago a *E. coli*. Esto resultó ser producido por dos enzimas, EcoKI y EcoBI, las cuales tenían la capacidad de reconocer secuencias en el genoma del bacteriófago e hidrolizar la cadena de ADN. En la actualidad se han descrito cinco tipos de endonucleasas de restricción, numeradas de I al V. En la Figura 7 se ilustra el mecanismo de acción de estas enzimas. El grupo más importante para las aplicaciones en biología molecular es el grupo de endonucleasas tipo II, ya que estas presentan una secuencia de reconocimiento y realizan la hidrólisis de la cadena en ambas hebras en una posición constante respecto a la secuencia de reconocimiento. La enzima Bstul es un ejemplo de estas enzimas (Sui-Hong *et al.*, 2010).

5. Secuenciación. La secuenciación Sanger es un proceso basado en el PCR en el cual se introducen deoxinucleótidos para la polimerización de ADN, sin embargo se introducen un porcentaje menor de dideoxinucleótidos, los cuales al incorporarse a la cadena de ADN detienen completamente la polimerización de ADN ya que no poseen el grupo hidroxilo 3’ necesario para la polimerización del siguiente nucleótidos.

Esta reacción de secuenciación acoplada a la electroforesis capilar permite la detección de todos los subfragmentos que se producen de los distintos tamaños mediante moléculas coloreadas ligados a cada dideoxinucleótidos. La Figura 8 ilustra el proceso.

Figura 8: Ilustración de cómo al incorporarse los dideoxinucleótidos marcados producen fragmentos de distintos tamaños marcados con distintos colores.





## V. MATERIALES Y MÉTODOS

### A. TIPO DE INVESTIGACIÓN

Esta es una investigación aplicada que pretende evaluar las metodologías publicadas para la detección del SNP rs12979860 y proponer herramientas moleculares óptimas para la detección de dicho SNP en las siguientes tecnologías: asPCR de punto final, as-qPCR, secuenciación, FRLP-PCR. Con esto se desea facilitar la implementación de esta prueba farmacogenética para resolver las necesidades inmediatas de la población guatemalteca.

### B. METODOLOGÍA

#### 1. Análisis de Homología de IL28B-IL28A

a. Determinación de las secuencias de trabajo. Se determinó la posición absoluta del SNP rs12979860 en el cromosoma 19 haciendo una búsqueda en la base de datos de SNPs (dbSNP) de NCBI, seguido se visualizó el SNP en la secuencia de referencia más reciente para el cromosoma 19 y se seleccionó una región 10Kb río arriba y 10Kb río abajo del SNP. Esta región se descargó bajo el nombre “Región SNP”.

Mediante la utilización del recurso Gene de NCBI se determinó el locus del interferón lambda. A continuación se visualizó dicho locus y se delimitó una región del cromosoma que contenga todos los genes del interferón lambda, especialmente IL28-A e IL28-B. Esta región se descargó bajo el nombre “Región IFNL”.

b. Determinación de la región homóloga y la región de interés. Se realizó la alineación de “Región IFNL” contra “Región SNP” para establecer los sitios de homología, utilizando la herramienta Blastn de NCBI. Se calculó la posición relativa del SNP en la homología restando la posición de inicio de la secuencia “Región SNP” de la posición absoluta del SNP rs12979860 en la secuencia de referencia del cromosoma 19. A continuación, se localizó el SNP en la alineación y se colocó un etiqueta sobre el SNP. Todas las subsecuencias de “Región IFNL” que contenían al SNP dentro de la región alineada se consideraron posibles regiones homólogas. Se revisó que todas las posibles regiones homólogas no se encontraban en el borde de los límites superior o inferior de “Región IFNL”. Si este hubiese sido el caso se hubiera descargado una nueva “Región IFNL” que contuviera un rango 2Kb río arriba más y 2Kb río abajo más, con la cual se repetiría todo el procedimiento descrito en esta sección. Se repetiría este procedimiento hasta que todas las posibles regiones homólogas se encontraran completamente dentro de “Región IFNL”.

Se determinó la posición absoluta en la secuencia de referencia del cromosoma 19 de la región homóloga dentro del IL28-B adicionando la posición relativa de dicha subregión alineada dentro de la secuencia “Región IFNL” a la posición de inicio de la secuencia “Región IFNL” en la secuencia de referencia del cromosoma 19. Igualmente, se determinó la posición absoluta de todas las regiones de “Región SNP” que presenten una región homóloga (regiones de interés) adicionando la posición relativa de la subsecuencia alineada dentro de “Región SNP” a la posición de inicio de la secuencia “Región SNP”. Estas regiones se buscaron en la secuencia de referencia del cromosoma 19 y se descargaron bajo los nombres “Región homóloga” +(número) y “Región interés”+(número correspondiente de su región homóloga), respectivamente.

2. Diseño de iniciadores específicos. Debido a que la prioridad en el diseño de los iniciadores es considerar la región homóloga, no se utilizó ningún programa o herramienta de diseño de iniciadores para extraer la secuencia de los iniciadores. Por esta razón, se visualizó la alineación para encontrar regiones que contengan múltiples no-homologías cercanas y que se encuentren a una distancia entre 200 y 600pb del SNP. En esta región se seleccionaron secuencias entre 20 y 30 nucleótidos como propuesta de iniciador con la mayor parte de las no-homologías hacia el extremo 3'. Otro factor que se consideró es la distancia entre el SNP y ambos iniciadores debe ser suficiente para poder diferenciarse en un gel de agarosa. Luego de elegidas las regiones para los iniciadores se utilizó la herramienta de PrimerBlast de NCBI para evaluar qué tan posible es la amplificación de otras regiones. Para los iniciadores con amplificación específica se analizó las características fisicoquímicas de cada pareja de iniciadores utilizando la herramienta Oligo Analyzer de Integrated DNA Technologies. Finalmente se utilizó la herramienta NEBcutter de New England Biolabs para producir los resultados de la FRLP de los alelos CC y TT, adicionalmente se utilizó un programa de edición de imágenes para producir el resultado del alelo CT al igual que producir los resultados de asPCR. Este proceso se realizó alterando el tamaño y colocación de los iniciadores para mejorar las características fisicoquímicas del PCR deseado.

3. Evaluación de iniciadores reportados. Se realizó una búsqueda en Internet sobre detección de IL28B o rs12979860 para construir una base de datos de los iniciadores reportados para la amplificación o detección de dicho SNP. En dicha base de datos se recopiló adicionalmente a la secuencia de los primers para amplificación, la técnica para la cual se utilizan, los autores y el año en que se produjo la publicación y si mencionan la existencia de una región homóloga.

Al usar la herramienta PrimerBlast de NCBI, se aplicó la secuencia “Región SNP” como la plantilla de amplificación y cada una de las parejas de iniciadores para evaluar su amplificación. Se recopiló resultados respecto al tamaño del producto de PCR, otras posibles amplificaciones y la posición relativa en la que se colocan los iniciadores en la secuencia “Región SNP”.

Al utilizar la alineación IL28B-IL28A se buscó los iniciadores para evaluar el sitio de colocación respecto a la homología respecto a qué tantas no-homologías abarca y que tan cercanas son al extremo 3'. A

continuación se hizo un PCR in silico con el software especializado para establecer si existen otras posibles amplificaciones y determinar el amplicón que producen.

4. **Determinación de iniciadores.** En base a los resultados recopilados en los incisos anteriores acerca de los iniciadores publicados y de los iniciadores diseñados se determinó la pareja óptima de iniciadores para la amplificación específica del SNP.

## 2. Diseño de métodos de detección

a. PCR-RFLP. Se buscaron de todas las enzimas que dentro de su sitio de reconocimiento contuvieran el SNP en cualquiera de sus alelos homólogos y se produjo resultados teóricos de la restricción enzimática en forma de imágenes de geles de Agarosa. Se utilizó la herramienta NEBcutter de NEW England Biolabs para producir los resultados de la restricción enzimática de los alelos CC y TT, adicionalmente se utilizó un programa de edición de imágenes para producir el resultado del alelo CT.

b. asPCR. Se diseñaron los iniciadores para el PCR alelo específico colocándolos sobre el SNP. El iniciador en sentido positivo se colocó con el nucleótido terminal C en el extremo 3' y el iniciador en sentido negativo se colocó con el nucleótido terminal T en el extremo 3'. Utilizando un programa de edición de imágenes se construyeron los resultados teóricos de la amplificación de los distintos genotipos utilizando asPCR.

c. as-qPCR. Se evaluaron los iniciadores publicados y los diseñados para determinar cual representa una mejor opción para la implementación de la prueba por PCR en tiempo real.

d. Secuenciación. Se determinó las condiciones óptimas de equipo y reactivos para realizar la secuenciación utilizando los iniciadores elegidos anteriormente y utilizados para asPCR.

e. Análisis de costos. Se recopilaron los costos estimados de los reactivos en base a las cantidades y concentraciones de los estándares comerciales y con base en esto se determinó el costo promedio por prueba para cada método.

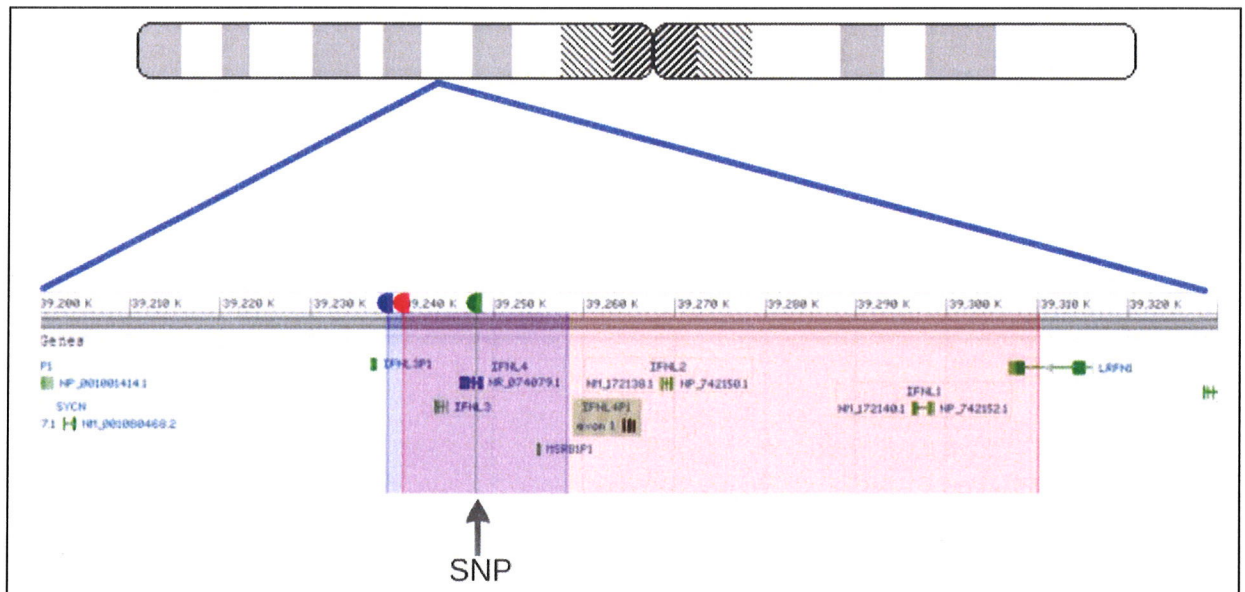


## VI. RESULTADOS

### A. Análisis de homología

Para obtener las secuencias de trabajo se utilizó la secuencia de referencia NC\_00019.10 del genoma de referencia GRCh38, la cual corresponde al cromosoma 19 completo. En esta se estableció que la posición del SNP rs12979860 era 39248147. En base a esta posición se determinó que la 'Región SNP' abarca del nucleótido 39238147 hasta el nucleótido 3925147 de dicha secuencia de referencia. Adicionalmente la 'Región SNP' (del nucleótido 39240000 al 39310000) se estableció luego de visualizar la posición de todos los genes del interferón lambda (Figura 9) para asegurarse de que dicha región abarcara todos los genes dicho interferón. La secuencias de trabajo que se utilizaron se encuentran en la sección de Anexos.

