

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN**

**FACULTAD DE MEDICINA**



PROYECTO DE TESIS

**“SECUENCIACIÓN GENÓMICA COMPLETA DE AISLADOS DEL VIRUS DE LA HEPATITIS C (VHC) PARA IDENTIFICAR VARIABILIDAD GENÓMICA, ESTABLECER RELACIONES FILOGENÉTICAS Y VARIANTES CIRCULANTES CONTEMPORÁNEAS”**

PRESENTA:

**ING. BIOTECNÓLOGA KAREN LÓPEZ RIOJAS**

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE:  
MAestrÍA EN CIENCIAS CON ORIENTACIÓN EN BIOLOGÍA MOLECULAR E  
INGENIERÍA GENÉTICA**

DIRECTOR DE TESIS:

**DRA. C. ANA MARÍA G. RIVAS ESTILLA**

**Secuenciación genómica completa de aislados del Virus de la Hepatitis C (VHC) para identificar variabilidad genómica, establecer relaciones filogenéticas y variantes circulantes contemporáneas**

Aprobación de la tesis:



---

**Dra. C. Ana María G. Rivas Estilla**  
Directora de tesis



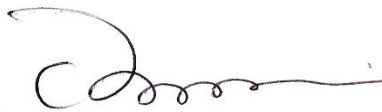
---

**Dr. C. Daniel Arellanos Soto**  
Co director de tesis



---

**Dra. C. Kame Alberto Galán Huerta**  
Miembro de la comisión



---

**Dr. Med. Felipe Arturo Morales Martínez**  
Subdirector de Estudios de Posgrado

El presente trabajo se realizó en coordinación con el Departamento de Bioquímica y Medicina Molecular y el Centro de Investigación e Innovación en Virología Médica (CIIVIM) ambos pertenecientes a la Facultad de Medicina de la Universidad Autónoma de Nuevo León bajo la dirección de la Dra. Ana María Rivas Estilla, como codirector el Dr. Daniel Arellanos Soto y como miembro de la comisión el Dr. Kame Alberto Galán Huerta.

Se resalta, además, el esfuerzo en conjunto que se realizó con el Hospital Universitario “José E. González” específicamente con la Unidad de Hígado y la Dra. Paula Cordero Pérez, Jefa de Laboratorio de la Unidad de Hígado, quienes han proporcionado las muestras de suero utilizadas para esta tesis.

Este trabajo se realizó acorde a la declaración de Helsinki y el protocolo fue aprobado por el comité de ética de la Facultad de Medicina y el Hospital Universitario de la Universidad Autónoma de Nuevo León con el siguiente número de registro: B122-00001.

*“La inteligencia sin la capacidad de dar y recibir un afecto conduce al derrumbe mental y moral, a la neurosis e incluso a la psicosis. Y digo que la mente absorbida en un interés egoísta tomado como un fin en sí mismo, con exclusión de toda relación humana, no puede conducir más que a la violencia y al dolor.”*

Charlie Gordon, Flores para Algernon

## Agradecimientos

A mi hermano **M. C. Francisco Javier López Riojas**, a quien le debo gran parte de lo que soy ahora, mis principios, mis valores, la resiliencia y determinación para cumplir mis metas. Mi mayor apoyo cuando más abajo estuve. Ahora, yo también tengo mi maestría y estoy segura de que seguiría siendo tu conejillo de indias a quien enseñarle programación y que así, tus alumnos pudieran aprender tanto como querías transmitirles.

A mi hermana, **M.COSE. Ariadne López Riojas**, por acercarme ese plato de comida cuando solo había café en mi sistema. Por escucharme en cada colapso y llevarme por más café. Gracias por cada raid a la Uni y asegurarte que siempre llegara sana y salva. Recordar vivir juntas de nuevo fue volver a la triada de la hermandad. Te amo. P.D. Te secuestraré cuando me vaya a Guanajuato, es un hecho. Fin.

A mis padres **Lic. Francisco de Asis López Barcena** y **Dra. Cecilia Riojas Flores**. No existen mejores padres que ustedes y no lo digo por decirlo. Ustedes nos han dado todo, casa, sustento, y sobretodo y lo más importante, su apoyo incondicional y amor. Soy ingenua y tan confiada porque crecí en un ambiente seguro, creo en el “para siempre” porque crecí con el mejor ejemplo de lo que es una pareja, un equipo de por vida. No me alcanzará nunca la vida para expresarles lo afortunada que fui al crecer en la familia que ustedes formaron.

A **Ing. Mario Gabriel Lozano Vega** mi futuro esposo. Te tocó la parte más difícil creo yo, porque escuchaste cada una de mis quejas, escuchaste por horas mis 500 estrategias hipotéticas para averiguar por qué no salían mis experimentos aunque no es tu área de conocimiento, me animaste en cada bajón, aguantabas mis cambios de humor y a distancia me cuidaste, gracias por esas cafiaspirinas a las 12 de la noche y esos desayunos y postres espontáneos. Mi apoyo, uno de mis pilares y también motivación para seguir. Te amo mi sol.

A **Dra. Daira Guadalajara Rubio Mendoza**, esta amistad es de lo más importante que me llevo de esta experiencia, gracias por cada comida juntas, cada viernes de gorditas, por escucharme y consolarme cuando más lo necesité. Me recordaste que siempre vale la pena esforzarse y me inspiraste para superarme más en cada paso de la maestría.

A **Lic. Biotecnóloga Sandra Jessica Olivera Hernández** mi apoyo moral a distancia en mis problemas personales, que afortunada soy de tener a una amiga y colega que me comparta sus conocimientos y me haya ayudado en los temas que se me complicaron aquí. Grande como siempre y bienvenida siempre al desestrés en Cancún.

A mis tíos **Lic. Sergio Riojas Flores** y **Carmen** por abrirme las puertas de su casa en cuanto llegué a Monterrey, jamás olvidaré que me hayan recibido a las 7 de la mañana para hacer mi examen de admisión porque por infortunios justo ese día no teníamos internet. Una comida en casa es lo más reconfortante para los tiempos difíciles, gracias por cada una de ellas.

Al equipo de trabajo del **CIIViM**, gracias por aceptarme e incluirme como una más de ustedes, me encantaron las comidas y todos los festejos, es muy bonito convivir en un ambiente así de armónico. Gracias en especial a **M.C. Tania Guadalupe Heredia Torres** por tan buena disposición que siempre me mostraste para resolver cada duda durante mis procedimientos, ojalá decidas enseñar y estoy

segura que acabarás con éxito tu doctorado; a la **Lic. Biotecnóloga Paulina Moreno Pascual** por tu paciencia para decirme dónde están los reactivos todo el tiempo y tus pláticas para el desestrés; a **M.C. Mayra Alejandra Meléndez Villanueva** quien me hizo sentir más incluida en el equipo y por los consejos también para mi tesis; a la **Q.C.B. Katya Estefanía Pérez Ibarra** por cuidar que mis bandas no se salieran del gel; al **Dr. Daniel Arellanos Soto**, de los mejores profesores que he tenido, gracias por resolver con tanto ánimo cada una de mis preguntas y los consejos antes de seminario; a la **Dra. Sonia Lozano Sepúlveda** por compartir sus conocimientos académicos y empíricos; a la **Dra. Ana María Rivas Estilla** por ser la responsable de integrarme en este equipo, por escuchar mis sugerencias y corregirme cuando fuese necesario.

Finalmente, gracias al **Dr. Kame Alberto Galán Huerta** quien fue mi Jedi en esta travesía, gracias por todos los conocimientos transmitidos, por cada mirada de regaño que me hizo aprender de mis errores, por estar en cada duda y paso, por confiar en mi trabajo y por entender que la vida personal es igual de importante que la vida académica. Puedo decir con todo el orgullo que tuve un excelente profesor y fue un gusto trabajar bajo su tutela. Espero verlo como SNI III tan pronto como se lo proponga.

## COLABORACIONES

**Dr.C. Gerardo Raymundo Padilla Rivas**

**Colaborador**

Departamento de Bioquímica y Medicina Molecular  
de la Facultad de Medicina de la Universidad  
Autónoma de Nuevo León.

**DEPARTAMENTO DE BIOQUÍMICA  
Y MEDICINA MOLECULAR**



**PhD. Linda Elsa Muñoz Espinosa**

**Dra. C. Paula Cordero Pérez**

Unidad de Hígado del Hospital Universitario

“Dr. José E. González”



**UANL**

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN®

## ÍNDICE

<b>COLABORACIONES</b> .....	6
<b>ÍNDICE DE FIGURAS</b> .....	10
<b>ÍNDICE DE TABLAS</b> .....	12
<b>ABREVIATURAS</b> .....	13
<b>RESUMEN</b> .....	16
<b>1. Introducción</b> .....	17
<b>1.1. Epidemiología de la infección por VHC</b> .....	17
<b>1.2. Sintomatología, diagnóstico y tratamiento</b> .....	19
1.2.1. Síntomas .....	19
1.2.2. Diagnóstico .....	21
1.2.3. Tratamiento .....	23
<b>1.3. Biología molecular del VHC</b> .....	26
1.3.1. Estructura y genoma .....	26
1.3.1.1. Estructura .....	26
1.3.1.2. Genoma viral .....	29
1.3.1.3. Ciclo de infección .....	36
<b>1.4. Variabilidad genética del VHC</b> .....	40
<b>1.5. Resistencia a antivirales de acción directa</b> .....	42
<b>1.6. Secuenciación con Oxford Nanopore</b> .....	46
<b>1.7. Análisis filogenético</b> .....	50
<b>2. Antecedentes</b> .....	51
<b>2.1. Antecedentes directos</b> .....	53
<b>3. Justificación</b> .....	54
<b>4. Hipótesis</b> .....	55
<b>5. Objetivos</b> .....	56
<b>5.1. Objetivo general</b> .....	56
<b>5.2. Objetivos específicos</b> .....	56
<b>6. Materiales y Métodos</b> .....	57



<b>8. Resultados</b> .....	74
<b>8.1. Muestras de pacientes</b> .....	74
<b>8.2. Diseño <i>in silico</i></b> .....	76
<b>8.2.1. Síntesis <i>in silico</i> de oligonucleótidos 1a obtenidos a partir de Primal Scheme</b> 76	
<b>8.2.2. Síntesis <i>in silico</i> de oligonucleótidos 1b obtenidos a partir de Primal Scheme</b> 78	
<b>8.2.3. Primer diseño de oligonucleótidos</b> .....	79
<b>8.2.3.1. Estandarización de PCR punto final mediante gradiente de temperatura y amplificación del genotipo 1b</b> .....	79
<b>8.2.3.2. Estandarización de PCR Múltiplex</b> .....	82
<b>8.2.3.3. Amplificación del genotipo 1a</b> .....	83
<b>8.2.4. Segundo diseño de oligonucleótidos</b> .....	86
<b>8.2.4.1. Amplificación del genotipo 1a</b> .....	86
<b>8.2.5. Rediseño y amplificación del genotipo 1b</b> .....	87
<b>8.2.6. Secuenciación</b> .....	94
<b>8.2.7. Análisis filogenético</b> .....	100
<b>8.2.7.1. Análisis filogenéticos en México</b> .....	100
<b>8.2.7.2. Análisis filogenéticos internacionales</b> .....	105
<b>8.2.8. Análisis evolutivo con TempEST</b> .....	108
<b>8.2.8.1. Análisis nacional</b> .....	108
<b>8.2.8.2. Análisis internacional</b> .....	110
<b>8.2.9. Análisis filogenético evolutivo con BEAST 2.0</b> .....	112
<b>8.2.9.1. Análisis nacional</b> .....	113
<b>8.2.9.2. Análisis internacional</b> .....	115
<b>9. Discusión</b> .....	119

<b>Conclusiones</b> .....	127
<b>10. Perspectivas</b> .....	128
<b>11. Bibliografía</b> .....	129
<b>12. Anexos</b> .....	143
<b>12.1. Anexo I Reactivos para la extracción de ARN</b> .....	143
<b>12.2. Anexo II Síntesis de cDNA</b> .....	144
<b>12.3. Anexo III Reactivos y condiciones para los ciclos de PCR punto final</b> .....	145
<b>12.4. Anexo IV Preparación de Buffer TBE 5x</b> .....	146
<b>12.5. Anexo V Condiciones de temperatura para PCR Multiplex</b> .....	147
<b>12.6. Anexo VI Primer diseño de oligonucleótidos GT 1a Primal Scheme</b> .....	148
<b>12.7. Anexo VII Primer diseño de oligonucleótidos GT 1b Primal Scheme</b> .....	150
<b>12.8. Anexo VIII Oligonucleótidos 1a sintetizados por Oligo T4</b> .....	152
<b>12.9. Anexo IX Oligonucleótidos 1b sintetizados por Oligo T4</b> .....	154
<b>12.10. Ejemplo de Resultados de evaluación en Amplify 4</b> .....	156
<b>12.11. Anexo XI Script para ensamblaje y limpieza de secuencias con Porechop y Flye</b>	157

## ÍNDICE DE FIGURAS

<b>Figura 1</b> Distribución regional de las muertes por hepatitis víricas del 2016.....	18
<b>Figura 2</b> Clasificación de los antivirales de acción directa de acuerdo a su mecanismo de acción.....	24
<b>Figura 3</b> Modelo del corte transversal del virión HCV. ....	26
<b>Figura 4</b> Composición del genoma del VHC. ....	29
<b>Figura 5</b> Ciclo de replicación del HVC. ....	39
<b>Figura 6</b> Tecnologías Oxford Nanopore.. ....	47
<b>Figura 7</b> Diagrama de flujo para el diseño de oligonucleótidos in silico y experimental. ...	64
<b>Figura 8</b> Flujo de trabajo para el tratamiento de datos crudos de secuenciación.....	69
<b>Figura 9</b> Muestras de pacientes infectados con VHC.....	74
<b>Figura 10</b> Diseño in silico de Primal Scheme para los GT 1a y 1b .. ....	77
<b>Figura 11</b> Primer diseño de oligonucleótidos 1b con 10 secuencias de NCBI.....	78
<b>Figura 12</b> Amplificación de los fragmentos de 400pb GT 1b.. ....	80
<b>Figura 13</b> Temperaturas óptimas de amplificación de los oligonucleótidos GT 1b.. ....	81
<b>Figura 14</b> Primera PCR Múltiplex GT 1b. ....	82
<b>Figura 15</b> Amplificación de fragmentos GT 1a. ....	85
<b>Figura 16</b> Segundo diseño de oligonucleótidos GT 1a.....	86
<b>Figura 17</b> Rediseño de oligonucleótidos GT 1b.....	87
<b>Figura 18</b> Primera secuenciación.. ....	89
<b>Figura 19</b> . Ejemplo del flujo de trabajo para el rediseño de oligonucleótidos GT1b. ....	91
<b>Figura 20</b> Segunda amplificación de fragmentos GT 1b.....	92
<b>Figura 21</b> PCR Multiplex y resumen de resultados del segundo diseño GT 1b.....	93