Figura 9: Esquema del cromosoma 19 y visualización de la región 19q13.13 obtenido de NCBI-Blast. El marcador verde indica la posición del SNP rs12979860, la región azul representa la secuencia de trabajo nombrada 'Secuencia SNP' y la región roja representa la 'Secuencia IFN'.



Al realizar la alineación inicial para observar todas las regiones homólogas que contengan al SNP rs12979860 (Figura 10) se puede observar que solamente dos subsecuencias presentan homología en dicha región. La primera es una región que presenta un 100% de homología y que representa la región entre 39240000 y 39258147 del cromosoma 19, la cual se puede visualizar en morado en la Figura R1 ya que es la región de intersección de ambas secuencias de trabajo. La segunda región es una alineación entre las subsecuencias de 7552 a 10896 de la 'Región SNP' y de 23954 a 27256 de la 'Región IFN'.

Con base en las posiciones relativas de las subsecuencias en cada una de las secuencias de trabajo se estableció las posiciones absolutas en el cromosoma 19 utilizando la Ecuación 1 en donde  $P_{abs}$  refiere a la posición absoluta, ya sea de inicio o final, de la subsecuencia en el cromosoma 19;  $P_{rel}$  refiere a la posición relativa, ya sea de inicio o final, de la subsecuencia en la secuencia de trabajo y  $P_o$  refiere a la posición inicial de la secuencia de trabajo en donde se encuentra la subsecuencia.

$$P_{abs} = P_{rel} + P_o - 1$$

$$P_{abs} = 7552 + 39238147 - 1$$

$$P_{abs} = 39245698$$

Al aplicar la misma deducción a las posiciones de inicio y final de las subregiones homólogas y de interés se estableció que la homología que contiene al SNP rs12979860 se encuentra entre 39245699-39249043 y 39263954-39267256. Por lo que estas secuencias se descargaron como 'Región interés' y 'Región Homóloga', respectivamente.

La alineación de la región de interés y la región homóloga se realizó en tres distintas herramientas bioinformáticas. En la Figura 11 se observan los resultados de dichas herramientas. La herramienta Blast de NCBI permite visualizar la posición de las no homólogas en una forma general de la secuencia, ya que se establecen como posiciones rojas en la secuencia homóloga y guiones rojos en las discontinuidades de la secuencia homóloga. El resultado de CLC Sequence Viewer permite visualizar los porcentajes de homología, ya que subdivide las secuencias en secciones cortas y les establece un porcentaje de homología, el cual se observa como un diagrama de barras debajo de la secuencia. El resultado que presenta ClustalW de EBI en una forma no gráfica de observar la homología que permite visualizar de una forma más fácil la discontinuidad de la homología, utilizando un asterisco debajo de los nucleótidos homólogos y un guión en las posiciones de los nucleótidos de las discontinuidades.

De las 3349pb de la región de interés se presentó 3076 bases alineadas con la región homóloga en sentido +/- . Adicionalmente, se determinó la presencia de 60 bases de discontinuidad de la región homóloga con la región de interés. Con estos datos se determinó un 93.64% de homología entre ambas regiones.

Figura 10: Alineación de la secuencia de 'Región SNP' contra subsecuencias de 'Región IFN' generadas por NCBI-Blast. La marca verde indica la posición del SNP rs12979860 y en la secuencia de 'Región SNP' y las sub-secuencias alineadas contra la misma. La coloración gris claro de las subsecuencias indica las bases homólogas, mientras que la coloración roja indica las bases no homólogas.

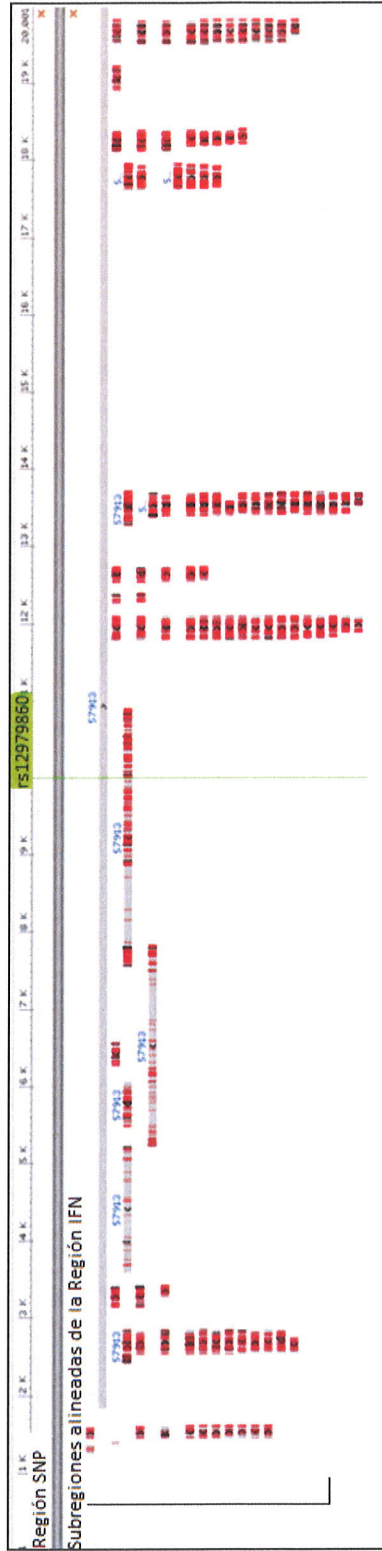
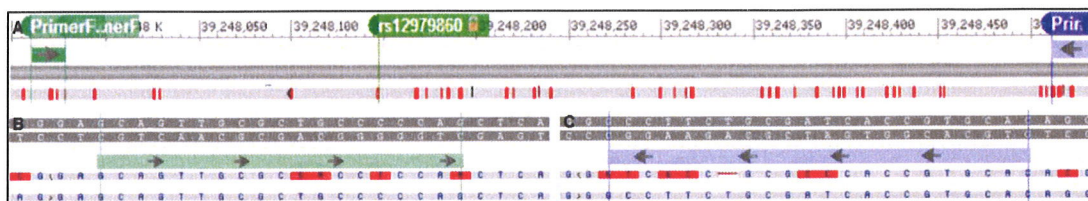






Figura 12: Esquema de la colocación de los iniciadores diseñados sobre la homología para observar A) el amplicón que se produciría como producto de PCR de la alineación de los iniciadores en la región de interés y la región homóloga y la alineación del iniciador B) Foward y C) Reverse sobre la secuencia de ambas regiones



Para el diseño de los iniciadores se observaron regiones ricas en no-homologías para colocar los iniciadores con la mayor cantidad de no- homologías en el extremo 3'. En la Figura 12 se observa la colocación de los iniciadores sobre la alineación y la posición asimétrica del SNP rs12979860 en el amplicón producido. El iniciador Foward se colocó sobre 4 no homologías, mientras que el iniciador Reverse se colocó sobre 6 no homologías y una discontinuidad de un nucleótido. En el Cuadro 1 se presentan los resultados termodinámicos del análisis de los iniciadores diseñados, los cuales producen un aplicón de 573pb. Se observa que el iniciador Foward presenta una temperatura de anidación de 65.1°C, mientras que el iniciador Reverse tiene una temperatura de anidación teórica de 63.2°C. Ambos iniciadores tienen un porcentaje de GC mayor al 60% y presentan una energía libre de Gibbs menor a -40Kcal/mol.

Cuadro 1: Parámetros termodinámicos de los iniciadores diseñados para amplificar la región de interés calculados utilizando Oligo Analyzer.

Delta G (kcal/mol)	TM (C)	Contenido GC	Largo	Secuencia	Iniciador
-44.34	65.1	73.70%	19	GCAGTTGCGCTGCCCCAG	<b>Forward</b>
-45.96	63.2	63.60%	22	GTGCACGGTGATCGCAGAAGGC	<b>Reverse</b>

Entre los heterodímeros determinados para el iniciador Foward se observa que el valor máximo de energía libre de Gibbs es de entre -9.89 (Figura 13), mientras que el valor máximo el iniciador Reverse es de -9.73kcal/mol (Figura 14).



Figura 13: Estructuras de homodímeros, la cantidad de nucleótidos que interactúan y los respectivos valores de Energía libre de Gibbs producidas para la interacción del Iniciador Forward.

Delta G (kcal/mol)	Bases Apareadas	5'	3'	Delta G (kcal/mol)	Bases Apareadas	5'	3'
-9.89	4	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG	-3.14	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG
		: :     :				:	
-6.69	4	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG	-3.14	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG
		:::				:    :	
-5.09	3	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG	-3.14	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG
		:				:    :	
-3.55	3	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG	-3.14	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG
		:				:    :	
-3.14	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG	-1.95	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG
						: :	
-3.14	2	GCAGTTGGGCTGCCCCAG	GACCCCGTCGCGTTGACG			: :	
		:    :				: :	

Figura 14: Estructuras de homodímeros, la cantidad de nucleótidos que interactúan y los respectivos valores de Energía libre de Gibbs producidas para la interacción del Iniciador Reverse.

Delta G (kcal/mol)	Bases Apareadas	5'	3'	Delta G (kcal/mol)	Bases Apareadas	5'	3'
-9.73	6	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG	-3.61	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG
						:     :: :	
-5.09	3	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG	-3.3	3	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG
		:				:    :	
-4.62	4	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG	-3.14	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG
		: : :     : :				: : : :	
-3.61	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG	-3.14	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG
		:    :				:	
-3.61	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG	-3.14	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG
		:    : : :					
-3.61	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG	-1.57	2	GTGCACGGTATCGCAGAAGGC	CGGAAGACGCTAGTGGCACGTG
		: :    :				:    : :	



Figura 15: Estructuras de heterodímeros, la cantidad de nucleótidos que interactúan y los respectivos valores de Energía libre de Gibbs producidas para la interacción del Iniciador Forward y Reverse diseñados.

Delta G (kcal/mol)	Bases Apareadas	5'	3'	Delta G (kcal/mol)	Bases Apareadas	5'	3'
-8.7	4	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-3.14	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			: :       :				: :     :
-6.69	4	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-3.14	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			: :				: :
-6.21	3	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-3.07	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			: :				:
-5.09	3	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-3.07	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
-5.09	3	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-3.07	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			:				
-5.09	3	5' <u>GTGCACGGTATCGCAGAAGGC</u>	3' GACCCCGTCGCGTTGACG	-3.07	2	5' <u>GTGCACGGTATCGCAGAAGGC</u>	3' GACCCCGTCGCGTTGACG
			:       : : :				
-3.61	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-1.95	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			:     :				
-3.14	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-1.95	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			: : : :				: : :
-3.14	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-1.60	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			: :				: : :
-3.14	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG	-1.34	2	5' GTGCACGGTATCGCAGAAGGC	3' GACCCCGTCGCGTTGACG
			:				:

Para la interacción entre el iniciador Forward y el iniciador Reverse se terminaron 20 heterodímeros con energía libre de Gibbs entre -8.7 y -1.34. Las interacciones que se observan en la Figura 15 involucran entre 4 y 2 nucleótidos. En el cuadro 2 se observan las conformaciones de Hairpin que se producen en cada iniciador, para los cuales la energía libre de Gibbs máxima calculada es de -2.9Kcal/mol.

Cuadro 2: Información de las estructuras de Haipin producidas por los iniciadores, con su respectivo Delta G y Tm.

Iniciador	Delta G (kcal/mol)	Tm (C)
Foward	-2.9	53.6
	-1.95	45.1
Reverse	-0.97	39.7
	-0.38	30.1
	-0.33	31
	-0.18	27.5
	-0.17	29.7

Luego de evaluar los parámetros de la amplificación de los iniciadores diseñados se evaluó los resultados que estos producirían para la detección del SNP con FRLP-PCR, asPCR y secuenciación. La digestión enzimática se realizó con Bst<sup>u</sup>I, como se ha reportado con anterioridad y para el alelo CC se produjeron 4 bandas de los siguientes tamaños: 272, 187, 89 y 25 pb, mientras que el alelo TT produjo solamente 3 bandas de 272, 212 y 89pb. Los resultados del asPCR se produjeron utilizando los iniciadores internos descritos en el Cuadro 5, en donde se produjeron dos amplicones para los alelos homocigotos (CC Y TT) y tres amplicones para el alelo heterocigoto CT.