<b>Figura 22</b> Gráficos de Genome Detective sobrelapados de la segunda secuenciación del GT1b.....	96
<b>Figura 23</b> Detección de resistencia a DAAs.....	97
<b>Figura 24</b> Genotipado muestra ID 213. NCBI Genotyping Tool.....	99
<b>Figura 25</b> Árbol filogenético de máxima verosimilitud con el modelo TPM2+F+G4 del gen NS3 en México.....	101
<b>Figura 26</b> Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 de los genes NS3 y NS4A en México.....	102
<b>Figura 27</b> Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del gen NS4A en México.....	103
<b>Figura 28</b> Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del gen NS5B en México.....	104
<b>Figura 29</b> Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del virus secuenciado en México con las secuencias de Estados Unidos de América. ....	106
<b>Figura 30</b> Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del virus secuenciado en México con las secuencias a nivel mundial. ....	107
<b>Figura 31</b> Análisis evolutivo TempEst con el método media residual heurística al cuadrado. ....	109
<b>Figura 32</b> Análisis evolutivo TempEst.....	111
<b>Figura 33</b> Filogenia evolutiva del Gen NS3 en México.. ....	113
<b>Figura 34</b> Filogenia evolutiva del Gen NS5B en México.....	114
<b>Figura 35</b> Análisis evolutivo de la secuencia completa del VHC en México con las de Estados Unidos.....	116
<b>Figura 36</b> Análisis evolutivo a nivel global de la secuencia completa del VHC en México.. ....	118

## ÍNDICE DE TABLAS

<b>Tabla 1</b> Síntomas de la infección aguda y crónica de Hepatitis C.....	20
<b>Tabla 2</b> Interpretación de Diagnóstico de VHC.....	23
<b>Tabla 3</b> Estructura y funciones del VHC .....	27
<b>Tabla 4</b> Variabilidad genética del VHC.....	40
<b>Tabla 5</b> RAS identificadas en los genes NS3, NS5A y NS5B debido a la administración de tratamientos con DAAs en 6 genotipos del VHC. Ceccherini-Silberstein et al., 2018. ..	43
<b>Tabla 6</b> Tipo de archivos utilizados en análisis bioinformático.....	49
<b>Tabla 7</b> Criterios de inclusión y exclusión de los pacientes donadores de muestras .....	58
<b>Tabla 8</b> Volúmenes utilizados para la síntesis de ADNc con enzima M-MLV RT .....	60
<b>Tabla 9</b> Volúmenes utilizados para la síntesis de ADNc con enzima SuperScript III RT... ..	61
<b>Tabla 10</b> Genomas del VHC genotipos 1a y 1b utilizados para el diseño de oligonucleótidos .....	62
<b>Tabla 11</b> Volúmenes utilizados para la PCR punto final.....	65
<b>Tabla 12</b> Condiciones de amplificación con gradiente de temperatura para oligonucleótidos 1a y 1b.....	66
<b>Tabla 13</b> Volúmenes para la preparación de PCR Multiplex .....	67
<b>Tabla 14</b> Oligonucleótidos utilizados en la PCR Multiplex .....	68
<b>Tabla 15</b> Muestras seleccionadas de pacientes infectados con VHC.....	75
<b>Tabla 16</b> Secuenciación GT 1b.....	95
<b>Tabla 17</b> RAS de las secuencias VHC GT 1b.....	98

## ABREVIATURAS

ASSLD	American Association for the Study of the Liver Diseases
μL	Microlitro
ALT	Alanina transaminasa
AST	Aspartato aminotransferasa
bDNA	Branched Deoxyribonucleic Acid
cLDL	Gotas Lipídicas Citosólicas
CLDN1	Claudina-1
CLIA	Chemiluminescence Immunoassay
DAA	Antivirales de Acción Directa
DGAT1	Enzima Diacilglicerol Acetiltransferasa-1
ELISA	Enzyme-Linked ImmunoSorbent Assay
FA	Fosfatasa alcalina
GGTP	Gamma-glutamyl transpeptidasa
GT	Genotipo
GTR	Genotipic Test Resistance
GWAS	Genome Wide Association Study
HSPG	Proteoglicanos de Sulfato de Heparán
igVR	Región Variable Intergenotípica
IRES	Internal Ribosome Entry Site
kb	Kilobytes
LDLr	Receptor de Lipoproteínas de Baja Densidad
LuLD	Gota Lipídica Luminal

Mb	Megabytes
min	Minutos
mL	Mililitro
NAT	Nucleic Acid Test
NGS	Tecnologías de Nueva Generación
OMS	Organización Mundial de la Salud
ONT	Oxford Nanopore Technologies
PEG	Polietilenglicol
PEG-IFN	Interferón Pegilado
pmol	picomol
pre-VLDL	Partícula lipoviral de muy baja densidad inmadura
qRT-PCR	quantitative RT-PCR
RAS	Resistance-associated substitutions
RBV	Rivabirina
RE	Retículo Endoplásmico
RIBA	Recombinant Immunoblot Assays
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
RVS	Respuesta Viroológica Sostenida
s	Segundos
SHARED	Surveillance of Hepatitis C Antiviral Resistance, Epidemiology and methoDologies
SNP	Single Nucleotide Polimorfism
SRB1	Receptor Scavenger B1

TG	Triacilglicerol
TGN	Red trans-Golgi
TM	Dominio Transmembrana
TMA	Transcription-Mediated Amplification
VHC	Virus de la Hepatitis C



## RESUMEN

Título del estudio: “**Secuenciación genómica completa de aislados del virus de la hepatitis C (VHC) para identificar variabilidad genómica, establecer relaciones filogenéticas y variantes circulantes contemporáneas**”

Número de páginas: 158

**Introducción:** La muerte por daño hepático causado por el virus de la hepatitis C (VHC) representa una de las amenazas sanitarias más frecuentes en México. El tratamiento actual de las infecciones por VHC incluye agentes antivirales de acción directa, sin embargo, este virus ha adquirido mutaciones que dificultan el efecto antiviral. En México se desconoce la presencia de estas mutaciones, y el genoma completo del VHC aún no ha sido secuenciado. El objetivo de este estudio fue obtener una secuencia completa del genoma del VHC aislado de pacientes en México, identificar las sustituciones asociadas a la resistencia y determinar las relaciones filogenéticas. **Métodos:** Se evaluaron pacientes con hepatitis C del "Hospital Universitario Dr. José Eleuterio" en Monterrey, México, durante mayo de 2016 a agosto de 2020. Se extrajo el ARN de muestras de suero y se amplificó el genoma completo del VHC con PCR Múltiplex. Los amplicones se secuenciaron con la tecnología de MinION. Las secuencias se ensamblaron con el flujo de trabajo medaka del pipeline bioinformático Artic. Las sustituciones asociadas a resistencia se identificaron con el servidor web HCV-Glue. Se prepararon 3 conjuntos de datos, genoma completo, NS3 y NS5B. Las secuencias parciales previamente reportadas en México se utilizaron en el conjunto de datos de un solo gen. Obtuvimos los árboles de máxima verosimilitud con IQtree y se analizaron a posteriori con Beast 1.10.4. **Resultados:** Se obtuvieron 4 genomas parciales y uno completo de VHC (genotipo 1b). Se detectó resistencia a daclatasvir, y una probable resistencia a ledipasvir y velpatasvir. Otro VHC con probable resistencia a daclatasvir y ombitasvir. La tasa de sustitución global fue de 6,655 sustituciones E<sup>-4</sup> por sitio y año. El nuevo genoma obtenido se agrupa con los virus aislados en EE.UU. El ancestro común más reciente del genoma obtenido se dedujo que circuló en 1927. La muestra no se agrupó con virus contemporáneos. El alineamiento del gen NS3 mostró que los virus recién secuenciados se agrupaban en un clado diferente de los virus previamente secuenciados en México. En cuanto al gen NS5B, las nuevas secuencias obtenidas se agruparon con los virus previamente secuenciados en México.

---

**Dra. C. Ana María G. Rivas Estilla**  
Firma del Asesor

## **1. Introducción**

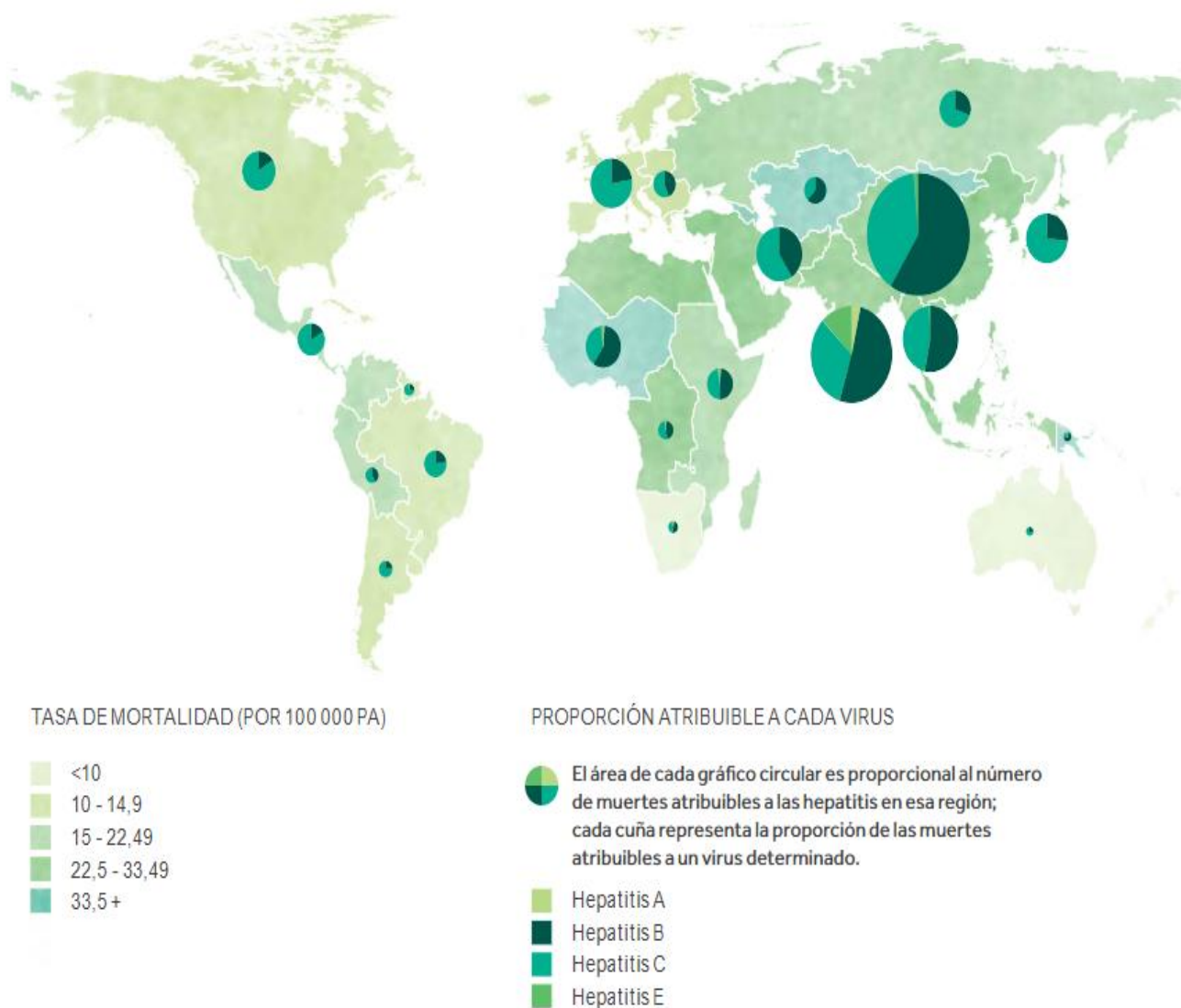
### **1.1. Epidemiología de la infección por VHC**

La hepatitis C es una enfermedad infecciosa provocada por el virus de la hepatitis C (VHC) que genera inflamación hepática aguda o crónica. Es la principal causa del desarrollo de una cirrosis y carcinoma hepatocelular (Koutsoudakis et al., 2013). Esta infección se transmite principalmente por vía parental y por contacto sexual. Existen prácticas de riesgo que elevan las probabilidades de adquirirla como compartir agujas en la administración de drogas inyectables, prácticas de venopunción poco seguras, transfusión de sangre sin análisis previos y prácticas sexuales de riesgo que llevan un contacto sanguíneo (OMS, 2020).

De acuerdo con la Organización Mundial de la Salud (OMS) se han reportado en el 2020 cerca de 325 millones de personas infectadas por el virus de hepatitis C y hepatitis B. En cuanto a la infección por hepatitis C, hay aproximadamente 71 millones de personas con infección crónica a nivel global (OMS, 2020). En México, durante el 2019, se registraron 2,378 casos de Hepatitis C y en el 2020, se registraron 1366 casos nuevos (DGE, 2020) y de acuerdo con Instituto Nacional de Salud Pública en el 2017, se estimó un total de 1.6 millones de personas infectadas.

Además de los contagios y los nuevos casos, en el reporte “Estrategia Mundial del Sector de la Salud contra las Hepatitis Víricas 2016-2021”, de la OMS, en el 2016 se reportó una tasa de mortalidad mayor al 22.5/100,000 habitantes causada por hepatitis en México, siendo la más agresiva la hepatitis C (Figura 1).

La hepatitis aguda suele ser asintomática y no detectarse en un 70%-80% de los casos (Salvatierra, 2017). Sin embargo, dentro del total de pacientes infectados, el 80% evoluciona a una fase crónica (Jang & Chung, 2010; Salvatierra, 2017). La OMS y Salvatierra reportan que en el 15%-30% de los casos, la infección se resuelve favorablemente. Esta variabilidad en la evolución del cuadro clínico hacia etapas más agresivas de la infección, genera un panorama importante a evaluar: se requiere mayor conocimiento sobre la sintomatología y serología de ésta para poder llevar a cabo un diagnóstico y tratamiento adecuados a la etapa de infección.



**Figura 1** Distribución regional de las muertes por hepatitis víricas del 2016. Extraído de “Estrategia mundial del sector de la salud contra las hepatitis víricas 2016-2021”

## 1.2. Sintomatología, diagnóstico y tratamiento

### 1.2.1. Síntomas

Las manifestaciones clínicas de la enfermedad se manifiestan gradualmente, de acuerdo con la etapa de evolución que presenta el individuo infectado. En cuadros agudos los síntomas pueden pasar desapercibidos y aumentar en un 80% de los casos. En una etapa crónica, los síntomas se agravan teniendo un alto riesgo de desarrollar cirrosis y carcinoma hepatocelular. Éstos se presentan en la Tabla 1.

Por otro lado, existen marcadores bioquímicos que indican el progreso de una enfermedad hepática, los cuales pueden conducir finalmente hacia el desarrollo de la cirrosis como lo son (Salvatierra, 2017; V.F. & A.López, 2015):

- Hiperbilirrubinemia: La hiperbilirrubinemia consiste en un aumento de la concentración de bilirrubina no conjugada en sangre. Cuando es de suficiente intensidad da lugar a una coloración amarillenta de la piel y las mucosas denominada "ictericia". La bilirrubina no conjugada es producto del catabolismo del Grupo Hemo, es poco soluble en plasma y solo puede ser transportada al hígado cuando se une a la albúmina. Si ésta no se conjuga no es excretada a través de la orina y aumenta su concentración. Ésta es conjugada dentro de las células hepáticas, en el citoplasma se une a la ligandina y se transporta al retículo endoplásmico en donde la glucoronil transferasa difosfato 1A1 cataliza la reacción de conjugación (Carvajal Carvaja, 2019). Cuando existe un daño hepático, la conjugación se ve afectada.
- Patrón de citólisis: Por lo general, la enzima ALT (alanina transaminasa) se encuentra en interior de las células hepáticas. En eventos de daño hepatocelular, se libera esta enzima al torrente sanguíneo. Los valores normales son de 10-40 UI/L y 7-35UI/L para hombres y mujeres respectivamente, y su elevación  $\geq 1.000$  UI/L es un importante indicador de daño hepatocelular.
- Patrón de colestasis: Aumento de GGTP (Gamma-glutamil transpeptidasa) y FA (Fosfatasa alcalina). La función de la GGTP es transferir grupos gamma-

glutamil, este fenómeno se da cuando hay un impedimento para la llegada de la bilis al intestino, éste puede ser por obstrucción por la destrucción de los conductos biliares del hígado.