En la Figura 3 de la sección de Anexos se observa el amplicón producido por los iniciadores y los sitios de restricción para la enzima Bst<sup>u</sup>I para los genotipos CC y TT, en donde los sitios de restricción se encuentran subrayados para facilitar su localización. Adicionalmente se observa un esquema de la digestión enzimática de los amplicones de los genotipos CC y TT.

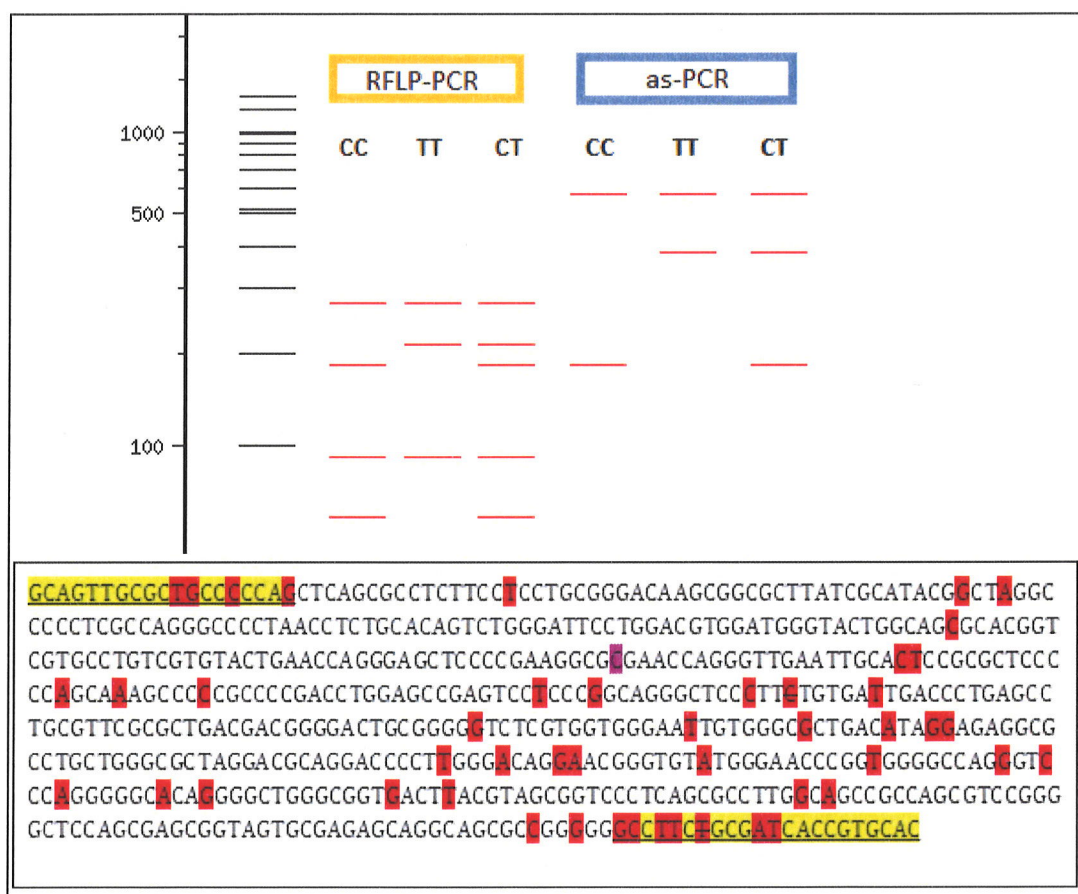
En la Figura 17 se muestran los resultados del análisis de 8 publicaciones entre 2010 y 2013 que produjeron un amplicón *in silico* en la IL28B. En esta revisión se establece la metodologías para genotipificar las muestras de pacientes que se utiliza en cada publicación. Entre los métodos utilizados se encuentran PCR con restricción enzimática, PCR alelo específico, PCR en tiempo real alelo específico, PCR en tiempo real alelo específico-sitio específico y Secuenciación. Adicionalmente, se lista el tamaño del amplicón producido y una referencia de si mencionan la presencia de la región homóloga en la publicación de donde se obtuvieron los iniciadores. No se presentan los resultados de la evaluación de los iniciadores publicados por Galmozzi (2010 y 2011) ya que no produjeron ningún amplicón *in silico*.

Es importante resaltar que varios autores presentan más de una reacción de PCR en la misma publicación, sin embargo en la Figura 17 solamente se presentan las parejas de iniciadores que produjeron una amplificación *in silico*. Adicionalmente algunos iniciadores no produjeron una amplificación *in silico*, por lo que no se procedió a analizarlos. Estos iniciadores presentan un guión en la casilla de amplicón. y los



iniciadores que presentan un asterisco fueron encontrados con una o más modificaciones en la secuencia de referencia del cromosoma 19.

Figura 16: A) Resolución en gel de agarosa al 2% de la digestión enzimática y PCR alelo específico utilizando los iniciadores diseñados anteriormente. B) Amplicón producido por los iniciadores diseñados, las bases subrayadas indican la posición de los iniciadores, las bases coloreadas de rojo indican las no homologías, la citosina magenta representa la posición del SNP rs12979860 y las bases tachadas representan los nucleótidos discontinuos en la secuencia homóloga.



Finalmente, la Figura 17 permite apreciar la colocación de los iniciadores sobre la alineación de la secuencia de interés y la homóloga. Esto permite visualizar la especificidad de cada uno de los iniciadores con amplificación *in silico* respecto a la homología IL28A-IL28B. Se puede observar que existe una alta variación entre la especificidad de los iniciadores publicados, respecto a la homología. Adicionalmente a esta especificidad, los iniciadores de Sharafi (2012) presentaron posible amplificación inespecífica de 495pb en el cromosoma 2; los de Lee (2013) presentaron la posible amplificación de 1426pb en el cromosoma 17 y 973pb en el cromosoma 4; y finalmente los iniciadores de Cielsa (2012) presentaron posible amplificación inespecífica en el cromosoma 16 de 3560pb.

Figura 17: Resumen del análisis de los iniciadores reportados para la genotipificación del rs12979860: Se visualiza la colocación de cada pareja de iniciadores respecto a la alineación de la región de interés (sobre cada flecha) y la región homóloga (debajo de cada flecha), junto con el tamaño de amplicón se que produjo con el PCR *in silico*, la tecnología para la cual fueron diseñados y el año de la publicación. \*Estos iniciadores fueron localizados manualmente mediante modificaciones de una base para que fuesen complementarios con la secuencia consenso que se utilizó, de esta forma se obtuvo una amplificación *in silico*.

	Frontal	Reverso	Autor, año	Mención Homología	Método de detección	Amplición
IL28B			Shafari, 2012	Si	PCF-RFLP	746
IL28A			Cook, 2012	No	qPCR	164
IL28B			Reynolds, 2012	Si	qPCR	593
IL28A			Lee, 2013	No	qPCR	66
IL28B			Ciesla, 2012	No	qPCR	190
IL28A			Younosi, 2012	No	qPCR	402
IL28B			Ezzikouri, 2013	No	qPCR	69
IL28A			Domagalski, 2013	No	PCR-RFLP	243

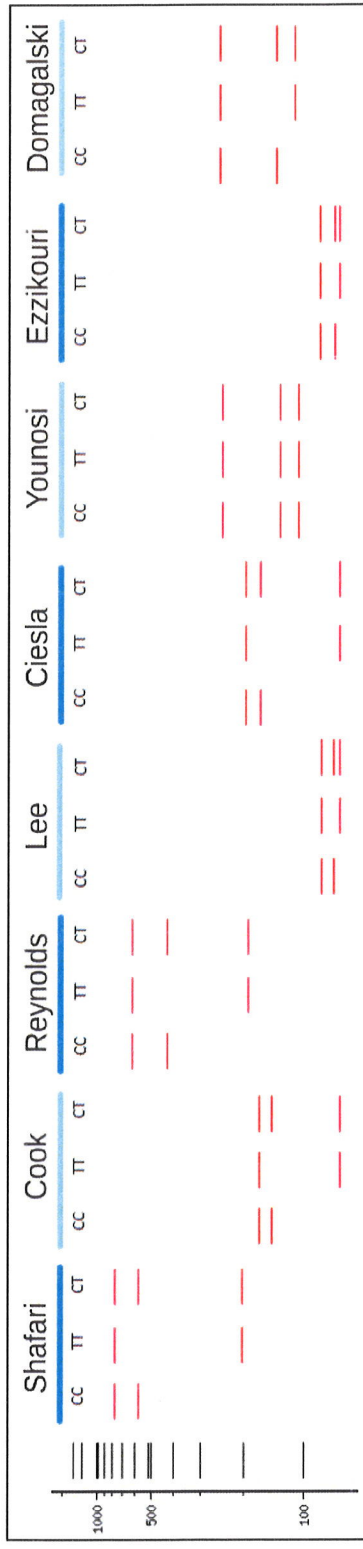
Para el diseño de los iniciadores internos del PCR alelo específico se decidió colocar el iniciador interno reverso (interno Reverse o iR) específico para el alelo C y el iniciador interno frontal (interno Forward o iF) específico para el alelo T. En el Cuadro 3 se observa la variación del cambio de energía libre de Gibbs dependiendo de si el nucleótido que da la especificidad se coloca en la posición 3' o una base antes del extremo 3' del iniciador. Se puede observar que la diferencia para el iniciador iR es de 3.71 y 6.75, respectivamente, mientras que para el iniciador iF la diferencia es de -1.34 y -3.3, respectivamente.

Cuadro 3: Posibles iniciadores internos para el PCR alelo específico con la interacción en la secuencia de interés con el alelo C y alelo T en versión gráfica y con delta G de la interacción. El nucleótido coloreado representa el alelo para el cual es específico dicho iniciador.

Iniciador	Interacción con alelo C	Delta G kcal/mol	Interacción con alelo T	Delta G kcal/mol
GGAGTGC AATTCAACCCCTGGTTCG <sup>+</sup>	5' GGAGTGC AATTCAACCCCTGGTTCG <sup>+</sup>       3' CCTCACGTTAAGTTGGGACCAAGC	-48.04	5' GGAGTGC AATTCAACCCCTGGTTCG <sup>+</sup>       3' CCTCACGTTAAGTTGGGACCAAGT	-44.33
GAGTGC AATTCAACCCCTGGTTCGC <sup>+</sup>	5' GAGTGC AATTCAACCCCTGGTTCGC <sup>+</sup>       3' CTCACGTTAAGTTGGGACCAAGC	-48.11	5' GAGTGC AATTCAACCCCTGGTTCGC <sup>+</sup>       3' CTCACGTTAAGTTGGGACCAAGT	-41.36
AGGGAGCTCCCCGAAGGCGT <sup>+</sup>	5' AGGGAGCTCCCCGAAGGCGT <sup>+</sup>       3' TCCTCGAGGGGCTTCGCGC	-44.98	5' AGGGAGCTCCCCGAAGGCGT <sup>+</sup>       3' TCCTCGAGGGGCTTCGCGA	-46.32
GGGAGCTCCCCGAAGGCGTG <sup>+</sup>	5' GGGAGCTCCCCGAAGGCGTG <sup>+</sup>       3' CCTCGAGGGGCTTCGCGC	-43.38	5' GGGAGCTCCCCGAAGGCGTG <sup>+</sup>       3' CCTCGAGGGGCTTCGCGAC	-46.68

La Figura 18 muestran los resultados de la amplificación multiplex de PCR alelo específico utilizando como iniciadores externos las parejas de iniciadores publicados que se evaluaron en la Figura 17.