**Tabla 1** Síntomas de la infección aguda y crónica de Hepatitis C.

	<b>Infección aguda</b>	<b>Infección crónica</b>
<b>Tiempo</b>	2 semanas -6 meses	>15 años
<b>Síntomas</b>	<ul style="list-style-type: none"> <li>• Fiebre</li> <li>• Fatiga</li> <li>• Nauseas</li> <li>• Vómito</li> <li>• Coluria</li> <li>• Presencia o no de ictericia</li> <li>• Falta de apetito</li> </ul>	<ul style="list-style-type: none"> <li>• Lesión hepática continua</li> <li>• Insuficiencia hepática (10% de casos)</li> <li>• Cirrosis</li> <li>• Descompensación hepática (13%-15%)</li> <li>• Hepatocarcinoma celular (2-7%)</li> <li>• Esteatosis</li> </ul>
<b>Respuesta celular</b>	<ul style="list-style-type: none"> <li>• Niveles altos hasta un 10% de ALT (2-4 semanas)</li> <li>• ARN-VHC detectable (2 semanas después de la infección)</li> <li>• Células T específicas (5-9 semanas)</li> <li>• Anticuerpos específicos (2- 5 meses)</li> </ul>	<ul style="list-style-type: none"> <li>• Niveles de ALT y/o AST altos o fluctuantes</li> <li>• Producción de <math>10^{12}</math> viriones/día</li> <li>• IgM anti <i>core</i> elevado (4ta semana)</li> </ul>

- 
- Título persistente de T CD4<sup>+</sup> y CD8<sup>+</sup> (aún después de décadas)
- 

\*ALT: Alanina transaminasa.

\*AST: Aspartato aminotransferasa

\*ARN-VHC: ARN del virus de hepatitis C

Extraído de NIH, 2017, (HHS et al., 2020; Lozano Mérida, 2005; Salvatierra, 2017).

Elaboración propia.

### 1.2.2. Diagnóstico

El diagnóstico consta de pruebas clínicas, principalmente divididos en dos categorías: diagnóstico con ensayos serológicos (indirecto) y diagnóstico mediante métodos de detección, caracterización y cuantificación de componentes de las partículas virales del VHC (directo) (Gupta et al., 2014).

En el primer grupo se encuentran las pruebas ELISA (Enzyme-Linked ImmunoSorbent Assay) para la detección de anticuerpos anti-VHC y para suplementar este tipo de pruebas se encuentra el RIBA (Recombinant Immunoblot Assays) y el CLIA (Chemiluminescence Immunoassay). Las pruebas ELISA y RIBA han evolucionado con los años desde el descubrimiento del virus en 1989, catalogándose como pruebas de primera, segunda, tercera y hasta cuarta generación para la detección de anti-VHC. Actualmente el ELISA de 4ta generación es capaz de detectar simultáneamente antígenos de la cápside y anticuerpos del núcleo de las regiones NS3, NS4 y NS5.

Sin embargo, debido a los avances en biología molecular las pruebas clínicas inmunológicas/indirectas han quedado obsoletas o se utilizan en laboratorios con recursos limitados y en su lugar, para la confirmación del ELISA, se utilizan métodos moleculares como NAT (Nucleic Acid Test) para detectar ARN del VHC y se ha

considerado como el estándar de oro para detectar la replicación del VHC. (Gupta et al., 2014; Salvatierra, 2017).

La etapa crónica se diagnostica cuando los anticuerpos contra el VHC que se generan en la persona infectada se encuentran presentes en la sangre durante más de 6 meses. Una de las desventajas del ELISA para detección de anticuerpos, es que tiene una sensibilidad menor al 95%, es decir, no detecta niveles inferiores a 20, 000UI/ml (Salvatierra, 2017).

La segunda categoría de diagnóstico se basa en la detección de componentes virales como fragmentos del genoma del VHC, evidenciando así, la existencia de la replicación vírica activa. Estas técnicas moleculares permiten la detección cualitativa y cuantitativa del ARN-VHC circulante en suero, plasma o tejido (Gupta et al., 2014; Salvatierra, 2017).

La detección cualitativa se utiliza para predecir precozmente la respuesta virológica sostenida (RVS), para confirmar la viremia y para pacientes inmunocomprometidos que no hayan desarrollado la respuesta inmune. Se utiliza convencionalmente métodos como RT-PCR (Reverse Transcription-Polymerase Chain Reaction) o TMA (Transcription-Mediated Amplification) (Gupta et al., 2014; Salvatierra, 2017).

La detección cuantitativa permite cuantificar la carga vírica basal y los métodos más utilizados son la qRT-PCR (quantitative RT-PCR) detectando hasta 30 copias/ml y la tecnología bDNA (Branched Deoxyribonucleic Acid)(Gupta et al., 2014; Salvatierra, 2017).

Dentro de las pruebas comerciales moleculares para identificación del genotipo viral y variantes genéticas más utilizadas en países de bajos recursos como México, se encuentran INNO-LiPA ®VHC II y Truprep® VHC 5'NC Genotyping Kit ambos de Bayer Diagnostics. La primera es la de mayor uso y consiste en un ensayo de hibridación inversa. Cuando se obtiene el producto amplificado de la región 5'UTR éste se desnaturaliza y se hibrida a una tira de nitrocelulosa que contiene sondas específicas del genotipo o subtipo. El segundo se basa en la secuenciación bidireccional de la región 5'UTR del ARN del VHC. Se purifica el amplicón y se

secuencia mediante cebadores marcados fluorescentemente (síntesis en dirección 5' y 3') y dideoxiterminadores. Para cada muestra se requieren de 4 reacciones de secuenciación para determinar en cada una, la secuencia de A,C,G y T. La interpretación se realiza mediante un software mediante alineamiento y homología.

De manera general en la práctica clínica, la presencia de anticuerpos anti-VHC y ARN-VHC son indicadores de la infección. Las posibles interpretaciones se resumen en la Tabla 2.

**Tabla 2** Interpretación de Diagnóstico de VHC

Anticuerpos anti-VHC	ARN VHC	Diagnóstico
-	-	Negativo
-	+	Hepatitis C aguda
+	-	Probablemente no es Hepatitis C aguda (hacer nuevamente los ensayos unas semanas después)
+	+	Positivo. Requiere pruebas adicionales para definir etapa de progreso

Chevaliez & Pawlotsky, 2017

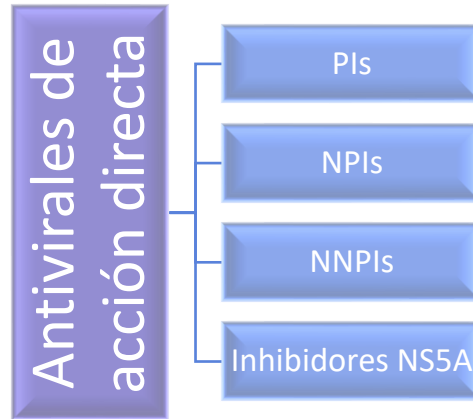
Una vez realizado el diagnóstico, el tratamiento se decide con base a la edad del paciente, presencia de síntomas, severidad de la enfermedad del hígado, probabilidad de respuesta al tratamiento, posibles efectos médicos y contradicciones en el uso de terapias con interferón.

### 1.2.3. Tratamiento

Los tratamientos para la hepatitis C tienen como objetivo curar la enfermedad. En el 2018 la OMS recomendó terapias que involucran el uso de DAA (Antivirales de Acción Directa). Los DAA son moléculas que tienen como blanco proteínas específicas no estructurales atacando pasos específicos del ciclo viral. Su clasificación se da de acuerdo con el tipo de mecanismo de acción que tienen



(Figura 2) (Paul J Pockros, 2021), éstos son capaces de curar a la mayor parte de los pacientes en un periodo entre 12-24 semanas. Sin embargo, éstos tienen un costo elevado haciéndolo poco accesible a todas las poblaciones.



**Figura 2** Clasificación de los antivirales de acción directa de acuerdo a su mecanismo de acción. Inhibidores de la proteasa (PI) de las proteínas no estructurales 3/4A (NS3/4A), inhibidores nucleósidos de la polimerasa (NPI) de la NS5B e inhibidores no nucleósidos de la polimerasa (NNPI) de la NS5B. (Poordad & Dieterich, 2012).

El tratamiento común hasta el 2011 consistía en una combinación con interferón pegilado (PEG-IFN) y ribavirina (RBV) (Bryan-Marrugo et al., 2015). El PEG-IFN es una modificación química del interferón a la cual se le han añadido cadenas de polietilenglicol (PEG). Éstas lo protegen del sistema metabólico humano prolongando la biodisponibilidad del fármaco en el organismo y así, aumentando su efecto antivírico (Carretero Colomer, 2006).

Después, en este mismo año (2011), surgieron tratamientos que incorporaron a los antivirales de acción directa (Bacon et al., 2011; Cummings et al., 2010; Fried et al., 2013; Jacobson et al., 2011; Sulkowski et al., 2013).

- Primera generación: Telaprevir y Boceprevir (inhibidores de la proteasa NS3).
- Segunda generación: Simeprevir, Sofosbuvir (inhibidor de la polimerasa NS5B) y Feldaprevir (inhibidor de la proteasa NS3/4A).

La administración y el tiempo de tratamiento depende del genotipo del virus. Antes de los antivirales de acción directa, se recomendaba primariamente, para los

pacientes que presentan genotipos 1 y 4, el tratamiento con PEG-IFN + Ribavirina y éste tenía una duración de 1 año, mientras que, para los genotipos 2 y 3 era de 6 meses (Jimenez-Mendez et al., 2010). El fin de este tratamiento es que el paciente alcance una respuesta viral sostenida (RVS) después de 6 meses de terminar el tratamiento. Es por ello que se requieren de pruebas serológicas que identifiquen el serotipo y pruebas genotípicas que identifiquen el genotipo para poder determinar el tratamiento y tiempo más adecuados.

Otra aspecto importante para el tratamiento es la carga viral. Los pacientes que tienen una carga viral menor a 200 000 UI/ml tienen una mayor probabilidad de alcanzar una RVS que los pacientes que tienen una carga >600 000 IU/ml (Jimenez-Mendez et al., 2010).

En México, en el 2018, la Cofepris aprobó el medicamento Mavyret de la compañía Abbvie para ser incorporado en el esquema de tratamiento para la hepatitis C. Este antiviral de acción directa es de tipo pangénotípico y es una combinación de Glecaprevir y Pibrentasvir. Éstos son inhibidores de las proteasas NS3/4A y NS5A respectivamente (COFEPRIS, 2018). Sin embargo, debido a su costo elevado, el esquema de DAAs en México incluye otros como el Sofosbuvir, Velpatasvir, Ledipasvir y Simeprevir.

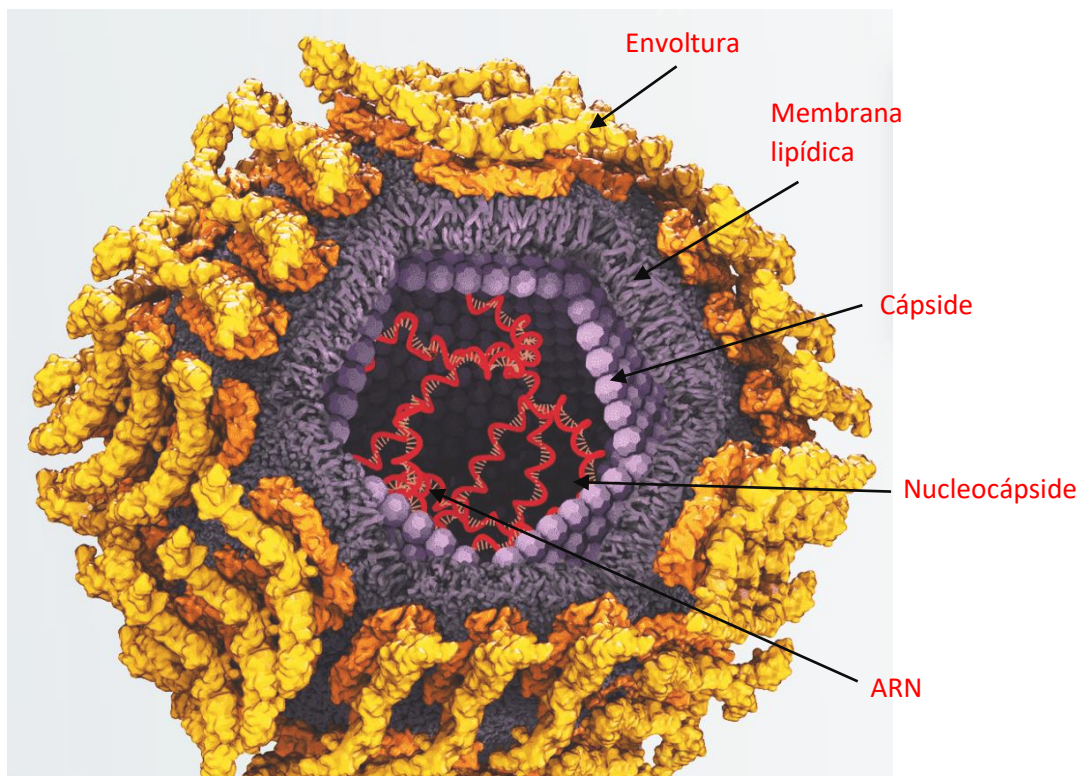
Actualmente no existe profilaxis ni vacuna para el VHC y, como ya se mencionó, es importante tomar en cuenta el genotipo del VHC que ha infectado al paciente para el pronóstico y duración del tratamiento.

### 1.3. Biología molecular del VHC

#### 1.3.1. Estructura y genoma

##### 1.3.1.1. Estructura

El VHC pertenece al género *Hepacivirus* y a la familia *Flaviviridae*. Es un virus de cadena sencilla de ARN de polaridad positiva que codifica para una poliproteína de aproximadamente 9,300 – 9,400pboc(Q. L. Choo et al., 1991). Las partículas virales son esféricas y heterogéneas en tamaño (Figura 3), regularmente entre 40-65nm de diámetro (Gastaminza et al., 2010; Kaito et al., 1994) y está asociado a lipoproteínas(Lavillette et al., 2006).



**Figura 3** Modelo del corte transversal del virión HCV. Creado por Cognition Studio

El virión se compone de una envoltura, membrana lipídica, cápside, nucleocápside el genoma de ARN y partículas lipovirales. Cada una de estas partes se describen en la Tabla 3.

**Tabla 3** Estructura y funciones del VHC

<b>Estructura</b>	<b>Descripción</b>	<b>Función</b>
<i>Envoltura</i>	Consiste en glicoproteínas E1 y E2 que forman una estructura oligomérica formada por trímeros de estos heterodímeros.	E2: protege a E1 del sistema inmune Juega un rol en la unión con el receptor, fusión de la membrana del endosoma de lípido y en el ensamblaje
<i>Membrana lipídica</i>	Se compone de colesterol, ésteres de colesterol, fosfatidilcolina y esfingomielina.	Tiene un rol en la entrada del VHC a la célula
<i>Cápside</i>	Se compone de proteínas nucleares ( <i>core</i> ). La cápside es heterogénea y esférica con un diámetro de 30nm	Las proteínas del núcleo se ensamblan en la cara citoplásmica del RE para formar la cápside; este proceso de ensamblaje se desencadena por las interacciones con el ARN del VHC y la proteína E1 de la envoltura.
<i>Nucleocápside</i>	Está formada por el ARN del virus y la cápside, la cual consiste exclusivamente de proteínas <i>core</i> .	Forma el escudo que encapsula y protege al ARN VHC

ARN

Contiene 1 marco de lectura: 3006-3037 codones flanqueado por regiones 5' y 3' UTR

Región 5' UTR: tiene 340 nucleótidos de longitud y tiene 4 dominios estructurales: I bucle y del II-IV IRES

Región 3'UTR: tiene 225 nucleótidos y consiste en 3 componentes: región variable, tracto largo poly U/UC y la región 3'X

Se utiliza para transcripción y traducción

Los IRES media la unión del ARN VHC al ribosoma del huésped

La región 3'X es un factor clave para cambiar entre la traducción y transcripción del ARN VHC

*Partículas Lipovirales*

Son partículas constituidas principalmente de triglicéridos, ARN VHC, proteínas de la cápside, E1 y E2, apolipoproteína B y E.