Figura 18: Resolución de fragmentos producidos por el PCR multiplex alelo específico producidos utilizando las parejas de iniciadores que evaluados



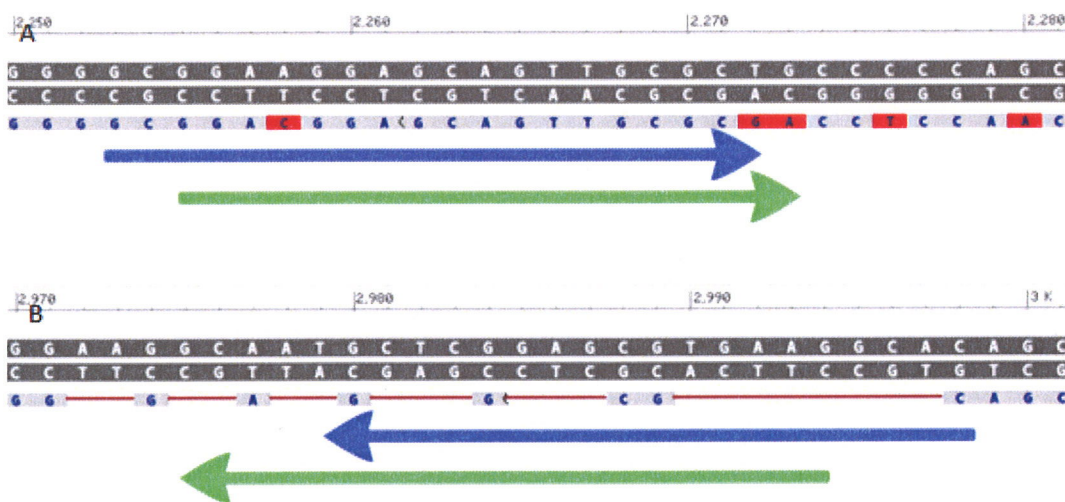


Cuadro 4: Iniciadores seleccionados para el PCR alelo específico diseñado y sus características fisicoquímicas como  $T_m$ , Tamaño y contenido de GC.

Nombre	Secuencia	$T_m$ (C)	Tamaño (nucleótidos)	Contenido GC
eForward (eF)	GGAAGGAGCAGTTGCGCTG	65.3	19	63.20%
eReverse (eR)	CTTCACGCTCCGAGCATTGC	65.6	20	60%
iForward (iF)	AGCTCCCCGAAGGCGTG	66.3	17	70.60%
iReverse (iR)	GTGCAATTCAACCCTGGTTCGC	66.3	22	54.50%

En el Cuadro 4 se resumen los iniciadores seleccionados para el PCR multiplex alelo específico. Nótese que los iniciadores internos no son exactamente los mismos que se muestran en el Cuadro 3. Los iniciadores externos se produjeron mediante modificaciones realizadas a los iniciadores publicados por Shafari *et al.* en el 2012.

Figura 19: Esquema gráfico de la colocación de A) El iniciador frontal y B) el iniciador reverso propuestos por Shafari en el 2012 (flechas azules) y la colocación de los iniciadores externos seleccionados (flechas verdes) para el PCR alelo específico.



En la Figura 19 se muestran las modificaciones que se les realizó a los iniciadores publicados por Sharafi *et al.*, en ésta se representan en la parte superior los iniciadores frontales y en la parte inferior se observan los iniciadores reversos respecto a la alineación. En esta figura se puede ver que los iniciadores que aquí se proponen producen un amplicón más corto; sin embargo, la diferencia de tamaño no es significativa para la resolución de los resultados.



Cuadro 5: Información de las estructuras de Hairpin de los iniciadores seleccionados para el PCR alelo específico.

Iniciador	Delta G (kcal/mol)	Tm (C)
eF	-1.58	50.5
	-1.57	44.2
	-1.35	39.3
	-0.97	34.6
	-0.76	35
eR	-1.7	46.1
	-1.53	44.2
	-0.91	40
iF	-1.26	53.1
	-0.85	36.6
	-0.35	32.4
iR	-0.065	34.7

En el Cuadro 5 se puede observar que la Tm mayor que se presenta es de 50.5C, sin embargo la mayoría de Tm se encuentran alrededor de 30-40C. Adicionalmente, todos los delta G son muy cercanos a 0, estos varían entre -1.58 y -0.35.

En las Figuras 20 y 21 se observan las estructuras de homodímeros y heterodímeros, respectivamente, que se producen al interactuar los cuatro iniciadores seleccionados. En estas figuras se presentan únicamente aquellas estructuras con un delta G menor a -4 kcal/mol. La interacción más estable que se presenta es la formación de un heterodímero entre el iniciador iF y eR, esta consta de 5 nucleótidos consecutivos y tiene un delta G -10.05 kcal/mol.

Figura 20: Estructuras de homodímeros producidas por los cuatro iniciadores seleccionados para el PCR alelo específico.

Delta G kcal/mol	Bases Apareadas	
<b>Homodímeros eF</b>		
-9.89	4	5' GGAAGGAGCAGTTGCGCTG         3'           GTCGCGTTGACGAGGAAGG
-5.09	3	5' GGAAGGAGCAGTTGCGCTG :       : : : : 3'           GTCGCGTTGACGAGGAAGG
-4.74	3	5' GGAAGGAGCAGTTGCGCTG             : : : 3'           GTCGCGTTGACGAGGAAGG
<b>Homodímeros eR</b>		
-6.31	4	5' CTTACGCTCCGAGCATTGC :         : : : : : 3'   CGTTACGAGCCTCGCACTTC
-5.09	3	5' CTTACGCTCCGAGCATTGC       : : : 3'           CGTTACGAGCCTCGCACTTC
<b>Homodímero iF</b>		
-6.34	4	5'                   AGCTCCCCGAAGGCGTG         3' GTGCGGAAGCCCCTCGA
<b>Homodímeros iR</b>		
-7.05	4	5'                           GTGCAATTCAACCCTGGTTCGC         3' CGCTTGGTCCCAACTTAACGTG
-6.36	4	5' GTGCAATTCAACCCTGGTTCGC         : : : : 3'           CGCTTGGTCCCAACTTAACGTG
-5.36	4	5'                           GTGCAATTCAACCCTGGTTCGC : :         : : 3' CGCTTGGTCCCAACTTAACGTG

Figura 21: Estructuras de heterodímero que se producen entre los cuatro iniciadores que seleccionaron para el PCR alelo específico.

Delta G kcal/mol	Bases Apareadas		Delta G kcal/mol	Bases Apareadas	
<b>Heterodímeros eF-eR</b>					
-9.38	5	5' GGAAGGAGCAGTTGCGCTG :           3' CTTACGAGCCTTGCACCTTC	-7.04	4	5' GGAAGGAGCAGTTGCGCTG :   :   :       : 3' CGCTTGGTCCCAACTTAACGTG
-6.75	3	5' GGAAGGAGCAGTTGCGCTG :   :   :       3' CTTACGAGCCTTGCACCTTC	-6.75	3	5' GGAAGGAGCAGTTGCGCTG :   :   :       3' CGCTTGGTCCCAACTTAACGTG
-5.12	4	5' GGAAGGAGCAGTTGCGCTG :   :   :       : 3' CTTACGAGCCTTGCACCTTC	-5.24	4	5' GGAAGGAGCAGTTGCGCTG :   :   :       : 3' CGCTTGGTCCCAACTTAACGTG
-5.09	3	5' GGAAGGAGCAGTTGCGCTG :       :   : 3' CTTACGAGCCTTGCACCTTC	-5.09	3	5' GGAAGGAGCAGTTGCGCTG :   :   :       3' CGCTTGGTCCCAACTTAACGTG
-5.09	3	5' GGAAGGAGCAGTTGCGCTG :       :   : 3' CTTACGAGCCTTGCACCTTC	-4.67	3	5' GGAAGGAGCAGTTGCGCTG :       :   : 3' CGCTTGGTCCCAACTTAACGTG
-4.74	3	5' GGAAGGAGCAGTTGCGCTG :       :       3' CTTACGAGCCTTGCACCTTC	-10.05	5	<b>Heterodímeros eR-IF</b> 5' CTTACGCTCCGAGCATTGC :         : 3' GTGCGGAAGCCCTCGA
-4.64	3	5' GGAAGGAGCAGTTGCGCTG :       :   :   : 3' CTTACGAGCCTTGCACCTTC	-6.31	4	5' CTTACGCTCCGAGCATTGC :   :         3' GTGCGGAAGCCCTCGA
<b>Heterodímeros eF-IF</b>					
-9.38	5	5' GGAAGGAGCAGTTGCGCTG :           3' GTGCGGAAGCCCTCGA	-5.12	4	5' CTTACGCTCCGAGCATTGC :         : 3' GTGCGGAAGCCCTCGA
-6.75	3	5' GGAAGGAGCAGTTGCGCTG :       :       3' GTGCGGAAGCCCTCGA	-4.74	3	5' CTTACGCTCCGAGCATTGC :       :       3' GTGCGGAAGCCCTCGA
-4.74	3	5' GGAAGGAGCAGTTGCGCTG :   :   :       3' GTGCGGAAGCCCTCGA	-8.51	5	<b>Heterodímeros eR-IR</b> 5' CTTACGCTCCGAGCATTGC :   :   :         3' CGCTTGGTCCCAACTTAACGTG
-4.64	3	5' GGAAGGAGCAGTTGCGCTG :       :       3' GTGCGGAAGCCCTCGA	-5.19	3	5' CTTACGCTCCGAGCATTGC :       :       3' CGCTTGGTCCCAACTTAACGTG
			-5.09	3	5' CTTACGCTCCGAGCATTGC :   :   :       3' CGCTTGGTCCCAACTTAACGTG
			-7.13	4	<b>Heterodímeros IF-IR</b> 5' AGCTCCCGAAGGCGTG :         : : 3' CGCTTGGTCCCAACTTAACGTG
			-6.75	3	5' AGCTCCCGAAGGCGTG :       :       3' CGCTTGGTCCCAACTTAACGTG
			-4.67	3	5' AGCTCCCGAAGGCGTG :       :       : 3' CGCTTGGTCCCAACTTAACGTG

Figura 22: Resultados teóricos del PCR alelo específico con los cuatro iniciadores seleccionados de los tres alelos posibles en gel de agarosa al 2%.



En la Figura 22 se ilustran las bandas que se producirían al procesar los distintos alelos del rs12979860 en el PCR multiplex alelo específico. En este se puede observar una banda superior, que representa un fragmento de 742pb, una inferior de 214pb y una intermedia de 561pb.

Al observar la Figura 23 se puede comparar la factibilidad de utilizar las parejas de iniciadores evaluados para aplicar un PCR con restricción de enzimas. Por la alta variación del tamaño de las bandas fue necesario establecer una marca de 40pb, para facilitar la visualización de las bandas de menor tamaño. Es importante resaltar que todos estos resultados fueron producidos en geles de agarosa al 2% in silico.

Como se mencionó anteriormente, otro posible método de detección es mediante la restricción enzimática de un producto de PCR. Para esta aplicación se ha reportado en la literatura la utilización de la enzima Bstul. En la Figura 24A se puede observar un ejemplo de dicha restricción, utilizando los iniciadores eF y eR, sin embargo, las enzimas AccII, Bsh1236I, BspFNI, BstFNI, FnuDII, MvnI y Thal también producirían un patrón de fragmentación igual al que se presenta en dicha figura. Adicionalmente las enzimas Hpy166II, Hpy8I y MjaIV también pueden ser utilizadas para el diagnóstico de dicho SNP pero estas producen un patrón de fragmentación distinto, el cual se presenta en la Figura 24B.



Figura 23: Resultados de la restricción enzimática con Bstul de los amplicones producidos por las parejas de iniciadores publicados. La franja lila representa la migración de un fragmento de 40pb en el gel

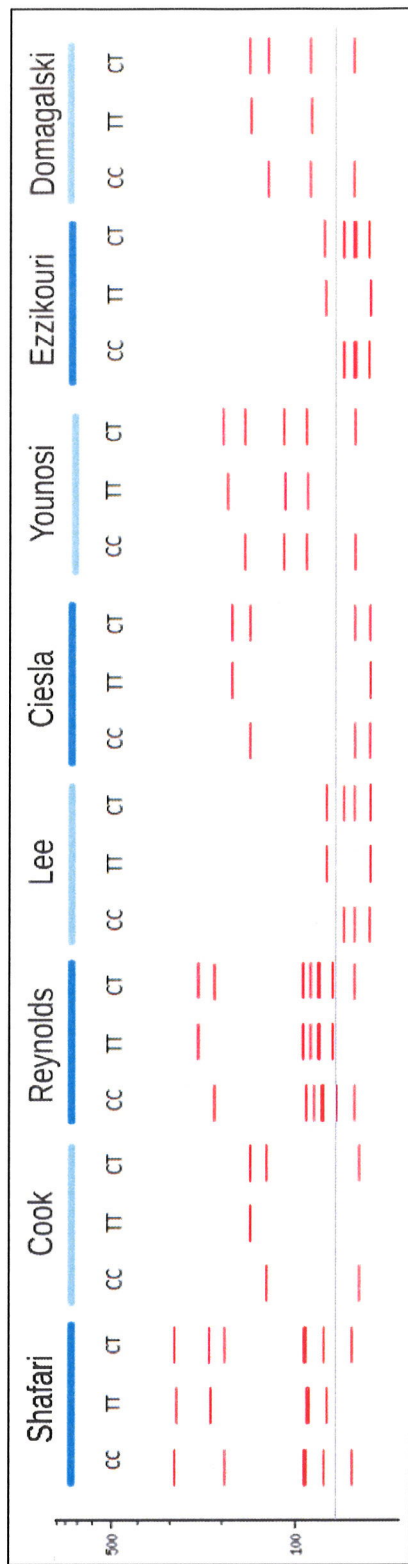
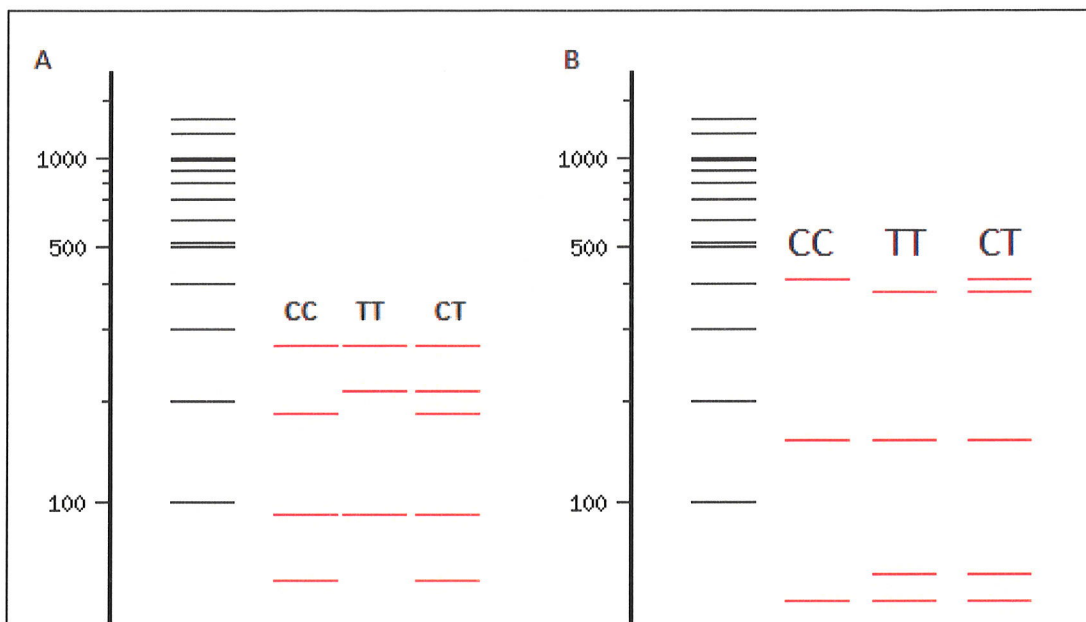


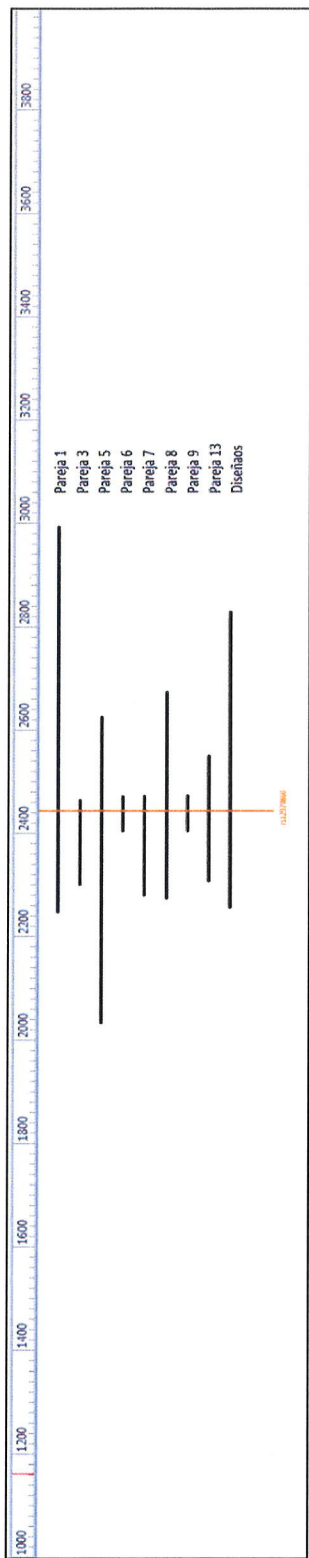
Figura 24: Esquema de la electroforesis en gel de agarosa al 2% de los distintos genotipos del rs12979860 con las enzimas con A) reconocimiento de 'CGCG' y B) 'GTNNAC'



La posición de los iniciadores respecto al SNP es un factor que determina las tecnologías en las cuales se puede utilizar una pareja de iniciadores para realizar la detección del SNP. En la Figura 25 se puede observar tanto la dimensión de los amplicones generados por cada pareja de iniciadores como la colocación relativa del SNP respecto a los iniciadores utilizados. Note la gran variedad de tamaños de amplicones, aquellos más cortos son utilizados para la detección por qPCR, mientras que los que se utilizan para FRLP-PCR producen amplicones de tamaños bastante mayores.

Los Cuadros 6 y 7 muestran el desarrollo de la evaluación de costos de la genotificación del SNP rs12979860 en asPCR, FRLP-PCR, qPCR y Secuenciación al extraer ADN mediante el Kit Maxwell semiautomatizado. Como se puede observar, los costos en las distintas tecnologías varían entre Q.187 hasta Q.802, siendo la aplicación más económica el asPCR y la más costosa la secuenciación.

Figura 25: Esquema de las regiones amplificadas con las distintas parejas de iniciadores evaluados según la numeración de Cuadro 4 y los iniciadores diseñados en este estudio. La línea naranja muestra la posición del SNP rs12979860.



Cuadro 6: Evaluación de los costos de implementar la detección del rs12979860 por PCR alelo específico, PCR seguido de restricción enzimática y PCR en tiempo real.

Rubro	Costo (Q)	Unidades		Costo por unidad		Unidades por reacción		Costo por reacción
<b>asPCR</b>								<b>187.41</b>
Kit Maxwell	4423.00	48	Muestras	92.1458333	Q/Muestra	1	Muestra	92.15
dNTPs	300.00	200	rxn	1.5	Q/rxn	1	rxn	1.50
Polymerase	818.00	100	u	8.18	Q/u	1.5	u	12.27
Primer eF	83.49	0.05	umol	1669.8	Q/umol	0.000025	umol	0.04
Primer eR	85.39	0.05	umol	1707.75	Q/umol	0.000025	umol	0.04
Primer iF	79.70	0.05	umol	1593.9	Q/umol	0.000025	umol	0.04
Primer iR	89.18	0.05	umol	1783.65	Q/umol	0.000025	umol	0.04
Buffer TAE	932.40	40000	mL	0.02331	Q/mL	2.5	mL	0.06
Agarosa	1659.00	100	g	16.59	Q/g	0.005	g	0.08
Escalera	1887.00	100	u	18.87	Q/u	0.1	uL	1.89
Termociclador				1.72	Q/h	2.5	h	4.30
Laboratorista	8000.00	160	h	50	Q/h	1.5	h	75.00
<b>RFLP</b>								<b>243.68</b>
Kit Maxwell	4423.00	48	Muestras	92.1458333	Q/Muestra	1	Muestra	92.15
dNTPs	300.00	200	rxn	1.5	Q/rxn	1	rxn	1.50
Polymerase	818.00	100	u	8.18	Q/u	1.5	u	12.27
Primer eF	83.49	0.05	umol	1669.8	Q/umol	0.000025	umol	0.04
Primer eR	85.39	0.05	umol	1707.75	Q/umol	0.000025	umol	0.04
Enzima	1269.00	200	u	6.345	Q/u	1	u	6.35
Buffer TAE	932.40	40000	mL	0.02331	Q/mL	2.5	mL	0.06
Agarosa	1659.00	100	g	16.59	Q/g	0.005	g	0.08
Escalera	1887.00	100	u	18.87	Q/u	0.1	uL	1.89
Termociclador				1.72	Q/h	2.5	h	4.30
Laboratorista	8000.00	160	h	50	Q/h	2.5	h	125.00
<b>qPCR</b>								<b>193.70</b>
Kit Maxwell	4423.00	48	Muestras	92.1458333	Q/Muestra	1	Muestra	92.15
qPCR TaqMan kit	1382.40	200	rxn	6.912	Q/rxn	1	rxn	6.91
Primer eF	81.59	0.05	umol	1631.85	Q/umol	0.000025	umol	0.04
Primer eR	81.59	0.05	umol	1631.85	Q/umol	0.000025	umol	0.04
Sonda C	2656.50	0.006	umol	442750	Q/umol	0.0000125	umol	5.53
Sonda T	2656.50	0.006	umol	442750	Q/umol	0.0000125	umol	5.53
qPCR				25	Q/h	2	h	50.00
Laboratorista	8000.00	160	h	50	Q/h	0.67	h	33.50



Cuadro 7: Evaluación de los costos de implementar la detección del rs12979860 por Secuenciación.

Rubro	Costo (Q)	Unidades		Costo por unidad		Unidades por reacción		Costo por reacción
<b>Secuenciación</b>								<b>802.82</b>
Kit Maxwell	4423.00	48	Muestras	92.1458333	Q/Muestra	1	Muestra	92.15
dNTPs	300.00	200	rxn	1.5	Q/rxn	1	rxn	1.50
Polymerase	818.00	100	u	8.18	Q/u	2.5	u	20.45
Primer eF	83.49	0.05	umol	1669.8	Q/umol	0.000025	umol	0.04
Primer eR	85.39	0.05	umol	1707.75	Q/umol	0.000025	umol	0.04
Buffer TAE	932.40	40000	mL	0.02331	Q/mL	2.5	mL	0.06
Agarosa	1659.00	100	g	16.59	Q/g	0.005	g	0.08
Escalera	1887.00	100	u	18.87	Q/u	0.1	uL	1.89
QIAquick PCR purification kit				23	Q/u	1	u	23.00
Big Dye Terminator kit	11931.46	100	rxn	119.3146	Q/rxn	1	rxn	238.63
Primer	83.49	0.05	umol	1669.8	Q/umol	0.0000032	umol	0.01
Big Dye X-Terminator kit 3.1	2590.95	100	rxn	25.9095	Q/rxn	1		51.82
POP 6	5878.66	7000	uL	0.83980857	Q/uL	40	uL	33.59
Arreglo del capilares 50cm	9079.29	200		45.39645		1		45.40
Buffer de corrida 25mL 10X	1306.37	250	mL	5.22548	Q/mL	0.5	mL	2.61
formamida				0.03		40	uL	1.20
Termociclador				1.72	Q/h	5	h	7.74
Seccuenciador				32.61	Q/h	1	h	32.61
Laboratorista	8000.00	160	h	50	Q/h	5	h	250.00

## VII. DISCUSIÓN

El impacto del SNP rs12979860 en el desarrollo de la respuesta a la hepatitis C mediante el tratamiento de ribavirina es el factor más importante en la farmacogenética del tratamiento de la Hepatitis C, por lo que el mecanismo de acción de dicho SNP en el cambio de la respuesta del paciente al tratamiento tradicional de Ribavirina e interferón es hasta el momento un factor de mucho interés. En la Figura 9 se puede observar que el SNP rs12979860 se encuentra fuera del marco de lectura de la interleucina 28B (codificada por el IFNL3 de 39243632 a 39245077), ya que como se mencionó la posición del SNP es 39248147 en la secuencia de referencia del cromosoma 19. Esto es un factor importante, ya que significa que el SNP no representa una alteración en la cadena de aminoácidos, la cual podría modificar su actividad en sí, sino que el mecanismo de acción de dicha mutación podría estar relacionado a un cambio en la regulación de los genes del interferón, específicamente la IL28B, los cuales como se mencionó con anterioridad se expresan en respuesta a una infección viral.