Su tamaño es de aproximadamente 100nm

Ayuda a que el VHC pueda circular por la sangre como un híbrido de partícula lipoviral. Facilita la entrada del VHC a los hepatocitos y lo protege de la neutralización por anticuerpos.

\*RE: Retículo Endoplásmico

\*IRES: InteARNI Ribosomes Entry Sites

Elaboración propia. Información extraída de

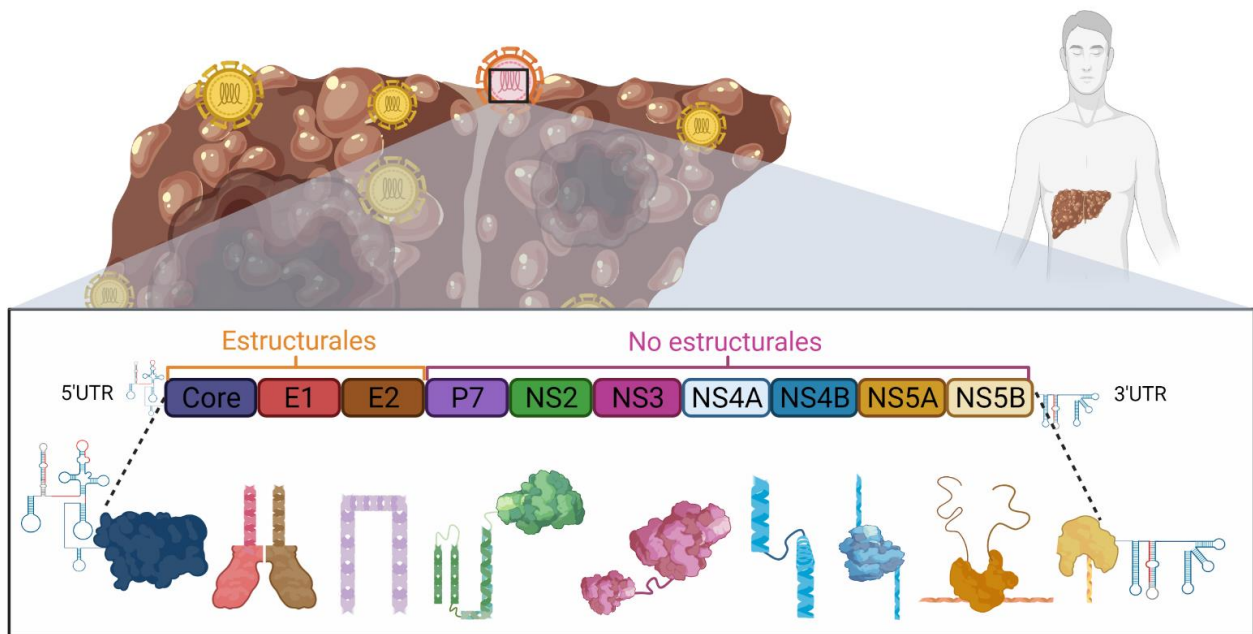
<https://www.hepatitisc.uw.edu/biology/structure#envelope>

### 1.3.1.2. Genoma viral

Como ya se mencionó anteriormente, el genoma del VHC contiene un solo marco de lectura abierto que codifica una poliproteína viral de 3006-3037 aminoácidos. El genoma no contiene caperuza de 5'-metilguanidina ni cola poli A que son característicos de los ARN mensajeros eucariotas, en su lugar se encuentran las regiones 5'UTR y 3'UTR. Éstas son las regiones más conservadas del genoma del VHC presentando menos mutaciones entre los subtipos (Bukh et al., 1992; Han et al., 1991).

La región 5'UTR contiene una secuencia IRES (sitio de entrada interno al ribosoma), la cual permite la traducción de la poliproteína en un mecanismo independiente a cap; seguido de la región codificante para los genes y el 3'UTR (C et al., 1993; K et al., 1992).

La secuencia codificante contiene información para la síntesis de proteínas estructurales y no estructurales comprendiendo un total de 9 regiones (Figura 4).

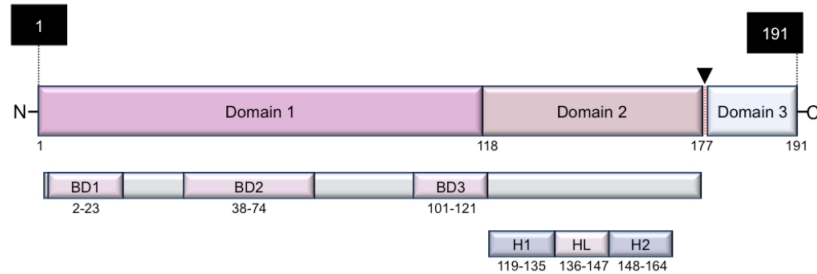
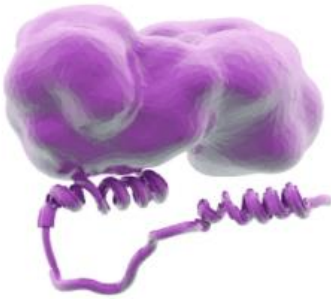


**Figura 4** Composición del genoma del VHC. Se muestra el marco de lectura abierto que codifica para una proteína de aproximadamente 3000aa. Elaboración propia Biorender

A continuación, se describen las funciones y estructura de las proteínas del genoma viral.

### <sup>1</sup>Core

Forma la cápside viral protegiendo el ARN viral. Es la primera proteína que se traduce. La región N terminal es la que se une al ARN del VHC. Consta de 177aa y tiene un peso de 71kDa. Tiene 3 dominios (Eng et al., 2017; Gawlik & Gallay, 2014; Kao et al., 2016; Klein et al., 2004; Strosberg et al., 2010).



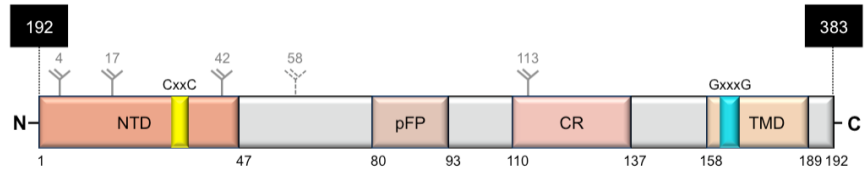
### <sup>2</sup>E1

Es una proteína transmembrana altamente glicosilada que se asocia con la E2 para formar el heterodímero. La unión con la célula huésped puede involucrar la unión a apoproteínas del huésped o a CD36. Tras la endocitosis de la partícula viral, el entorno ácido del endosoma estimula los cambios de conformación de E1 y E2 generando la fusión de membrana endosomal con la envoltura viral y liberación del ARN al citoplasma. Consta de 192 aa y tiene un peso de 23kDa. Tiene 4 dominios principales ( Beeck et al., 2000; Omari et al., 2014; Haddad et al., 2017; Mazumdar et al., 2011; Moustafa et al., 2018)



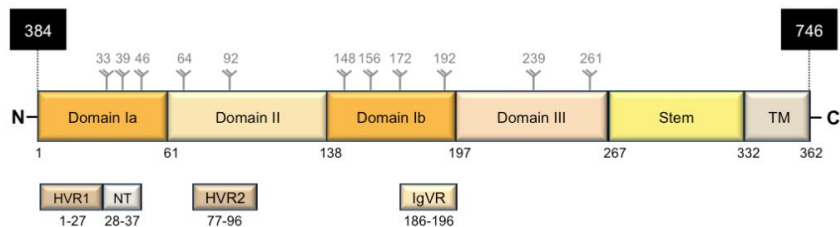
<sup>1</sup> Todos los esquemas mostrados pertenecen a Cognition Studio

<sup>2</sup> Todos los esquemas mostrados pertenecen a Cognition Studio



### <sup>3</sup>E2

Es una proteína transmembranal altamente glicosilada, asociada con la glicoproteína E1, formando heterodímeros que conforman la envoltura del VHC. Tiene un papel en la unión al receptor del huésped, la entrada y la fusión con la membrana endosomal. Se une a receptores CD81y el scavenger SR-B1. Consta de 363 aa y tiene un peso de 30kDa. Tiene una región hipervariable 1 (HVR1), la HVR2, la región variable intergenotípica (igVR), una región madre y un dominio transmembrana (TM). Tiene 18 residuos de cisteína conservados. Las regiones altamente glicosiladas de E2 contienen epítopos de neutralización inmunodominantes que aparentemente funcionan como señuelos inmunológicos para proteger epítopos de neutralización más conservados. Por otro lado, el extremo N de la glicoproteína E2 puede ser muy variable y cambiar rápidamente incluso dentro de un mismo paciente(Albecka et al., 2011; Prentoe et al., 2016; Sabahi et al., 2014).



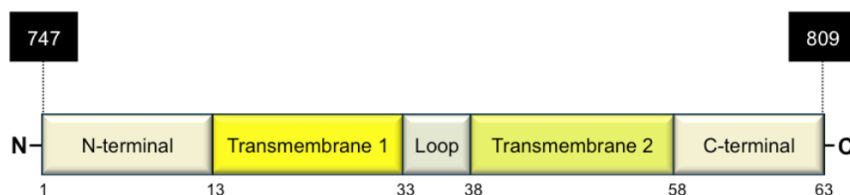
<sup>3</sup> Todos los esquemas mostrados pertenecen a Cognition Studio



#### **<sup>4</sup>P7**



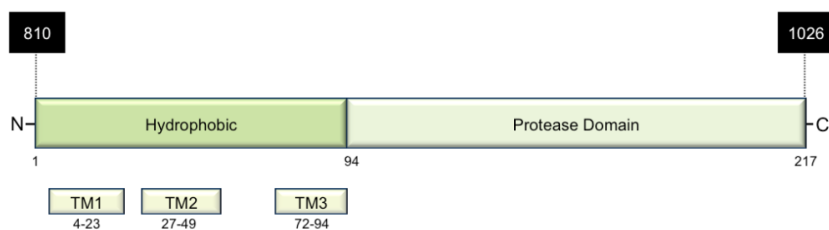
Es una proteína transmembranal hidrofóbica que participa en el ensamblado y la liberación del virus. Se oligomeriza para formar canales iónicos en las membranas de la célula huésped. Ayuda a descargar las proteínas del núcleo en conjunto con proteínas del R.E. Se compone de 63aa y tiene un peso de 7kDa. Tiene dos dominios transmembranales y un *loop* (Atoom et al., 2014; Chandler et al., 2012; Denolly et al., 2017; Khaliq et al., 2010; Ouyang et al., 2013).



#### **<sup>5</sup>NS2**

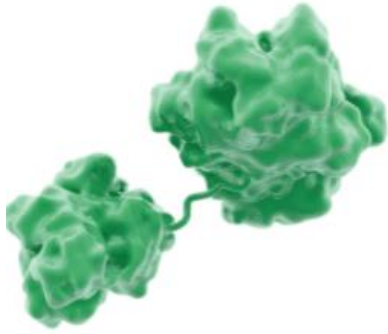


Tiene dos funciones en el ciclo viral del HVC: como cisteína proteasa para catalizar un único corte entre las proteínas NS2 y NS3, la liberación de estas proteínas es esencial para la replicación del ARN VHC y como cofactor en el proceso de ensamblaje. Se compone de 217 aa y tiene un peso de 21kDa. Tiene 3 dominios transmembranales, una pequeña hélice alfa y un dominio de proteasa (C. T. Jones et al., 2007; Lorenz, 2010; Lorenz et al., 2006; Popescu et al., 2011; Yi et al., 2009).



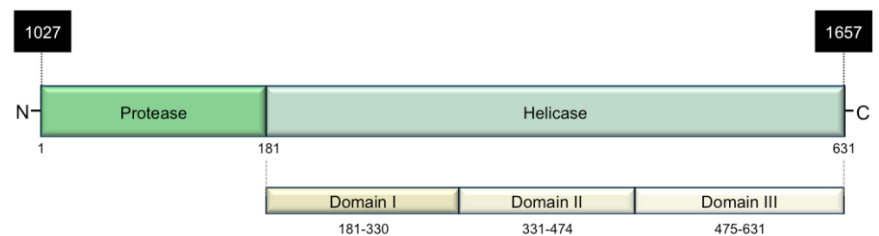
<sup>4</sup> Todos los esquemas mostrados pertenecen a Cognition Studio

<sup>5</sup> Todos los esquemas mostrados pertenecen a Cognition Studio



### <sup>6</sup>NS3

Es una enzima bifuncional que tiene actividad de serina proteasa y de helicasa. La proteasa realiza los cortes para liberar proteínas no estructurales. Además, esta enzima participa en la subversión de la respuesta inmune innata inactivando proteolíticamente factores de la célula huésped que pudieran bloquear la replicación viral. Se compone de 631aa y tiene un peso de 70kDa. Tiene dos dominios: en la región N terminal está el de la serina proteasa y en la región C terminal está el de la helicasa de ARN dependiente de nucleósido-trifosfatasa más grande (Brass et al., 2008a; de Francesco et al., 1999; Love et al., 1996; Morikawa et al., 2011a; Raney et al., 2010).



### <sup>7</sup>NS4A

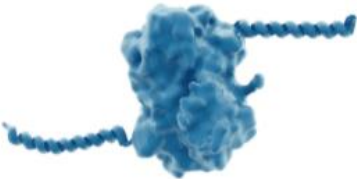
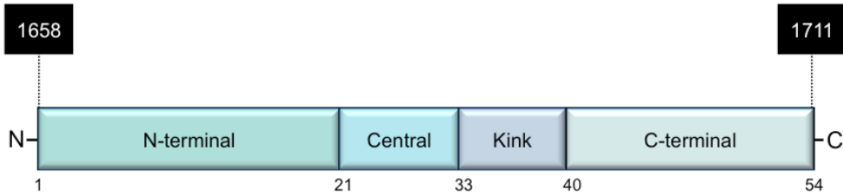


Tiene 4 funciones principales: (1) anclar el complejo NS3-4A a la pared externa del retículo endoplásmico y a la membrana externa mitocondrial, (2) servir de cofactor para la serina proteasa NS3A, (3) aumentar la actividad de la helicasa NS3A, y (4) regular la hiperfosforilación de NS5A y la replicación viral. NS4A y NS4B controlan la replicación; NS4A y NS3 participan en el ensamblaje. Se compone de 54 aa y tiene un peso de 27kDa. Tiene 3 dominios: la porción N terminal hidrofóbica, una región central para el plegamiento de NS3 y una porción ácida C terminal que se involucra en la replicación viral regulando

<sup>6</sup> Todos los esquemas mostrados pertenecen a Cognition Studio

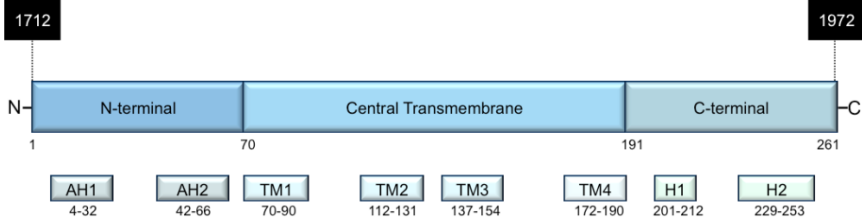
<sup>7</sup> Todos los esquemas mostrados pertenecen a Cognition Studio

NS5A(Brass et al., 2008b; Lin et al., 1997; Lindenbach et al., 2007; Morikawa et al., 2011b; Phan et al., 2011).