En la Figura 10 se puede observar la alta homología que existe dentro de los genes de la familia del Interferón lambda, ya que, se encontraron varias regiones homologas a lo largo del fragmento analizado. En esta figura se puede observar la presencia de regiones con mayor homología que la región con la que se trabajó, como las regiones entre 3.5- 5Kb y 5-7.5Kb de la alineación, sin embargo, debido a que estas homologías no intervienen con la genotipificación del rs12979860 no se profundizó más en dichas regiones.

Al realizar una alineación de la región de interés y la región homóloga, en la Figura 11, se pueden comparar los resultados obtenidos utilizando distintas herramientas bioinformáticas para procesar los datos. Entre estos tres resultados la alineación producida por Blast de NCBI resultó presentar la forma más sencilla de representar la distribución de la homología entre ambas secuencias debido a que permite determinar fácilmente la posición de las no homologías, mientras que los resultados de la herramienta CLC Sequence viewer a pesar de que parecieran mostrar la misma información pueden dar resultados engañosos si se desea analizar la secuencia a un nivel más detallado, es decir a nivel de nucleótidos. Esto es debido a que los resultados se presentan como porcentaje de homología de una subregión y a pesar de que dicho porcentaje sea bajo, si la subregión es lo suficientemente grande esta puede tener segmentos altamente homólogos muy cortos seguida de segmentos poco homólogos más largos y de esta forma no es posible visualizar dicha homología con esta herramienta. Finalmente la herramienta Clustal W de EBI permite un buen análisis a nivel de nucleótidos, sin embargo, los resultados podrían ser más cómodos si se utilizara algún código de color que resalte las homologías de las no homologías. Por esta razón se trabajó con la herramienta de Blast de NCBI.

Al calcular el porcentaje de homología de ambas regiones se estableció un 93.6% de homología, lo cual es cercano al 95% publicado por Reynolds *et al.* en el 2012. Esta diferencia en el porcentaje de homología puede haberse producido por una diferencia en la limitación de las regiones de trabajo, ya que en la Figura 11A se observa que en los extremos 3' y 5' de la alineación existe un incremento en las no homologías. A pesar de

este factor, la función de la alineación es brindar información referente a la genotipificación del SNP rs12979860, por lo que, a pesar de abarcar un marco mayor que el publicado por Reynolds *et al.*, la alineación presentada proporciona la información de interés para el diseño y análisis de iniciadores para la detección del SNP.

El factor crítico en el impacto de la región homóloga en la genotipificación del rs12979860 es que el nucleótido homólogo a dicho SNP es convencionalmente una Timina. La amplificación inespecífica de ambas regiones un PCR producirá dos bandas de aproximadamente el mismo tamaño, las cuales no se podrán diferenciar en un gel de agarosa. Por esta razón la amplificación de la región homóloga por PCR seguida de la digestión con enzimas de restricción o por asPCR producirán un patrón de bandas típico de un alelo T el cual puede crear interferencia en el diagnóstico, convirtiendo un genotipo CC en un genotipo CT.

En el diseño de los iniciadores específicos se consideró dos factores para la colocación de los iniciadores. El primero fue que la región de colocación fuera rica en no homologías, estas no homologías se establecieron más cercanas al extremo 3' que al extremo 5' ya que el inicio de la traducción ocurre en este extremo, por lo que la hibridación del iniciador en este extremo es más importante que la hibridación en el extremo 5'.

El segundo factor fue la colocación del iniciador respecto al SNP y respecto al posible otro iniciador. Debido a que se desea evaluar la implementación de estos iniciadores en un diagnóstico por PCR de punto final seguido de restricción enzimática o por un PCR multiplex alelo específico, se priorizó la producción de un amplicón grande en el cual el SNP se encontrase en una posición asimétrica para que luego de la restricción enzimática o del asPCR se pudiera presentar una buena resolución en un gel de agarosa por cualquiera de los dos métodos de detección mencionados. Adicionalmente, la producción de un amplicón grande permite la verificación del método por medio de secuenciación, tanto de la genotipificación en sí como de amplificación específica por el PCR optimizado.

En la Figura 12A se observa que la colocación de los iniciadores sobre la alineación permite la producción de un fragmento de suficiente tamaño para obtener una buena resolución en el SNP utilizando secuenciación y la posición asimétrica del SNP dentro del amplicón permite la buena resolución de los fragmentos producidos por restricción enzimática y PCR alelo específico. Adicionalmente, las posiciones de las no homologías en los iniciadores mostrados por la Figura 12B y C permiten la amplificación por PCR sitio específico. Sin embargo un factor muy importante en dicha amplificación específica es la optimización de la reacción de PCR. Por esta razón la Temperatura media de ambos iniciadores debe ser muy cercana, debido a que mientras más se aleja el programa de termociclado de las temperaturas de alineación óptimas de los iniciadores se incrementa la probabilidad de la amplificación inespecífica. Por esta razón se cambió el tamaño y colocación de los iniciadores hasta obtener una diferencia de  $T_m$  de aproximadamente 2 grados o menor (ver Cuadro 1).

En el Cuadro 1 se observa también el delta G para la reacción de hibridación de los iniciadores con su secuencia complementaria, para ambos iniciadores esta fue menor a  $-44\text{kcal/mol}$ . Esto es importante como



parámetro de comparación de las estructuras de homodímeros (Figuras 13 y 14) y heterodímeros (Figura 15). La interacción con energía libre de Gibbs más fuerte calculada fue de  $-9.88\text{kcal/mol}$  de un homodímero del iniciador Forward que involucra la interacción de 4 nucleótidos consecutivos, seguida de  $-9.73\text{kcal/mol}$  de un homodímero del iniciador Reverse, la cual involucra la interacción entre 6 nucleótidos. A pesar de que la primera interacción involucra menos nucleótidos consecutivos que la segunda esta presenta una mayor estabilidad debido a que son nucleótidos internos, a comparación de la interacción del iniciador Reverse en el cual interactúan los nucleótidos de las orillas. Adicionalmente, en la interacción del iniciador Forward existen otros nucleótidos aislados que se alinean para aumentar la estabilidad de la interacción. Respecto a las interacciones unimoleculares, estas presentan una  $T_m$  al menos 10 grados menor a la  $T_m$  promedio de ambos iniciadores, por lo que no debería de afectar significativamente la eficiencia de la reacción.

Al analizar los datos de la Figura 17 es interesante observar que seis de las ocho publicaciones analizadas proponen la detección del SNP por PCR en tiempo real, ya sea utilizando una sonda sobre el SNP o utilizando dos iniciadores reversos con los alelos C o T marcados con distinto fluoróforo (Cook, 2012). A pesar de que el PCR en tiempo real permite eficientemente la identificación de dicho SNP, esta técnica es realmente aprovechada para determinar mutaciones en poblaciones mixtas, es decir cuando en una muestra existen células con distintos genotipos o cuando dicho genotipo esperado se encuentra en una menor proporción en la muestra. Sin embargo la mutación del rs12979860 es una mutación somática, por lo que todas las células de la muestra presentarán el mismo genotipo, cualquiera que sea. Por esta razón, a pesar de que la mayoría de las publicaciones utilizan esta tecnología para determinar el genotipo del paciente, no se recomienda la implementación de este método para el diagnóstico de la población guatemalteca, debido a que es un sobrecargo por la avanzada tecnología requerida que no es justificable.

Otro factor interesante que se observa en la Figura 17 es que solamente 2 publicaciones consideran la existencia de la región homóloga para el diseño y optimización del PCR. Ambas publicaciones son del año 2012 y las publicaciones posteriores producidas por Lee *et al.* y Domagalski *et al.* tampoco consideran la existencia de esta homología. Sin embargo, el no considerar la existencia de la homología no necesariamente significa que los iniciadores que diseñaron no sean específicos. En la actualidad las herramientas de diseño de iniciadores facilitan estos procesos considerando la presencia de regiones similares en el genoma humano para la colocación de los iniciadores. Por esta razón los resultados de la Figura R9 muestran que en varios de los iniciadores se encuentran algunas no homologías, no necesariamente segregadas al extremo 3'. Sin embargo, al buscar una amplificación específica es importante recordar que esta no depende únicamente de el diseño de los iniciadores, sino que uno de los factores más importantes es la optimización de la reacción de PCR, ya que las condiciones que se establezcan en este proceso dictarán si es estable o no la interacción con la región homóloga, por lo que es importante que el autor considere esto al realizar o verificar dicha optimización. A continuación se realiza un análisis de los resultados producidos por los iniciadores publicados.



Los iniciadores propuestos por Sharafi, *et al.* (2012) produce un amplicón grande (746pb) debido a que fue diseñada para la implementación de detección por FRLP-PCR. Los autores de la publicación consideraron la región homóloga en el diseño de sus iniciadores, lo cual se evidencia claramente en el iniciador reverso ya que se encuentra colocado en una región de la alineación en donde existen varias discontinuidades, básicamente esta región existe solo parcialmente en la región homóloga. Respecto al iniciador frontal se encuentra colocado sobre una no homología en el extremo 3', lo cual le proporciona cierto grado de especificidad, sin embargo debido a la alta especificidad del iniciador reverso y a la aplicación para la cual se diseñaron estos iniciadores, el producto de la reacción de PCR posiblemente será muy específico.

Respecto los iniciadores propuestos por Cook *et al.*(2012), estos son iniciadores para un PCR alelo específico en tiempo real, por esta razón el amplicón que producen es bastante corto (164pb). El iniciador frontal está colocado sobre 2 no homologías cercanas al extremo 3', por lo que esto le puede dar un poco de especificidad, sin embargo el iniciador reverso está colocado sobre el SNP en el extremo 3' pero no presenta ninguna otra no homología que le dé especificidad respecto a la región homóloga. La amplificación de este PCR podría ser específica en base al diseño de los iniciadores.

Los iniciadores propuestos por Reynolds *et al.* (2012) son otro ejemplo de iniciadores diseñados considerando la región homóloga. Estos, a pesar de que están diseñados para qPCR, producen un amplicón de 593pb. Esto es debido a que las regiones muy cercanas al SNP no contienen muchas no homologías, por lo que la colocación de los iniciadores debió ser más alejada del SNP. A pesar de que ambos iniciadores no son tan específicos como la pareja 1 se observan varias no homologías en el extremo 3' del iniciador reverso y una en el iniciador frontal. En combinación y con la optimización correcta estos iniciadores podrían producir una amplificación específica.

El caso de las parejas de iniciadores diseñados por Lee, *et al.* (2013) y Ciesla, *et al.* (2012) es muy similar, ambos producen un fragmento corto (66 y 99pb, respectivamente) ya que están diseñados para qPCR, sin embargo ninguno de los cuatro iniciadores presenta una no homología en los 10 nucleótidos más cercanos al extremo 3'. El iniciador frontal de Lee, *et al.* (2013) no está colocado sobre ninguna no homología, mientras que el iniciador frontal de Ciesla, *et al.* (2012) se encuentra sobre una sola no homología a la mitad del iniciador. El caso de los iniciadores diseñados por Ezzikouri, *et al.* (2013) es parecido ya que el iniciador frontal no se encuentra sobre ninguna no homología y el reverso tiene una no homología hasta el noveno nucleótido desde el extremo 3', por lo que la aplicación de la amplificación de dichos iniciadores en la genotipificación del SN rs12979860 es cuestionable.

La pareja de iniciadores propuestos por Younosi, *et al.* (2012) también está diseñada para PCR en tiempo real pero esta si presenta algunas no homologías cercanas al extremo 3' de ambos iniciadores, por lo que podría producir una amplificación específica si se optimiza adecuadamente la reacción de PCR.

La pareja de iniciadores diseñados por Domagalski, *et al.* (2013) produce un amplicón de 243 pb, sin embargo, por la posición de los iniciadores éstos podrían aplicarse para digestión enzimática o asPCR (ver

resultados en Figuras 16 y 23), adicionalmente, a pesar de que no está diseñado considerando la región homóloga si presenta varias no homologías cerca del extremo 3', por lo que su aplicación podría producir amplificación específica.