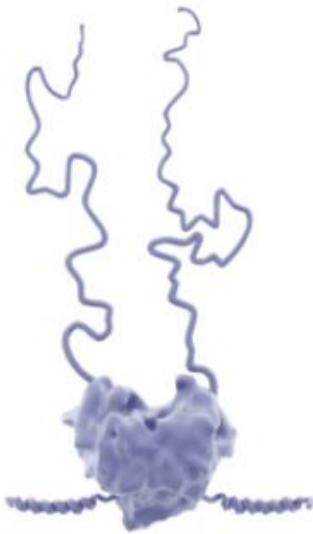


**<sup>8</sup>NS4B**

Es una proteína hidrofóbica que media las interacciones virus-huésped. Forma el complejo de replicación y es esencial para la formación de la red membranosa la cual se encarga de soportar la replicación viral. Se compone de 261 aa y tiene un peso de 27kDa. Antes de la escisión proteolítica, la NS4B está unida a la NS4A, con los componentes N-terminal y C-terminal situados en el lumen citosólico (Esser-Nobis et al., 2013; Gouttenoire et al., 2010; Lundin et al., 2006; Paul et al., 2013).

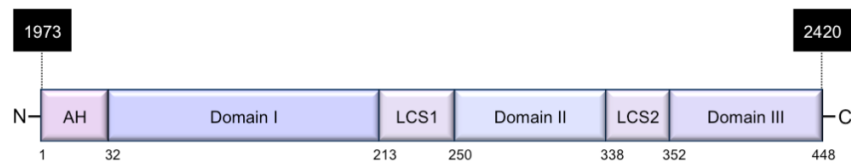


<sup>8</sup> Todos los esquemas mostrados pertenecen a Cognition Studio



### <sup>9</sup>NS5A

Es una fosfoproteína hidrofílica rica en prolina que se une al zinc. Interactúa con NS4B, NS5B, ARN y con proteínas de la célula huésped para regular la replicación y ensamblaje viral. Se compone de 448 aa y tiene un peso de 56kDa. Dentro de los hepatocitos infectados, NS5A se localiza en el R.E. donde forma vesículas de múltiples membranas inducidas por el virus, en esta red membranosa, a vesícula de múltiples membranas alberga complejos de replicación de ARN y gotas de lípidos (Appel et al., 2008; Berger et al., 2014; Bukh, 2016; Dustin & Rice, 2007; Zayas, Madan, et al., 2016).



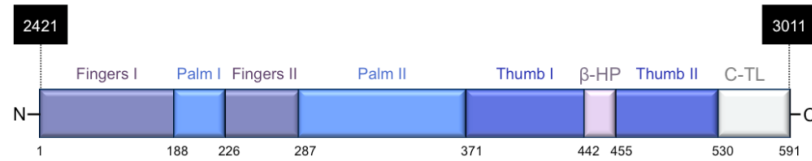
### <sup>10</sup>NS5B

Es una ARN polimerasa dependiente del ARN (RdRp) que desempeña un papel fundamental en la replicación del VHC. cataliza la polimerización de los ribonucleósidos trifosfatos (rNTP). Existen dos subclases principales de inhibidores de la polimerasa NS5B: (1) análogos de nucleótidos que imitan el sustrato natural e inducen la terminación de la cadena cuando se incorporan al nuevo ARN y (2) inhibidores no nucleotídicos que se unen a los sitios alostéricos de la enzima y perjudican su función. Se compone de 591 aa y tiene un peso de 66kDa (Ago et al., 1999; Appleby et al., 2015; Boyce et al., n.d.; Lesburg

<sup>9</sup> Todos los esquemas mostrados pertenecen a Cognition Studio

<sup>10</sup> Todos los esquemas mostrados pertenecen a Cognition Studio

et al., 1999; Lohmann et al., 2000; Love et al., 2003; Soriano et al., 2013).



### 1.3.1.3. Ciclo de infección

El ciclo de infección del VHC (Figura 5) comienza con una unión de baja afinidad a dos receptores del huésped en la superficie de un hepatocito: el receptor de lipoproteínas de baja densidad (LDLR) y los proteoglicanos de sulfato de heparán (HSPG). Esta interacción inicial desencadena que la proteína de membrana del heterodímero E1/E2 del VHC se una al receptor scavenger B1 (SRB1) y a la proteína tetraspanina CD81. Las interacciones entre estas proteínas crean un movimiento en la membrana lipídica, impulsando la partícula del VHC hacia una unión estrecha entre los hepatocitos (Agnello et al., 1999; Dubbuison, 2014; Lindenbach & Rice, 2013).

Cuando el VHC llega a la unión de oclusión, el CD81 interactúa con la claudina-1 (CLDN1), iniciando el plegamiento hacia el interior de la partícula vírica. Este proceso genera un endosoma interno compuesto por una partícula viral recubierta por la membrana de la célula huésped; todo el endosoma está rodeado por una jaula de clatrina (Blanchard et al., 2006; Evans et al., 2007; Farquhar et al., 2012; Lindenbach & Rice, 2013).

Posteriormente sucede la fusión en donde al entrar el virus en la célula, la jaula de clatrina que rodea el endosoma se dispersa, dejando la vesícula endosomal libre dentro del citosol. El pH ácido dentro del endosoma desencadena la fusión entre las membranas del virus y del huésped, en un proceso denominado fusión endosomal. Este acontecimiento permite que se desprenda la cubierta de la cápside y que el

ARN del VHC se libere en el citosol para su traducción y replicación (Koutsoudakis et al., 2006; Lindenbach & Rice, 2013; Sharma et al., 2011).

El proceso de traducción de la poliproteína del VHC se inicia cuando las subunidades ribosómicas se unen al ARN del VHC en el retículo endoplásmico rugoso. Posteriormente, el complejo ribosoma-ARN se une a la membrana del retículo endoplásmico y se completa la traducción de la poliproteína del VHC (Fraser & Doudna, 2007; Niepmann, 2013; Niepmann et al., 2018).

Una vez dentro del retículo endoplásmico rugoso, sucede el proceso proteolítico en donde las proteasas celulares escinden las proteínas estructurales, E1, E2 y p7. Después la cisteína proteasa NS2, junto con el extremo N-terminal de la proteína NS3, escinde la NS2 de la NS3. Por último, la NS3, ayudada por la NS4A unida a la membrana, forma un complejo de proteasas NS3-4A que escinde el resto de las proteínas (NS3, NS4A, NS4B, NS5A, NS5B). El resultado final son 10 proteínas maduras del VHC, que incluyen proteínas estructurales y no estructurales (Lindenbach & Rice, 2005).

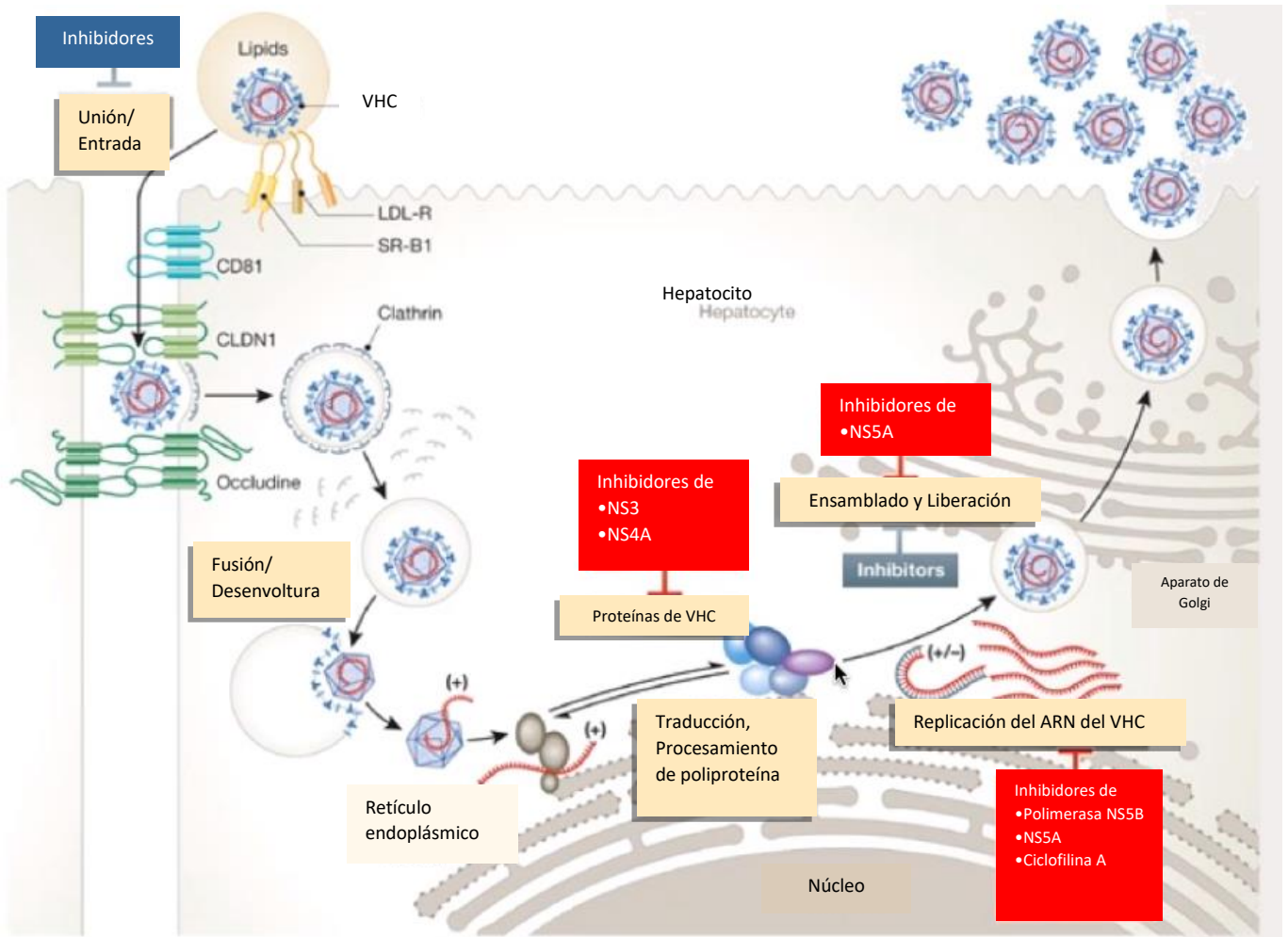
Posteriormente sucede la replicación del ARN. Varias proteínas del VHC inducen la reorganización de las membranas de la célula huésped para formar la red membranosa. En la red membranosa, la ARN polimerasa dependiente del ARN NS5B cataliza la síntesis de un ARN intermedio de sentido negativo (plantilla) que se utiliza para crear numerosas copias de ARN del VHC de sentido positivo. Estos ARN del VHC recién sintetizados se incorporan a las partículas de la nucleocápside o se utilizan para la traducción y replicación del ARN (Lohmann et al., 2013; Romero-Brey et al., 2012; Romero-Brey & Bartenschlager, 2014).

Una vez terminada la replicación, comienza el ensamblaje. Las partículas del VHC se ensamblan cerca de las gotas lipídicas citosólicas (cLDL) y, con la ayuda de la enzima diacilglicerol acetiltransferasa-1 (DGAT1) del huésped, forman la nucleocápside. Las proteínas del núcleo celular utilizan el ARN del VHC como andamio, ya que éstas forman una cubierta protectora alrededor del ARN del VHC. Una vez formada la cubierta protectora, la partícula inmadura del VHC se fusiona

con una gota lipídica luminal (LuLD) cargada con proteínas ApoE para crear un precursor del VHC de alta densidad. Simultáneamente, el RE sintetiza proteínas de muy baja densidad (pre-VLDL); el precursor del VHC de alta densidad y el pre-VLDL transitan hacia el Golgi, donde maduran antes de ser empaquetados y liberados (Bartenschlager et al., 2011; Crouchet et al., 2017; Lee et al., 2014; Paul et al., 2014; Shi & Suzuki, 2018; Stewart et al., 2016; Zayas, Long, et al., 2016).

Antes de ser liberadas las nuevas partículas virales en conjunto con la gota de lípidos, éstas necesitan madurar. En el Golgi, existe evidencia de que las pre-VLDL se fusionan con grandes gotas lipídicas ricas en triacilglicerol (TG) para formar las VLDL. A continuación, las VLDL se fusionan con los precursores del VHC de alta densidad para formar la partícula lipoviral del VHC. Esta partícula lipoviral del VHC de baja densidad abandona la red trans-Golgi (TGN) en vesículas de transporte especializadas conocidas como cuerpos multivesiculares. La maquinaria de secreción celular transporta los cuerpos multivesiculares a la superficie celular (Coller et al., 2012a; Gastaminza et al., 2008a; D. M. Jones & McLauchlan, 2010a; Syed et al., 2017).

Tras el transporte de los cuerpos multivesiculares que contienen las partículas lipovirales del VHC a la superficie celular, las vesículas se fusionan con la membrana celular del hepatocito. Durante este proceso tanto las partículas lipovirales del VHC como las partículas VLDL se liberan en los compartimentos extracelulares (Coller et al., 2012b; Corless et al., 2010; Gastaminza et al., 2008b; D. M. Jones & McLauchlan, 2010b; Lindenbach, 2013).



**Figura 5** Ciclo de replicación del HVC. Lange et al., EMBO Mol Med, 2013



## 1.4. Variabilidad genética del VHC

En 1995 Simmonds clasificó aislamientos del VHC basados en la variación de la secuencia NS5B y se observó que se agrupaban en 6 tipos las cuales a su vez se dividen en otros subtipos (Hajime et al., 1998). Sin embargo, en el 2015 se encontró la circulación, en individuos de la República Democrática del Congo, un nuevo genotipo (QC69) proponiéndolo como el séptimo grupo (DEV, 2021; Murphy et al., 2015) y en Punjab, India en el 2018 se identificó el octavo genotipo (Borgia et al., 2018).

Las diferencias que mejor caracterizan a los genotipos se encuentran en las regiones NS5b, E1 y el *core* (del Carmen et al., 2015). Las secuencias nucleotídicas varían entre genotipos en un 31%-33% y entre subtipos en un 20%-25% (Tabla 4). Estas variaciones son debidas a la alta tasa de mutación viral ( $1.1 - 1.5 \times 10^3$  mutaciones o sustituciones de nucleótidos por sitio por genoma por año) y a la falta de actividad correctora de la ARN polimerasa (Zulay Sulbarán et al., 2016) la cual tiene un error de  $10^{-4}$  (del Carmen et al., 2015).

**Tabla 4** Variabilidad genética del VHC

	<b>% Homología</b>	<b>Símbolo</b>	<b>Origen</b>
Genotipo	66-69	Números (1-8)	Heterogeneidad del virus
Subtipo	77-80	Letras	Virus relacionados dentro del mismo genotipo
Aislado	81-90	---	Especies dentro del mismo subtipo en distintos individuos
Cuasiespecie	91-99	---	Mutaciones en el individuo durante la replicación viral

\*Modificado de Llanes et. al 2015

Los genotipos se identifican con números y el VHC se clasifican en 8 (1:8) y los subtipos se identifican con letras en minúscula (a,b,c...) y de éstos son 67 (Kuiken & Simmonds, 2009; Simmonds, 1995; Simmonds et al., 2005; Smith et al., 2014). También existen las cuasiespecies, las cuales son secuencias que tienen hasta una homología del 98%.

Estas cuasiespecies se originan dentro del mismo huésped dando lugar a secuencias genómicas heterogéneas del virus (del Carmen et al., 2015) como consecuencia de un mecanismo de evasión de la respuesta sistema inmune del hospedero. Por lo tanto, la población viral que circula en el individuo afectado tiene un espectro de mutación que está en continuo cambio (Lapa et al., 2019).

Las tecnologías de secuenciación, a través de la filogenética, son esenciales para poder diseñar estrategias de predicción, prevención y tratamiento ya que permiten la detección y control de brotes epidemiológicos.

### **1.5. Resistencia a antivirales de acción directa**

Como se ha mencionado anteriormente, el surgimiento de los antivirales de acción directa ha marcado las pautas para el inicio de la nueva era de un tratamiento más eficaz contra esta enfermedad.

Antes del 2011, los tratamientos se administraban de seis meses a un año, ahora el tratamiento se ha reducido a un periodo de entre 8-24 semanas (Pawlotsky et al., 2020). Otra de las ventajas brindadas por este nuevo esquema, es la eficacia para alcanzar una RVS. Se ha reportado una eficacia de entre el 80-95% para obtener una RVS con el tratamiento (Chen et al., 2017; Childs-Kean & Hong, 2017; Ji et al., 2016; Krüger et al., 2019), no obstante, el éxito o no de lograr una RVS, se ha visto asociado con el tratamiento previo del paciente y los polimorfismos que presenta el paciente (Dawood et al., 2021; Pár et al., 2014; Thomas et al., 2009).

El problema emergente que enfrenta la eficacia del tratamiento, son las mutaciones que ocurren durante la replicación viral, las cuales pueden otorgar potenciales resistencias al tratamiento farmacológico. Estas mutaciones se conocen como sustituciones asociadas a resistencia (RAS, por sus siglas en inglés), las cuales son una clase de polimorfismo que varía de acuerdo con el genotipo, subtipo y antiviral utilizado. Se ha reportado, a lo largo de la última década, que la frecuencia de aparición de estas sustituciones ha incrementado en 15-40% tras la incorporación de los DAA al esquema de tratamiento (Howe et al., 2022), por lo tanto es de importancia clínica la identificación de éstas para tener un panorama más asertivo en la toma de decisiones para el tratamiento del paciente.

De acuerdo con la guía médica de la AASLD (American Association for the Study of the Liver Diseases) estas sustituciones se pueden clasificar en 3 clases: sencilla (single-class resistance), doble (dual-class resistance) y triple (triple-class). La clase de resistencia es la resistencia que presenta el virus contra una clase específica de DAAs. Las sustituciones sencillas son resistencias a los inhibidores de la NS5A (INS5A) o IN o IP; las sustituciones dobles son resistentes a INS5A e IN, o INS5A e IP o IN e IP; las sustituciones triples son resistencias a las tres clases de DAA (Howe et al., 2022).

Algunas de estas sustituciones de importancia clínica se presentan en la Tabla 4 (Ceccherini-Silberstein et al., 2018). También existe la base de datos SHARED de la universidad *The University of British Columbia*, la cual contiene metadata de las secuencias del VHC incluyendo las mutaciones que le confieren resistencia al virus a los DAA.

**Tabla 5** RAS identificadas en los genes NS3, NS5A y NS5B debido a la administración de tratamientos con DAAs en 6 genotipos del VHC. Ceccherini-Silberstein et al., 2018.

Table 1					
Summary of substitutions associated with resistance to protease NS3 inhibitors, NS5A inhibitors, and nucleoside and non-nucleoside NS5B inhibitors					
NS3 RASs	Glecaprevir	Grazoprevir	Paritaprevir	Simeprevir	Voxilaprevir
<b>GT1a</b>	V36M R155T A156G/T/V D/Q168A/V	V36L/M Y56H/E Q80K R155K/I/Q/S/T A156G/M/S/T/V V158A D168A/C/E/G/K/N/T/V/Y	V36A/M F43L Y56H Q80K/R R155G/K/S/T/W A156G/S/T D168A/E/F/H/N/V/Y	V36L/M T54S Q80K/R S122N/R R155G/K/T/Q A156G/T/V D168A/E/ H/N/V I170T/V	V36G/L Q41R F43S Q80K R155G/K/W A156L/T/V D168A/F/I/K/L/R/T/V
<b>GT1b</b>	A156T/V D/Q168V	F43S Y56H/E Q80L R155G/K/L/T A156T/V D168A/E/G/H/K/V/Y	Y56H Q80K/R R155K/T A156T/V D168A/E/H/K/N/T/V/Y	V36A/M Q41R F43I/S/V T54A/S Q80G/H/K/R S122R/I S138T R155G/K/Q/T/W A156G/S/T/V D168A/G/E/F/H/I/N/Q/T/V/Y V170T	V36A/M F43S S122D R155W A156S/T/V D168V/Y V170A
<b>GT2</b>	D/Q168A/E/V  Y56H/N Q80K/R	NR  NR	NR  NR	NR  NR	F43V A156L/T/V Q41K Q80K/R A156T/V L175M
<b>GT3</b>	A156G A166T D/Q168K/L/R/T				
<b>GT4</b>	R155C D/Q168V	Y56H R155K A156S/T D168A/V	Y56H R155C/K A156H/K/S/T/V D168A/E/H/V	Q80R R155K A156G D168A/E/V	Q41R A156L/T/V D168E/T/V
<b>GT5</b>	None		NR	NR	A156T/V D168A/H/K/R/Y
<b>GT6</b>	D/Q168Y	F43C A156S D168A/E/G/N/V/Y	NR	NR	Q41K/R Y56H A156T/V D168A/H

NS5A RASs	Daclatasvir	Elbasvir	Ledipasvir	Ombitasvir	Pibrentasvir	Velpatasvir
<b>GT1a</b>	K24R M28 <u>A/T</u> Q30 <u>D/E/G/H/K/N/R/Y</u> L31 <u>I/M/V</u> P32L H58 <u>D</u> Y93 <u>C/E/H/N</u>	M28 <u>A/G/S/T</u> Q30 <u>D/E/G/H/K/R/Y</u> L31 <u>F/I/M/V</u> H58D Y93 <u>C/H/N/S</u>	K24 <u>R</u> M28 <u>A/G/T/V</u> Q30 <u>E/G/H/K/L/N/R/Y</u> L31 <u>F/I/M/V</u> P32L S38 <u>F</u> H58 <u>D</u> A92T Y93 <u>C/E/H/N/S</u>	K24 <u>R</u> M28 <u>T/V</u> Q30 <u>E/H/K/L/R/Y</u> L31 <u>V</u> H58 <u>D</u> Y93 <u>C/E/H/L/N/S</u>	K24 <u>R</u> M28 <u>A/G</u> Q30 <u>D/K/R</u> H58 <u>D</u> Y93 <u>H/N</u>	K24 <u>R</u> M28 <u>A/G/T/V</u> Q30 <u>E/G/H/K/L/R</u> L31 <u>F/I/M/V</u> P32 <u>L</u> H58 <u>D</u> A92K Y93 <u>C/H/L/N/R/S/T/W</u>
NS5A RASs	Daclatasvir	Elbasvir	Ledipasvir	Ombitasvir	Pibrentasvir	Velpatasvir
<b>GT1b</b>	L28M/T R30 <u>G/H/Q</u> L31 <u>F/I/M/W/V</u> P32L/ <u>del</u> P58 <u>A/L/S</u> A92K Y93 <u>C/H/I/N/R</u> F/L28C/S L31M	L28M R30Q L31 <u>F/M/V</u> Y93 <u>H</u>	L28M/T R30 <u>H</u> L31 <u>F/I/M/V</u> P32L P58D A92K/T Y93 <u>C/H</u>	L28M/T R30Q L31 <u>F/V</u> P58S A92T Y93 <u>H/N/S</u>	P32 <u>del</u>	Q24K R30S L31 <u>F/I/M/V</u> P58 <u>R/T</u> A92K Y93 <u>C/H/R/N/S/T</u>  F/L28S/F L/M31I/M/V P58A C92R/S/T Y93F/H/N
<b>GT2</b>	C92R Y93H	NR	NR	NR	None	

Table 1 (Continued)

NS5A RASs	Daclatasvir	Elbasvir	Ledipasvir	Ombitasvir	Pibrentasvir	Velpatasvir
<b>GT3</b>	M28I A30 <u>K</u> L/V31F/ <u>I/M/V</u> A62L Y93 <u>H</u>	NR	NR	NR	S24F M28 <u>G/K</u> A30 <u>G/K</u> L/V31F/ <u>I/M</u> P58T Y93 <u>H</u>	A30H/ <u>K/V</u> L/V31F/ <u>M/P/V</u> P58G Y93 <u>H/N/R</u>
<b>GT4</b>	L/I28M/ <u>V</u> L/R30G/ <u>H/R/S</u> L/M31V Y93 <u>H/R/S/W</u>	L/I28I/ <u>M/S</u> L/R30F/ <u>H/R/S</u> L/M31M P/T58D Y93C/ <u>H/S</u>	L/I28M/ <u>V</u> L/R30H/ <u>R/S</u> M/L31I/ <u>L/V</u> P/T58L Y93C/ <u>H/S/W</u>	K24Q L/I28A/ <u>C/I/S/V</u> L/R30R M/L31I/ <u>L</u> Y93H/ <u>S</u>	None	L/I28T/ <u>V</u> L/R30R L/M31M Y93 <u>H/N/S</u>
<b>GT5</b>	L31F/V	NR	None	NR	None	None
<b>GT6</b>	Q/K24H L/I31M P32L/S T/S/G58A/N/S	NR	None	NR	None	F/T/V28M/V L/I31I/M/V P32A/L/Q/R T/S/G58G/H A/P92T T93A/H/N/S
<b>GT8</b>			Q30S Y93S			
NS5B RASs					Dasabuvir	Sofosbuvir
<b>GT1a</b>					L314H C316 <u>Y</u> A395G M414I/ <u>I/V</u> N444K E446K/Q Y448 <u>C/H</u> A553 <u>T/V</u> G554 <u>S</u> S556G/ <u>N/R</u> G558R D559G Y561H S565F	L159F E237 <u>G</u> S282 <u>R/T</u> C316 <u>F</u> L320F V321A
NS5B RASs					Dasabuvir	Sofosbuvir

**GT1b**

C316H/N/Y  
S368T  
N411S  
M414I/T/V  
C445F  
Y448C/H  
C451S  
A553V  
S566G  
D559G/N

L159F  
S282G/T  
C316F/H/N  
V321I

Para la identificación de estas mutaciones se pueden llevar a cabo análisis de resistencia genotípica (GTR, por sus siglas en inglés) los cuales utilizan las tecnologías de secuenciación. Las más comunes son la secuenciación de Sanger, que es capaz de identificar variantes con una prevalencia >15-20% y las de tercera generación, que identifican variantes con una prevalencia >0.1-1% (Pawlotsky et al., 2020).

## 1.6. Secuenciación con Oxford Nanopore

El avance en las tecnologías para los análisis moleculares ha tenido un crecimiento exponencial en la última década. Estas tecnologías han permitido identificar factores y/o características específicas para entender mejor el entorno objetivo estudiar, los cuales van desde estudios poblacionales como GWAS (Estudio de Asociación de Genoma Completo) hasta el transcriptoma de una célula y secuenciación de genomas para la construcción de análisis filogenéticos.

Una de estas grandes innovaciones son los secuenciadores de tercera generación como el Nanopore (Figura 7). Éstos consisten en una pieza de silicón o grafeno que contiene nanoporos de 1nm de diámetro. Esta placa contiene enzimas que “guían” al ADN para que pase a través del nanoporo, el cual, debido a su tamaño, obliga al ADN a atravesarlo como una sola hebra base por base.

Los dispositivos ONT (Oxford Nanopore Technologies) son equipos que convierten y toman miles de lecturas de corriente por segundo. La translocación procesiva de una molécula de ADN o ARN conduce a un cambio de corriente característico que está determinado por múltiples nucleótidos consecutivos (es decir, k-mer) definidos por la longitud de la región de detección de nanoporos. La medición de corriente sin procesar se puede segmentar según el cambio de corriente para capturar señales individuales de cada k-mer. Cada segmento de corriente contiene múltiples mediciones, y la media, la varianza y la duración correspondientes de las mediciones de corriente juntas forman los datos del “evento”. La dependencia de los datos de eventos en los nucleótidos vecinos es similar a la cadena de Markov, lo que hace que los métodos basados en HMM (Hidden Markov Model) sean una combinación natural para decodificar los cambios actuales en la secuencia de nucleótidos (Wang et al., 2021).

El Nanopore se sumerge en un fluido conductor y se aplica un voltaje que lo atraviesa. La corriente es muy sensible al tamaño y forma del nanoporo, por lo tanto, la cantidad de corriente que pasa a través del nanoporo en un momento específico varía dependiendo de la base (A, C, G o T) que lo esté bloqueando. El cambio en la corriente en el nanoporo cuando pasa la molécula de ADN representa una medición

directa de la secuencia de ADN, identificando un orden correcto de las bases (Raza & Ameen, 2017).



**Figura 6** Tecnologías Oxford Nanopore. A) Principio de funcionamiento del Nanopore. La hebra de DNA (azul) pasa a través del nanoporo guiada por una enzima (morado) lo cual genera una corriente eléctrica detectable (cyan). B) Se muestran 5 dispositivos disponibles por Oxford Technologies, de izquierda a derecha: Flongle, MinION, MinION Mk1C, GridION, PromethION.

Dentro de las aplicaciones más conocidas se encuentran: secuenciación de genoma completo, Secuenciación para identificación de secuencias objetivo, Secuenciación metagenómica, Epigenética, Microbiología, Genética humana, Genética clínica, Investigación de plantas e Investigación en animales (Oxford Technologies, 2021).

El minION es un dispositivo portable de secuenciación de la casa comercial Oxford Technologies que permite la secuenciación de lecturas largas en tiempo real. En el 2019 se obtuvo un promedio de lecturas de secuencias largas de hasta 23kb, una lectura individual de 2,273kb y un máximo de lectura de hasta 2,273Mb con protocolos específicos de extracción de ADN. La calidad de las lecturas puede



mejorarse secuenciando la misma molécula de ADN de doble cadena varias veces (Wang et al., 2021).

En comparación con otras tecnologías de nueva generación para secuenciación (NGS) como Illumina, la secuenciación mediante nanoporo MinION representa una ventaja económica en cuanto a la reducción del costo por muestra de casi el 50%(Riaz et al., 2021).

En el caso de las plataformas de pirosecuenciación 454, ésta ofrece lecturas más largas que Illumina, su precisión para diferenciar SNP (single nucleotide polimorfism) es reducida. Illumina ofrece esta ventaja de diferenciación de SNPs sin embargo, las lecturas cortas pareadas impiden la reconstrucción de las variantes virales. Por otro lado, la segunda opción de las NGS de tercera generación, PacBio, ofrece lecturas largas y voluminosas que exceden el genoma de muchos virus de ARN y requieren de condiciones sofisticadas en el laboratorio elevando a su vez los costos para la secuenciación(Rodrigo et al., 2017).

Uno de los inconvenientes que presenta la secuenciación de nanoporo es que las lecturas largas presentan una tasa de error relativamente alta (5-20%) de acuerdo con el tipo de molécula y los métodos de preparación de la librería, estos errores incluyen inserciones y eliminaciones, por lo que la corrección de errores normalmente requiere de datos de secuencia de lectura corta adicionales (Kono & Arakawa, 2019). Es por ello que para efectos esta tesis se ha seleccionado un tamaño de fragmento a secuenciar de 400pb.