Para el diseño de los iniciadores internos para el PCR alelo específico se evaluó inicialmente el cambio en la energía libre de Gibbs entre la interacción con el alelo para el cual fue diseñado y el otro alelo para iniciadores con el SNP en la última base en el extremo 3' y en la penúltima base. Los resultados muestran que hay una mayor pérdida de estabilidad si se coloca el SNP en la penúltima base (Cuadro 3). Esto se puede explicar de dos formas: Una es que adicional a las parejas de nucleótidos descritas por Watson y Crick existen otras interacciones entre nucleótidos que involucran puentes de hidrógeno menos estables o la interacción mediada por moléculas de agua, estas interacciones son más estables en los extremos ya que involucra un cambio tridimensional de la hebra de ADN. La segunda forma de explicar este evento es mediante la teoría del vecino más cercano, lo cual significa que la estabilidad de cada pareja de nucleótidos depende de la estabilidad de los nucleótidos vecinos, por esta razón desestabilizar al nucleótido del extremo solamente afecta a un nucleótido vecino, mientras que si se desestabiliza el siguiente se está desestabilizando indirectamente el nucleótido terminal 3' y el tercero desde el extremo 3'.

Debido a que el iniciador reverso de Shafari *et al.* (2012) presentó la mejor colocación sobre la alineación se decidió utilizar este como iniciador reverso externo para la implementación del PCR alelo específico. Sin embargo, ya que uno de los amplicones del PCR alelo específico utilizará otro iniciador reverso sobre el SNP es necesario incrementar la especificidad del iniciador frontal. Debido a que el iniciador frontal de Shafari solamente presentó una no homología en el extremo 3' se prefirió utilizar el iniciador frontal diseñado en este estudio. Se modificó el tamaño de ambos iniciadores para que las características fisicoquímicas de éstos fuesen compatibles con los iniciadores internos en un solo PCR multiplex (Ver figuras 2-6 de Anexos). Sin embargo esto no produjo un juego de iniciadores que funcionasen apropiadamente. Por esta razón se evaluó nuevamente el diseño de iniciadores externos, para esto se modificó la colocación y el tamaño de los iniciadores publicados por Shafari *et al.* Para producir una pareja de iniciadores externos que fueran compatibles con los iniciadores internos que se diseñaron.

En la Figura 19 se observa que el iniciador frontal fue corrido para abarcar dos no homologías en el extremo 3'. A pesar de que la colocación en el nucleótido siguiente hubiese producido una mayor especificidad respecto a la región homóloga no se utilizó ya que alteraba negativamente la  $T_m$ . Adicionalmente, Se modificó el tamaño de ambos iniciadores para que las  $T_m$  de ambos estuvieran por debajo de 63.6 C, más adelante se explica el porqué.

Los iniciadores internos fueron modificados para disminuir la estabilidad de estructuras de heterodímeros que se producían, especialmente entre los iniciadores iR y eR. Adicionalmente, estas modificaciones ayudaron a establecer la diferencia menor posible entre la  $T_m$  de ambos, esto se logró ya que la  $T_m$  de ambos iniciadores internos es de 63.6C, por lo que al optimizar la PCR se partirá de esta temperatura para

asegurar la especificidad de ambos iniciadores respecto al alelo que detectan. Es importante resaltar que a pesar de que estos iniciadores son alelo específicos, estos no presentan especificidad respecto a la región homóloga, por lo que toda la especificidad del PCR multiplex recae en la iniciadores externos. Por esta razón se diseñó el PCR de forma en que la  $T_m$  de los iniciadores externos fuera menor a la de los internon, ya que la optimización se realizaría alrededor de la  $T_m$  de los iniciadores internos, por lo que el porcentaje de iniciadores externos interactuando inespecíficamente a una temperatura mayor a su  $T_m$  disminuye.

Los resultados de las Figuras 23 y 18 presentan la genotipificación por restricción enzimática y asPCR, respectivamente. En estos geles se puede observar que los iniciadores diseñados para qPCR no producen buenos resultados en estas dos metodologías. Adicionalmente, dado a los resultados recopilados en estas dos figuras se puede observar que la mayoría de iniciadores, a pesar de producir fragmentos de distintos tamaños presentan un patrón de fragmentación similar, esto es debido a que tanto las enzimas con sitio de reconocimiento 'CGCG' como las que reconocen 'GTNNAC' presentan más de un sitio de restricción los fragmentos amplificados (Ver Figura 1 de Anexos). Por esta razón, las restricciones realizadas con las enzimas que reconocen la secuencia 'CGCG' sobre las distintas parejas de iniciadores presentan al menos un fragmento de aproximadamente 89pb, uno de aproximadamente 62 y uno de aproximadamente 25pb. Este último fragmento se produce únicamente en presencia del alelo C. Sin embargo al contener más de un sitio de reconocimiento en la secuencia amplificada esto produce un incremento en las bandas que se observan, haciendo la optimización de la resolución de la electroforesis una prioridad en la implementación de la detección por esta técnica.

Al observar los patrones de digestión de las Figuras 16 y 23 se observa que la pareja de iniciadores que produce la resolución más sencilla es la pareja 3, ya que por su corto tamaño solamente contiene un sitio de restricción, por lo que la distinción de 1 a 2 bandas parece un diagnóstico sencillo, sin embargo es necesario resaltar que la banda inferior es de solamente 21pb, por lo que es posible que esta banda no sea tan fácilmente visualizable. Adicionalmente, se debe considerar que la especificidad de la pareja de iniciadores 3 respecto a la región homóloga es bastante baja. Por estas razones se descarta la pareja 3 como iniciadores para diagnóstico por FRP-PCR.

La resolución de los fragmentos producidos por la digestión enzimática de los amplicones de los iniciadores presentados en el Cuadro 1 es similar a la de las parejas 1,5 y 13, por lo que todos estos podrían ser adecuados para esta aplicación. Respecto a su especificidad, estos iniciadores parecen producir una amplificación específica, la cual es necesaria para producir un diagnóstico confiable, sin embargo, debido a que la diferencia entre las dos bandas que producen el diagnóstico homocigoto CC o TT es de solamente 25pb se vuelve importante asegurarse de que la migración de las muestras en todo el gel sea uniforme, por lo que es recomendable utilizar un marcador molecular en ambos extremos del gel.



Para implementar el qPCR se observa que existe una publicación que presenta un juego de iniciadores muy específico, por lo que si se desea implementar el diagnóstico en esta tecnología se recomienda utilizar los iniciadores publicados por Reynolds en 2012 con la sondas de dicha publicación.

Ya sea que se desee establecer el diagnóstico rutinariamente por secuenciación o que se utilice ésta como parte del proceso de verificación de la especificidad durante la optimización de la reacción de PCR la secuenciación es un paso clave en la verificación de la especificidad de los resultados al implementar el diagnóstico del SNP rs12979860 por cualquiera de los métodos mencionados anteriormente.

La Figura 25 muestra el tamaño de amplicón producido por cada una de las parejas de iniciadores al igual que la posición relativa del SNP respecto a los iniciadores. Estos factores son importantes al elegir la química que se utilizará para la reacción de secuenciación. Debido al análisis anterior respecto a los iniciadores se recomienda que si se desea genotipificar el rs12979860 rutinariamente mediante secuenciación se utilicen los iniciadores descritos por Sharafi, *et al.* (2012) o Reynolds, *et al.* (2012) o los descritos en el Cuadro 1, ya que estos parecen producir los amplicones más específicos. Debido a que estos producen un amplicón relativamente grande (746, 593 y 57epb, respectivamente) se recomienda utilizar la química BigDye Terminator 3.1 con el equipo (ABI3130) con que cuenta en el laboratorio del Instituto de Investigaciones. La tecnología del BigDye Terminator presenta dos proporciones de dNTPs y ddNTPs, las cuales han sido optimizadas por Applied Biosystems para dos aplicaciones distintas, el BigDye 3.1 permite una mayor resolución en fragmentos largos, por lo que es la más adecuada para esta aplicación.

Al realizar el análisis de costos se observa que la tecnología más económica para implementar la prueba es el asPCR, con un costo por muestra de Q187.41, mientras que la más costosa es la reacción por secuenciación, ya que el método disponible en el país es secuenciación Sanger. A pesar de esto es necesario considerar que durante el proceso de implementación de la prueba por cualquier método es necesaria, como se mencionó anteriormente, la verificación de la especificidad por secuenciación, por lo que la implementación en cualquiera de las técnicas debe asumir este costo.

A pesar de que la implementación de la prueba por PCR en tiempo real resulta bastante económica, Q. 193.70, es importante resaltar que la implementación por este método representa una inversión mayor de equipo ya que el valor de un equipo de qPCR es aproximadamente diez veces mayor al de un termociclador de la misma marca, igualmente, se necesitan técnicos de laboratorio mucho más preparados y el mantenimiento del equipo es más costoso. Por estas razones se recomienda la implementación de la prueba en PCR alelo específico.



## VIII. CONCLUSIONES

El análisis de homología estableció un 93% de homología entre la región que contiene el SNP rs12979860 y la región homóloga. A pesar de que estos resultados varían contra los publicados anteriormente la función del análisis de homología es brindar información para el diseño y evaluación de los iniciadores, por lo que los resultados obtenidos son confiables y proporcionan la información necesaria para este estudio.

Las parejas de iniciadores propuestos por Sharafi *et al.* (2012), Reynolds *et al.* (2012) y Domagalski *et al.* (2013) son los que parecen presentar una especificidad adecuada respecto a la región homóloga, por lo que su utilización para la implementación del diagnóstico por PCR seguido de restricción enzimática es altamente recomendable. Por el contrario, la utilización de los iniciadores descritos por Lee *et al.* (2013), Ciesla *et al.* (2012) y Ezzivouri *et al.* (2013) de iniciadores en cualquier tecnología es poco recomendable debido a que estos iniciadores no presentan especificidad respecto a la región homóloga.

Es posible implementar el diagnóstico del SNP rs12979860 mediante un PCR alelo específico mediante la utilización de los iniciadores externos 5'- GGAAGGAGCAGTTGCGCTG-3' y 5'- CTTACGCTCCGAGCATTGC-3, adicional a los iniciadores internos 5'-AGTCCCCGAAGGCGTG-3' y 5'- GTGCAATTCAACCCTGGTTCGC-3.

La implementación de la genotipificación del SNP rs12979860 por PCR en tiempo real más favorable es aquella descrita por Reynolds *et al.*, 2012, por lo que si se desea implementar el diagnóstico por esta tecnología se recomienda implementarlo como se describe en esta publicación.

Considerando el costo de la prueba, la especificidad de los iniciadores requeridos, las condiciones de equipo y laboratoristas requeridos, y la resolución del diagnóstico se concluye que la implementación del diagnóstico por PCR alelo específico es la opción más favorable para la implementación rutinaria de la genotipificación del SNP rs12979860 en el diagnóstico de la farmacogenética de la hepatitis C para la población guatemalteca.

## IX. RECOMENDACIONES

Adicionalmente a la implementación de la genotipificación del SNP rs12979860 es necesario implementar la genotipificación del virus y la evaluación del índice de predicción de estos marcadores para poder optimizar el tratamiento de Hepatitis C en los pacientes de la población Guatemalteca.

Luego de montar el método de detección, ya sea por la tecnología que se desee, es necesario realizar un estudio para probar la correlación del marcador con el desarrollo de respuesta viral sostenida en pacientes de la población para la cual se desea implementar la prueba.

La verificación de la amplificación específica mediante secuenciación del producto de cualquier tecnología basada en PCR es un paso clave para asegurar la detección confiable del SNP rs12979860 y no debe ser omitido.

## X. BIBLIOGRAFÍA

Afdhal, Nezam; *et al.*. 2010. Hepatitis C Pharmacogenetics: State of the Art in 2010. *Hepatology*. 53(1): 336-345.

Amstutz, Ursula y Carleton, B. 2011. Pharmacogenetic Testing: Time for Clinical Practice Guidelines. *Nature*. 89(6): 924- 927.

Bartenschlager, Ralf y Lohmann, V. 2000. Replication of hepatitis C virus. *Journal of General Virology*. 81:1631-1648.

Becker, Robert. 2007. Pharmacogenetic Test and Genetic Test for Heritable Markers. FDA.