El tratamiento de datos bioinformáticos se realiza a través del software MinKNOW el cual no requiere de recursos informáticos de alta gama ni habilidades avanzadas para el procesamiento de datos. Este software administra la adquisición de datos, el *base calling* (decodifica la señal actual a la secuencia de nucleótidos) y el análisis en tiempo real, generando archivos binarios en formato fast5 para almacenar metadatos e información de lectura (Wang et al., 2021). En la Tabla 6 se resumen los resultados obtenidos durante el tratamiento de los datos crudos de secuenciación y el formato de archivo correspondiente a cada análisis.

**Tabla 6** Tipo de archivos utilizados en análisis bioinformático

<b>Data</b>	<b>Formato</b>
<b>Datos crudos</b>	.fast5 y .fastq
<b>Alineamiento</b>	.sam / .bam / .cram
<b>Genoma de referencia</b>	.fasta
<b>Archivos de anotación</b>	.gtf / .gff / .bed
<b>Resultados finales</b>	.vcf / .tsv / .bcf / .gVCF

Oxford Nanopore “London Basecalling 2022”

Para la epidemiología de las enfermedades es importante realizar análisis filogenéticos que sean capaces de identificar clústers del virus y, con base en esto, lograr la identificación de tratamientos eficaces y preventivos.

## 1.7. Análisis filogenético

Uno de los retos más grandes a considerar para el desarrollo de una vacuna global para la infección de la hepatitis C es la variabilidad genética que tiene el virus y la creación de cuasiespecies que se originan en el huésped como mecanismo de defensa.

La filogenética permite el “rastreo” de la rápida evolución del virus mediante la identificación de similitudes genómicas virales en diferentes huéspedes. Para la construcción de las relaciones filogenéticas se requiere principalmente de dos pasos bioinformáticos: filtrado y alineamiento de secuencias.

El alineamiento múltiple de secuencias es capaz de detectar residuos funcionales para inferir la historia de evolución de una proteína. En comparación con otros métodos de alineamiento como CLUSTALW, MAFFT es más rápido, tiene una mejor precisión para hacer los alineamientos, se pueden introducir sets de datos amplios (~1,000,000bp) y, dentro de sus mejores características, el algoritmo empleado es capaz de ajustar automáticamente la decisión de algoritmos utilizados en el programa para proveer el alineamiento más adecuado para cada set de datos (D.S. Guttman & N.J. Provart, 2021; Katoh et al., 2002).

Como parte de las aportaciones que estos análisis pueden ofrecer, se menciona el trabajo de Rodrigo et. al. (2017) quienes realizaron un estudio filogenético transcontinental en el cual se ha reportado que las variables genéticas que llevan a la creación de los clústers, se deben principalmente a las regiones de envoltura E1 y E2 y la NS5B. Además, demostraron que las relaciones filogenéticas con secuencias del genoma completo tienen mejor resolución a diferencia de las construidas únicamente con las secuencias core, E2 y NS5B.

## 2. Antecedentes

La identificación del tipo de VHC que se ha caracterizado en México comenzó en la primera década del 2000 y estudios posteriores no han demostrado una divergencia en cuanto al genotipo que prevalece en el país.

En el 2007, Dehesa-Violante et. al, realizaron un estudio para conocer la genotipificación del VHC en México. Se reportó que en la región norte (Culiacán, Monterrey, Ciudad Obregón y Tijuana) se encontraron 68 genotipos tipo 1 (en su mayoría correspondientes al subtipo 1b) y 26 no pertenecientes al genotipo 1. Mientras que en la región central-sur (Guadalajara, Puebla, Veracruz, León y Ciudad de México), se reportaron 229 pacientes con el genotipo 1 y 98 no pertenecientes al genotipo 1. El genotipo mayor encontrado después del 1 fueron subtipos pertenecientes al genotipo 2, posteriormente el 3 y al último, con 2 muestras, el genotipo 4. La confirmación de la infección se realizó mediante el ensayo ELISA.

Después, en el 2010, Jiménez-Méndez et. al, hicieron una identificación de la distribución geográfica del VHC y su carga viral. En este estudio se reportan 5 genotipos de VHC en México (1,2,3,4 y 5) siendo el de mayor prevalencia el genotipo 1 en todo México, seguido del genotipo 3 en las regiones Sur, Este y la Península y los genotipos 4 y 5 únicamente en el Noreste, Noroeste y Península. La prevalencia del genotipo 1, siendo el subtipo más frecuente el 1b, coincidiendo con el estudio de Dehesa. Esto se debe a que este es el genotipo que presenta una mayor resistencia al tratamiento con PEG-INF + Ribavirina. La genotipificación se realizó mediante el ensayo INNO-LiPA® VHC II.

Durante el 2019 se realizó un estudio de actualización de las cepas circulantes en México por Sedeño-Monge et. al. en el cual se observó una prevalencia del subtipo más frecuente fue el 1b, seguido del 1a y por último 2b. De igual forma se menciona que el subtipo más abundante en la región norte es el 1a seguido del genotipo 2.

Junto con los avances de las tecnologías de nueva generación, también van de la mano los avances en diagnóstico. Tal es el caso de un estudio realizado en un paciente de origen chino para lograr la secuenciación del genoma viral del VHC en donde se identificó el genotipo 3b mediante la tecnología de secuenciación de

tercera generación Oxford Nanopore, obteniendo una profundidad de 73.6 NS y una lectura total de 9442 pb (Uchida et al., 2019).

Actualmente solo se ha reportado la secuenciación de un genoma de VHC del genotipo 2b en Latinoamérica realizado por Spitz et. al. (2019) en Brasil. La secuenciación se realizó mediante el kit Big Dye Terminator v3.1 de Applied Biosystems (Nátalia et al., 2019).

## **2.1. Antecedentes directos**

En nuestro de trabajo se ha llevado a cabo la investigación de los genotipos circulantes en el Norte de México.

El grupo de investigación de la doctora Rivas-Estilla ha monitoreado en las últimas décadas la prevalencia de los genotipos circulantes del VHC en el norte del país.

En el primer trabajo de investigación (Rivas-Estilla et. al, 1998) se reportó que el GT 1a fue el prevalente en el oeste de México. La genotipificación se realizó con el ensayo INNO-LiPA ® VHC II y la presencia viral fue detectada con RT-PCR.

Diez años después, esta prevalencia cambia, reportando el GT 1b como el prevalente en el noreste del país (Rivas-Estilla et al., 2008). La genotipificación se realizó igualmente con el ensayo INNO-LiPA ® VHC II pero el método para la detección del ARN viral cambia, utilizando el ensayo ELISA 3ª generación.

Hasta ahora no se ha secuenciado el genoma completo del VHC circulante en México. No obstante, gracias a los resultados previos, se conocen los genotipos circulantes en el norte (GT 1a y GT 1b). Secuenciar los genotipos circulantes, permite la construcción filogenética de los VHC para la predicción de brotes epidemiológicos. Además, conocer las secuencias virales aporta conocimiento para entender mejor la biología molecular del virus y permite la identificación de las fallas a los tratamientos antivirales actuales.

### **3. Justificación**

En México la infección por el VHC es una de las tres causas más frecuentes de cirrosis hepática, ésta representa una de las principales causas de muerte en la población mexicana y la principal indicación de trasplante hepático. Actualmente, no existen estudios de secuenciación de genoma completo de los genotipos circulantes en México del VHC.

En la nueva era de la Medicina Molecular es necesario conocer las características de los subtipos virales circulantes para así asociarlas en árboles filogenéticos y evaluar las mutaciones que se han presentado en el genoma viral mediante la valoración de áreas específicas de éste para que nos puedan brindar información antigénica y llevar a cabo la identificación de posibles resistencias a los nuevos antivirales específicos, su evolución y prevalencia.

Por otra parte, los análisis filogenéticos han demostrado un cambio constante de clados o linajes en la dinámica del VHC. Estos reemplazos de linajes pueden estar asociados a brotes epidémicos y evolución del virus. Además, las mutaciones pueden mejorar la capacidad replicativa y una mejor adaptabilidad en la población mexicana, así como establecer resistencia a los antivirales específicos.

El uso de las nuevas plataformas de secuenciación para la elucidación de genomas virales permite el estudio de la filogenia, transmisión y evolución de estos virus patógenos para el humano. Esto habilita la generación de nuevo conocimiento y bases de datos para futuras predicciones de brotes, estimar tasas de evolución, tiempos de divergencia y su respuesta oportuna.

#### **4. Hipótesis**

La secuencia genética del VHC circulante en México durante el periodo 2016-2020 es igual a las reportadas en los últimos años a nivel internacional y carece de mutaciones asociadas a resistencia a antivirales.



## **5. Objetivos**

### **5.1. Objetivo general**

Secuenciar el genoma completo, con la tecnología de tercera generación Oxford Nanopore, del virus de la hepatitis C (VHC) a partir de aislados de suero de pacientes infectados para identificar la variabilidad genómica, establecer las relaciones filogenéticas y variantes circulantes contemporáneas.

### **5.2. Objetivos específicos**

- 1) Aislar, amplificar e identificar el ARN viral del VHC a partir de muestras de suero de pacientes con diagnóstico confirmado de hepatitis C en México.
- 2) Secuenciar los genomas virales mediante tecnología de tercera generación Oxford Nanopore, de virus aislados de pacientes infectados.
- 3) Realizar análisis filogenéticos de los virus secuenciados para establecer las relaciones filogenéticas entre las variantes circulantes en México y las variantes internacionales e identificar mutaciones asociadas a resistencia a antivirales en los virus contemporáneos.

## **6. Materiales y Métodos**

Los reactivos y volúmenes utilizados se mencionan en la descripción de cada procedimiento y en los Anexos de este trabajo.

### **6.1. Materiales**

#### Materiales biológicos

Muestras de suero de pacientes con diagnóstico confirmado de Hepatitis C por RT-PCR sin coinfecciones. Se obtuvieron 53 sueros de pacientes de la Unidad de Hígado del Hospital Universitario “José E. González” de la Universidad Autónoma de Nuevo León. Los sueros permanecieron preservados con EDTA a  $-80^{\circ}\text{C}$  en la seroteca de la Unidad de Hígado. Se descongelaron las muestras y se tomaron alícuotas de  $500\mu\text{L}$  para posteriormente ser almacenados nuevamente a  $-80^{\circ}\text{C}$  hasta su uso.

#### Reactivos

Los reactivos utilizados se mencionan en los anexos correspondientes

#### Equipos

- Secuenciador Oxford Nanopore MinION No. Serie 35308
- Computadora Huawei con procesador AMD Ryzen 7 3700U con Radeon Vega Mobile Gfx 2.30 GHz Windows 11 RAM 8GB

## 6.2. Metodología

### 6.2.1. Muestras biológicas

Las muestras de suero fueron proporcionadas por la Unidad de Hígado del Hospital Universitario “José Eleuterio González”. Los pacientes tenían un diagnóstico confirmado de hepatitis C en estado crónico mediante RT-PCR. El rango de la carga viral para la selección de las muestras fue  $>200,000\text{UI}$  y se escogieron los genotipos más frecuentes los cuales incluyen 1b y 1a.

Los criterios de selección se describen en la Tabla 7.

**Tabla 7** Criterios de inclusión y exclusión de los pacientes donadores de muestras

Criterios de inclusión	Criterios de exclusión	Criterios de eliminación
<ul style="list-style-type: none"><li>▪ Pacientes <math>&gt;18</math> años</li><li>▪ Diagnóstico confirmado Hepatitis C</li><li>▪ Carga viral detectable de VHC</li></ul>	<ul style="list-style-type: none"><li>▪ No existe carga viral de VHC sérica detectable.</li><li>▪ La carga viral de VHC <math>&lt;100,000\text{UI/ml}</math> en sangre.</li><li>▪ Coinfección de otras hepatitis víricas, VIH u otros virus de ARN.</li></ul>	<ul style="list-style-type: none"><li>▪ No tener acceso al expediente clínico del sujeto de investigación</li></ul>

### 6.2.2. Extracción de ARN

La extracción de ARN se llevó a cabo a partir de muestras de suero mediante el protocolo TRIzol LS Reagent de Invitrogen. Los materiales requeridos se mencionan en el anexo 1. Brevemente, se añadieron 750µL de TRIzol LS Reagent por cada 250µL de muestra en tubos Eppendorf de 1.5mL y se homogeniza la muestra mediante pipeteo. Se incubaron a temperatura ambiente durante 5 min y posteriormente se añadieron 200 µL de cloroformo por cada 750 µL de trizol y se incubó durante 2-3min a temperatura ambiente. Se centrifugó durante 15 min a 12,000 × g a 4°C. Se transfirió la fase acuosa que contiene el ARN a un nuevo tubo Eppendorf de 1.5ml evitando transferir la interfase. Se añadieron 500 µL de isopropanol a la fase acuosa y se incubó durante 10 min a temperatura ambiente. Se centrifugó durante 10 min a 12,000 × g a 4°C y se descartó el sobrenadante. Para el lavado del ARN se re suspendió la pastilla en 1mL de etanol al 75% por cada 750 µL de TRIzol LS Reagent utilizado para la lisis. Se agitó mediante vórtex y se centrifugó a 5 min a 7500 × g at 4°C, se descartó el sobrenadante mediante micropipeta. Finalmente se secó la pastilla de ARN a temperatura ambiente durante 5-10 min, se resuspendió en 50 µL de agua libre de ARNsas y se incubó en termo block a 55-60°C durante 10-15 min.

### 6.2.3. Síntesis de ADN complementario (ADNc)

La síntesis de ADNc se llevó a cabo mediante el protocolo M-MLV Reverse Transcriptase de Invitrogen. Antes de comenzar se programó el termociclador con las temperaturas a utilizar. En un microtubo se añadieron los oligonucleótidos, dNTP Mix 10m y la plantilla de ARN. Se incubó a 65°C durante 5 min., después de los 5 min se enfría rápidamente en hielo y se da una centrifugación breve. Posteriormente se añadió el Buffer 5X-First Strand, 0.1M DTTM y ARNseOUT Recombinant Ribonuclease Inhibitor. Se mezcla suavemente y se incuba a 37°C durante 2 min. Finalmente, se añade 1µL de enzima M-MLV RT y se incuba a 25°C durante 10 min, después a 37°C durante 50 min y por último a 70°C durante 15 min para inactivar la reacción. Las cantidades de cada reactivo utilizadas se muestran a continuación:

**Tabla 8** Volúmenes utilizados para la síntesis de ADNc con enzima M-MLV RT

Reactivo	µL
5X-First Strand Buffer	4
dNTP Mix 10mM	1
0.1M DTT Mix	1
M-MLV RT (200 U/µL)	1
Random Primers (oligonucleótidos) (250ng/uL)	1
Agua libre de nucleasas	a 12
Plantilla de RNA	5-10 µL

Para la preparación del stock de oligonucleótidos se utilizó la ecuación 1 para lograr una concentración deseada dentro de un rango de 50-250ng/ µL. La memoria de cálculo se encuentra en el Anexo 2.

$$C1V1 = C2V2$$

[1]

En donde:

C1= Concentración inicial de la solución madre (stock) [ $\mu\text{g}/\mu\text{L}$ ] \*Convertir a ng/  $\mu\text{L}$

V1 = Volumen que no se conoce [ $\mu\text{L}$ ]

C2 = Concentración deseada (50-250 ng/  $\mu\text{L}$ )

V2 = Volumen de alícuota [ $\mu\text{L}$ ]

### Oligonucleótidos específicos

La síntesis de cDNA que se realizó a partir de oligonucleótidos específicos, se utilizó una concentración de 1pmol/ $\mu\text{L}$ , diluyendo 1:10 el primer específico en concentración 10  $\mu\text{M}$ . La enzima utilizada fue RT SuperScript III. Se añadieron los oligonucleótidos específicos, dNTP 10  $\mu\text{M}$  y la plantilla de ARN. Se llevó a 65°C durante 5 min e inmediatamente después se enfrió en hielo durante 1 min. Posteriormente se añadieron el First Strand Buffer, DTT 0.1M, 1  $\mu\text{L}$  de RNase Out y SuperScript III RT. Se mezcló por pipeteo y se llevó a 55°C durante 60min y, finalmente, a 70°C por 15 min. Las cantidades de cada reactivo utilizadas se muestran a continuación:

**Tabla 9** Volúmenes utilizados para la síntesis de ADNc con enzima SuperScript III RT

Reactivo	$\mu\text{L}$
5X-First Strand Buffer	4
dNTP Mix 10mM	1
0.1M DTT Mix	1
SuperScript III RT (200 U/ $\mu\text{L}$ )	1
Gene Specific Primer (2pmol/L)	2
Agua libre de nucleasas	a 12
Plantilla de RNA	5-10 $\mu\text{L}$

#### 6.2.4. Diseño de oligonucleótidos

En el diseño de oligonucleótidos se utilizó la herramienta en línea *Primal Scheme*. Primero se obtuvieron los genomas completos de referencia de la página web del NCBI (*National Center of Biotechnology*) del virus de la hepatitis C genotipos 1a y 1b. Una vez ingresando a la página web *Primal Scheme*, se siguieron las indicaciones para su diseño las cuales constan de tres pasos principales: 1) Se realizó un alineamiento de las secuencias de interés por medio del servidor web Clustal Omega o MAFFT; 2) Se revisó la matriz de homología; 3) Se creó un archivo .fasta del alineamiento. Para el primer paso, las secuencias utilizadas para los genotipos 1a y 1b se muestran en la Tabla 9. El tamaño de amplicón seleccionado fue de 400pb.