Booth, David; Ahlenstiel, G; George, J. 2012. Pharmacogenomics of hepatitis C infections: personalizing therapy. *Genome Medicine*. 4(12).

Brocker, Chad; *et al.*. 2010. Evolutionary divergence and functions of the human interleukin (IL) gene family. *Human Genomics*. 5 (1): 30-55.

Caruntu, Florin y Benea, L. 2006. Acute Hepatitis C Virus Infection: Diagnosis, Pathogenesis, Treatment. *Journal of Gastrointestinal Liver Diseases*. 15(3): 249-256.

Chevaliez, Stéphan y Pawlotsky, J. 2006. Hepatitis C Viruses: Genomes and Molecular Biology. Horizon Bioscience. UK

Choo, Qui; *et al.*. 1989. Isolation of a cDNA Clone Derived from A Blood-Borne Non-A, Non-B Viral Hepatitis Genome. *Science*. 244:359-362.

División de hepatitis Viral. 2010. Hepatitis C Factsheet. Centers for Disease Control and Prevention. Publication No. 21-1075.

Freeman, Anthony; *et al.*. 2001. Immunopathogenesis of hepatitis C virus infection. *Immunology and Cell Biology*. 79: 515-536.

Ghany, Marc; *et al.* 2009. Diagnosis, Management, and Treatment of Hepatitis C: An Update. *Hepatology*. 49(4):1335-1374.

Hanafiah, Khayriyyah; Groeger, J; Flaxman, A y Wiersma, S. 2013. Global Epidemiology of Hepatitis C Virus Infection: New Estimates of Age-Specific Antibody to HCV Seroprevalence. *Hepatology*. 57(4): 1333-1342.

Hydes, Theresa y Khakoo, S. 2012. Innate Immunity: A new chapter for hepatitis C. *Annals of Gastroenterology*. 25: 232-240.

Kato Nobuyuki. 2001. Molecular Virology of Hepatitis C Virus. *Acta Medica Okayama*. 55(3): 133-159.

Kim, Chang y Chang K. 2013. Hepatitis C virus: virology and life cycle. *Clinical and Molecular Hepatology*. 19(1):17-25.

Lin, Chun; *et al.* 2011. IL28B SNP rs12979860 is a critical predictor for on-treatment and sustained Virologic Resonse in patients with hepatitis C virus genotype-1 infection. 2011. 6(3).

Marco, Vito; *et al.* 2012. IL28B polymorphisms influence stage of fibrosis and spontaneous or interferon-induced viral clearance in thalassemia patients with hepatitis C virus infection.. *Hematologica*. 97 (5). 679-686

Meyer, Urs. 2004. Pharmacogenetics- five decades of therapeutic lessons from genetic diversity. *Nature Reviews. Genetics*. 5: 669-676.

Pearlman, Brian y Traub, N. 2010. Sustained Virologic Response to Antiviral Therapy for Chronic Hepatitis C Virus Infection: A Cure and So Much More. *Clinical Infectious Diseases*. 52(7): 889-900.

Pearlman, Brian y Traub, N. 2010. Sustained Virologic Response to Antiviral Therapy for Chronic Hepatitis C Virus Infection: A Cure and So Much More. *Clinical Infectious Diseases*. 52(7): 889-900.

Pineda, Juan; *et al.* 2010. Prediction of Response to Pegylated Interferon plus Ribavirin by IL28B Gene Variation in patients Coinfected with HIV and Hepatitis C virus. *Clinical Infectious Disease*. 51(7):788-795.



Raptopoulou, Sinakos; Sinakos, M; Orfanou, H; Lalla Th y Gigi, H. (2006). Monitoring of patients with sustained virological response treated with standard or pegylated interferon in combination with ribavirin for chronic hepatitis C. *Hippokratia*. 10(1): 32-34.

Rauch; Andri; *et al.* 2010. Genetic Variation in IL28B Is Associated with Chronic Hepatitis C and Treatment Failure: A genome-Wide Association Study. *Gastroenterology*. 138: 1338-1345.

Sharma, Suresh. 2010. Hepatitis C virus: Molecular biology and current therapeutic options. *Indian Journal of Medical Residence*. 131:17-34.

Shepard, Colin; Finelli, L y Alter, M. 2005. Global epidemiology of hepatitis C virus infection. *Lancet of Infectious Diseases*. 5: 558-567.

Soriano, Vicent; *et al.* 2012. Pharmacogenetics of hepatitis C. *Journal of Antimicrobial Chemotherapy*. 67: 523-529.

Wakita, Takaji; *et al.* 2005. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nature Medicine*. 11(7):791-796.

World Health Organization. 2003. Hepatitis C. WHO.

## XI. ANEXOS

Figura 1: Secuencia de los amplicones producidos por los iniciadores diseñados para los alelos homocigotos CC y TT con sus respectivos esquemas de digestión.



Cuadro 1: Propuesta anterior de Iniciadores internos y externos para el PCR alelo específico.

<b>Delta G (kcal/mol)</b>	<b>TM (C)</b>	<b>Contenido GC</b>	<b>Largo</b>	<b>Secuencia</b>	<b>Iniciador</b>
-45.69	71.8	70%	20	ACAGTTGCGCTGCCCCCAGC	<b>eForward'</b>
-36.35	63.7	64.70%	17	CCTTCACGCTCCGAGCA	<b>eReverse'</b>
-46.68	66.6	72.20%	18	GAGTCCCCGAAGGCGTG	<b>iForward''</b>
-48.04	67.3	67.30%	24	GAGTGCAATTCAACCCTGGTTCGC	<b>iReverse</b>

Figura 2: Estructuras de los homodímeros producidos por el iniciador eForward'

Delta G (kcal/mol)	Bases Apareadas	Delta G (kcal/mol)	Bases Apareadas
-9.89	5' ACAGTTGGCTGCCCCCAGC 4 :      : 3' CGACCCCGTCGGGTGACA	-3.14	5' ACAGTTGGCTGCCCCCAGC :    : 3' CGACCCCGTCGGGTGACA
-6.89	5' ACAGTTGGCTGCCCCCAGC 4 :      : : : : 3' CGACCCCGTCGGGTGACA	-3.14	5' ACAGTTGGCTGCCCCCAGC :    : : : : 3' CGACCCCGTCGGGTGACA
-3.55	5' ACAGTTGGCTGCCCCCAGC 3 :     : : : 3' CGACCCCGTCGGGTGACA	-3.14	5' ACAGTTGGCTGCCCCCAGC :    : : : : 3' CGACCCCGTCGGGTGACA
-3.14	5' ACAGTTGGCTGCCCCCAGC 2 :    : 3' CGACCCCGTCGGGTGACA	-3.14	5' ACAGTTGGCTGCCCCCAGC :    : 3' CGACCCCGTCGGGTGACA
3.14	5' ACAGTTGGCTGCCCCCAGC 2 :    : 3' CGACCCCGTCGGGTGACA	-3.14	5' ACAGTTGGCTGCCCCCAGC :    : : : : 3' CGACCCCGTCGGGTGACA
-3.14	5' ACAGTTGGCTGCCCCCAGC 2 :    : : : : 3' CGACCCCGTCGGGTGACA	-1.95	5' ACAGTTGGCTGCCCCCAGC :    : : : : 3' CGACCCCGTCGGGTGACA
-3.14	5' ACAGTTGGCTGCCCCCAGC 2 :    : : : : 3' CGACCCCGTCGGGTGACA	-1.95	5' ACAGTTGGCTGCCCCCAGC :    : : : : 3' CGACCCCGTCGGGTGACA



Figura 3: Estructuras de los homodímeros producidos por el iniciador eReverse'

Delta G (kcal/mol)	Bases Apareadas	Delta G (kcal/mol)	Bases Apareadas
-6.31	4 5' CCTTCACGCTCCGAGCA     : : : 3' ACGAGCCTCGCACTTCC	-3.14	2 5' CCTTCACGCTCCGAGCA : :    : : 3' ACGAGCCTCGCACTTCC
-3.61	2 5' CCTTCACGCTCCGAGCA :    : 3' ACGAGCCTCGCACTTCC	-3.14	2 5' CCTTCACGCTCCGAGCA    3' ACGAGCCTCGCACTTCC
-3.61	2 5' CCTTCACGCTCCGAGCA : :    : : : 3' ACGAGCCTCGCACTTCC	-1.57	2 5' CCTTCACGCTCCGAGCA    : : 3' ACGAGCCTCGCACTTCC
-3.61	2 5' CCTTCACGCTCCGAGCA    3' ACGAGCCTCGCACTTCC		

Figura 4: Estructuras de los homodímeros producidos por el iniciador iForward'

Delta G (kcal/mol)	Bases Apareadas	Delta G (kcal/mol)	Bases Apareadas
-9.49	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG 	-3.07	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG :    : : :
-3.61	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG :    : :	-3.07	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG    : :
-3.61	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG :    : :	-3.07	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG : :    : : :
-3.61	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG 	-1.6	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG    : : :
-3.14	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG    : : :	-1.57	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG :    : :
-3.14	5' GAGCTCCCCGAAAGGCGTG 3' GTGGGAAAGCCCTCGAG :    :		

Figura 5: Estructuras de los homodímeros producidos por el iniciador iReverse'

Delta G (kcal/mol)	Bases Apareadas	Delta G (kcal/mol)	Bases Apareadas
-7.05	4 5' GGAGTGCATTCACACCCCTGGTTTCG : :         : : 3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.95	2 5' GGAGTGCATTCACACCCCTGGTTTCG    3' GCTTGGTCCCAACTTAAAGCTGAGG
-6.36	4 5' GGAGTGCATTCACACCCCTGGTTTCG         : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.95	2 5' GGAGTGCATTCACACCCCTGGTTTCG : :     : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG
-5.36	4 5' GGAGTGCATTCACACCCCTGGTTTCG : :         : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.94	2 5' GGAGTGCATTCACACCCCTGGTTTCG : :     : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG
-3.61	2 5' GGAGTGCATTCACACCCCTGGTTTCG    3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.94	2 5' GGAGTGCATTCACACCCCTGGTTTCG    3' GCTTGGTCCCAACTTAAAGCTGAGG
-3.07	2 5' GGAGTGCATTCACACCCCTGGTTTCG    3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.6	2 5' GGAGTGCATTCACACCCCTGGTTTCG    : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG
-3.07	2 5' GGAGTGCATTCACACCCCTGGTTTCG     : : : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.57	2 5' GGAGTGCATTCACACCCCTGGTTTCG     : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG
-3.07	2 5' GGAGTGCATTCACACCCCTGGTTTCG : :     : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG	-1.57	2 5' GGAGTGCATTCACACCCCTGGTTTCG    : : : : 3' GCTTGGTCCCAACTTAAAGCTGAGG

Figura 6: Interacciones de heterodímeros producidas por los iniciadores propuestos originalmente para el PCR alelo específico.

Delta G (kcal/mol)	Bases Apareadas	5'	3'	Delta G (kcal/mol)	Bases Apareadas	5'	3'
<b>Heterodímero eForward-iForward</b>							
-6.75	3	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG	-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG
-6.21	3	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG	-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG
-4.74	3	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG	-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG
-4.74	3	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG	-3.07	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG
-3.61	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG	-3.07	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG
-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG	-3.07	2	5' GTTGGGCTGCCCCAAGC 	3' GTGGGAAAGTCCCTCGAG
<b>Heterodímeros eForward-iReverse</b>							
-7.04	4	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG	-7.04	4	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG
-6.62	4	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG	-6.62	4	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG
-5.24	4	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG	-5.24	4	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG
-3.61	2	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG	-3.61	2	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG
-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG	-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG
-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG	-3.14	2	5' GTTGGGCTGCCCCAAGC 	3' GOTTGGTCCCAACTTAAAGTGAAG
<b>Heterodímeros eReverse-eForward</b>							
-6.75	3	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA	-8.61	2	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA
-5.09	3	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA	-3.14	2	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA
-5.09	3	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA	-3.14	2	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA
-4.74	3	5' COTTCAAGTCCGAGCA 	3' GSAACCCCTGCGGTTGACA				





Figura 7: amplificación del PCR alelo específico de cada uno de genotipos posibles con los iniciadores originales