**Tabla 10** Genomas del VHC genotipos 1a y 1b utilizados para el diseño de oligonucleótidos

Genotipo	Número de acceso NCBI	
1a	NC_004102.1	EU781820.1
	EU781824.1	EU781816.1
	EU781823.1	EU781809.1
	EU781822.1	EU781811.1
	EU781821.1	
1b	B828701.1	EU781828.1
	LC011931.1	EU781831.1
	LC011928.1	EU781830.1
	LC011927.1	EU781829.1

Una vez obtenidas las secuencias de los oligonucleótidos, se realizó una prueba de PCR *in silico* mediante el software para Mac OSX Amplify4©. Para realizar las pruebas de PCR se tomó un genoma completo de cada genotipo al azar y las modificaciones de las secuencias de los oligonucleótidos se realizaron si el

oligonucleótido presentaba una homodimerización. En caso de que no se observara la banda deseada, se vuelve a rediseñar el oligonucleótido hasta que se observará en el programa.

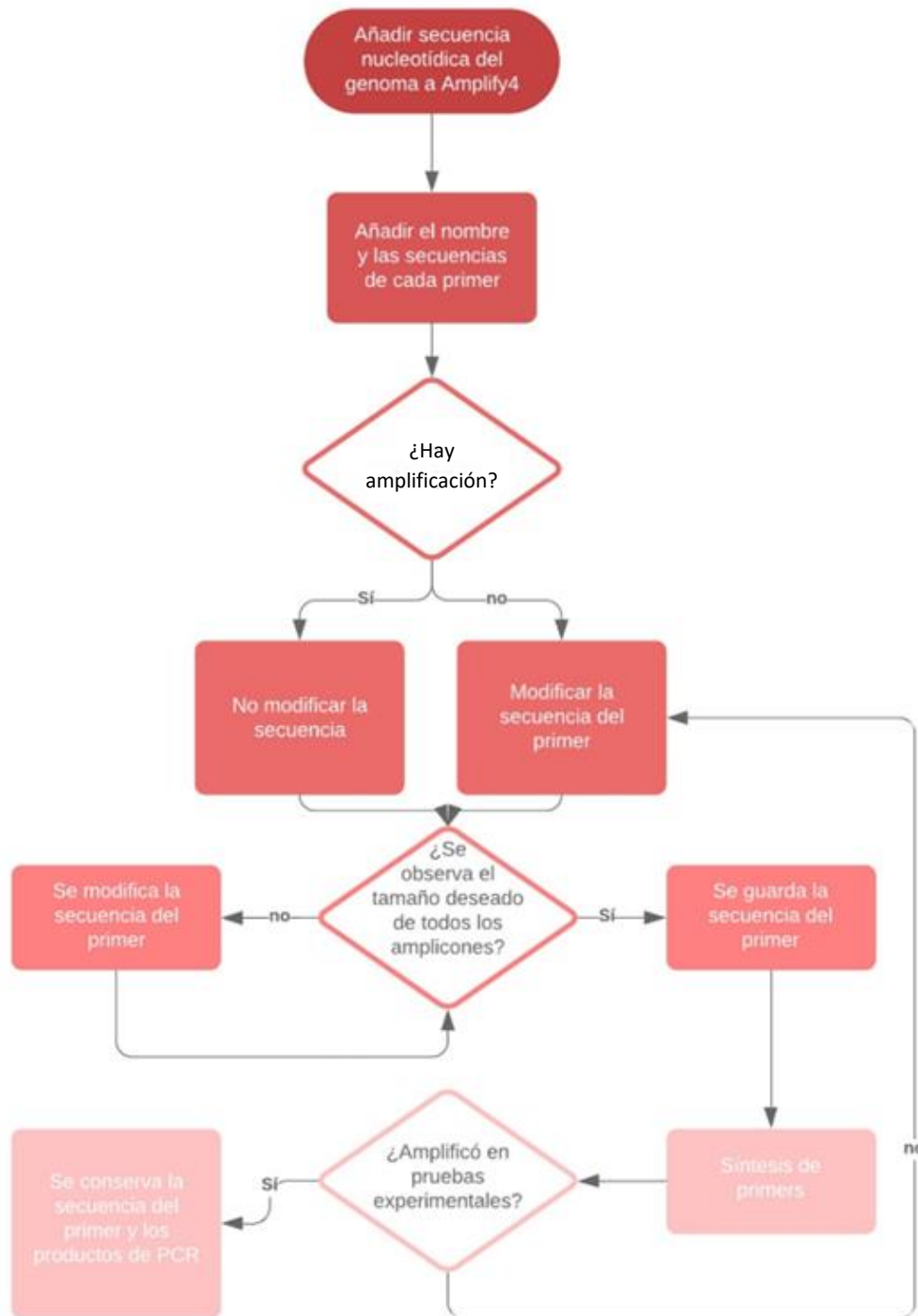
Para confirmar que el diseño de oligonucleótidos funciona, se tomaron genomas aleatorios diferentes a los usados para su diseño. El flujo de trabajo para este diseño *in silico* y en el trabajo experimental se observa en la Figura 9.

Para el rediseño de los oligonucleótidos se tomaron secuencias de genomas completos del VHC genotipo 1b de Estados Unidos de la base de datos Los Alamos y se realizó un alineamiento para encontrar las posiciones conservadas de las regiones variables del genoma en donde los oligonucleótidos fallaron. Se eliminaron secuencias que no mostraban una cobertura completa. Posteriormente se realiza un alineamiento del oligonucleótido previamente sintetizado y se modificó la secuencia sustituyendo las bases por bases degeneradas que la empresa IDT Oligo Analyzer emplea. Finalmente, la comprobación *in silico* se realiza nuevamente mediante el software Amplify 4 con una base de datos diferente a la usada para el rediseño.

La síntesis de los oligonucleótidos se llevó a cabo mediante la empresa Oligo T4®. Los certificados de los oligonucleótidos sintetizados se encuentran en el Anexo 4. Para ambos genotipos se obtuvieron 32 pares de oligonucleótidos.

Los oligonucleótidos sintetizados se resuspendieron en agua libre de nucleasas para obtener una concentración molar de 100µM y posteriormente se realizó una dilución 1:10 en un volumen total de 100µL. Esta dilución final (concentración molar de 10µM) fue la que se utilizó para llevar a cabo las reacciones de PCR.





**Figura 7** Diagrama de flujo para el diseño de oligonucleótidos *in silico* y experimental. Si los oligonucleótidos fallan las pruebas experimentales se rediseña el oligonucleótido, se sintetiza nuevamente, se realizan pruebas *in silico* y finalmente se lleva a cabo la síntesis del oligonucleótido y su prueba experimental.

### 6.2.5. Reacción en cadena de polimerasa (PCR) punto final

La PCR se llevó a cabo de acuerdo siguiendo las instrucciones del protocolo de uso de la DNA Taq Polimerasa de Thermo Scientific. Brevemente, se mezclaron las cantidades de todos los reactivos de la Tabla 11 en un microtubo, éstos se mencionan igualmente en el anexo II junto con sus cantidades. Antes de comenzar la reacción, se descongelan y se usa brevemente el vórtex para mezclar cada reactivo. Todos los pasos se llevan a cabo en frío en un plato de enfriamiento para PCR.

En un microtubo de PCR se adicionaron los siguientes volúmenes:

**Tabla 11** Volúmenes utilizados para la PCR punto final

Reactivo	$\mu\text{L}$
Taq Buffer 10X	2.5
dNTP Mix 10 mM	0.5
MgCl <sub>2</sub> 50 $\mu\text{M}$	0.75
Oligonucleótido Forward 10 $\mu\text{M}$	1.25
Oligonucleótido Reverse 10 $\mu\text{M}$	1.25
Taq DNA Polimerasa (200 U/ $\mu\text{L}$ )	0.4
Agua libre de nucleasas	16.35
Plantilla de ADNc	2

Una vez mezclados los componentes se llevó a cabo la PCR con las condiciones de temperatura de cada ciclo mencionadas en el Anexo II.

Se realizó un procedimiento de estandarización de la PCR, por medio de una evaluación de condiciones óptimas de amplificación. Dicha evaluación consistió en realizar gradientes de temperaturas de 64°C – 70°C. Las condiciones de prueba para estos gradientes se resumen en la Tabla 12. Los productos de PCR fueron observados mediante electroforesis en gel de agarosa al 1% con Buffer TBE al 1x.

**Tabla 12** Condiciones de amplificación con gradiente de temperatura para oligonucleótidos 1a y 1b

Genotipo	Oligonucleótido	Temperaturas (°C)				
1b	<sup>1</sup> Pares nones	64		66	68	70
	<sup>2</sup> Pares pares	64		66	68	70
1a	Pares nones	*62	64	66	68	70
	Pares pares	64		66	68	70 72*

\*Solo se probó esta temperatura para un par de oligonucleótidos

<sup>1</sup>Oligonucleótidos etiquetados con números nones

<sup>2</sup>Oligonucleótidos etiquetados con números pares

### 6.2.6. Preparación de muestras para secuenciación

Para la secuenciación se realizó una PCR Multiplex, combinando los oligonucleótidos nones en un microtubo y los oligonucleótidos pares en otro microtubo (Tabla 13). Ambas reacciones fueron probadas a dos temperaturas: 64°C y 66°C. Se realizaron dos mezclas, una de oligonucleótidos nones y otra de oligonucleótidos pares, añadiendo a cada tubo 5µL de los oligonucleótidos diluidos (1:100). Posteriormente se llevó a cabo la preparación de la mezcla para la PCR Multiplex añadiendo en 2 microtubos los siguientes volúmenes:

**Tabla 13** Volúmenes para la preparación de PCR Multiplex

Reactivo	[µL]
5X Q5 Reaction Buffer	5
10mM dNTPs	0.5
Q5 Hot Start DNA Polymerase (200 U/µL)	0.25
Oligonucleótidos Nones o Pares (10µM)	1.3
Nuclease-free water	15.45
<b>Total</b>	<b>22.5</b>

Los oligonucleótidos utilizados para las PCR Múltiplex se muestran en la Tabla 14.

**Tabla 14** Oligonucleótidos utilizados en la PCR Múltiplex

Nones		Pares	
Nombre	Secuencia	Nombre	Secuencia
1b_1_LEFT	TAGTATGAGTGTCGTGCAGCCT	1b_2_LEFT	TCGTTGGTGGAGTTTACCTGTT
1b_1_RIGHT	ACCGCTCGGAAGTCTTCCTA	1b_2_RIGHT	TAATTCACGCCGTCTCCAGAA
1b_3_LEFT	GGTCATCGATACCCTCACATGC	1b_4_LEFT	TCACGAACGACTGCTCCAAC
1b_3_RIGHT	CGTCGTATTGTCGTAGTGGGGA	1b_4_RIGHT	GATCCGGAGTAACTGCGACA
1b_5_LEFT	TATCCATGGTGGGGAAGTGGG	1b_6_LEFT	CACATCAACAGAAGTGCCTGA
1b_5_RIGHT	AGGTGCGTAGTGCCAGCAATA	1b_6_RIGHT	TTCATCCATGTACAGCCGAACC
1b_7_LEFT	GACGTGCTGCTCCTCAACAA	1b_8_LEFT	ACTACCCCTGCACTGTCAACTT
1b_7_RIGHT	GGCTGAGCTCTGATCTATCCCT	1b_8_RIGHT	ACCACCAAGTTCTCTAAGGCGG
1b_9_LEFT	GCAATACCTGTACGGCATAGGG	1b_10_LEFT	TTATGCCATGGACCGGGAGAT
1b_9_RIGHT	AGTATGGTGCCAAGGTCAAGAAT	1b_10_RIGHT	AGCCTCATGAGAGCCATTTGGAC
1b_11_LEFT	CTGGGGGCCATTATGTCCAAAT	1b_12_LEFT	AGTAGAGCCCGTCGTCTTCT
1b_11_RIGHT	GAAACCACTTGAACCTCCCCCT	1b_12_RIGHT	AGGTCCTGGTCTACATTGGTGT
1b_13_LEFT	TTAACGGCGTGTGTTGGACTG	1b_14_LEFT	TGTGGCAGCTCGGACCTTTA
1b_13_RIGHT	TGGTTTCCATAGACTCAACGGG	1b_14_RIGHT	TCAGGACGAGTACCTTGTACCC
1b_15_LEFT	TTCCAAGTGGCCCATCTACAC	1b_16_LEFT	ATCATAATATGTGATGAGTGCCACTC
1b_15_RIGHT	ATGTTGGGATGTGGCACGGT	1b_16_RIGHT	CCGGTAAAGCCCGTCATTAGAG
1b_17_LEFT	TAGCGTATTACCGGGGTCTTGA	1b_18_LEFT	TGTGTGAGTGCTATGACGCG
1b_17_RIGHT	GCAACCCTGGTGTGTTTAGGTA	1b_18_RIGHT	CAGCCGACATACATGCCATGAT
1b_19_LEFT	ACACCCCTGCTGTATAGGCTAG	1b_20_LEFT	TACATCGAACAGGGAATGCAGC
1b_19_RIGHT	AAGGGTTCGCCACTTGGACT	1b_20_RIGHT	ATGTCCACAAGCACCTTCCCA
1b_21_LEFT	ACCCAAAACACCTCCTGTTTA	1b_22_LEFT	CAGTGGATGAACCGGCTGATAG
1b_21_RIGHT	CGTCGCTTTCAGGCACATAGT	1b_22_RIGHT	GGAACCGTTTTTGACATGTCCG
1b_23_LEFT	TCATGTCAACGTGGGTACAAGG	1b_24_LEFT	TAAATGCCCATGCCAGGTTCC
1b_23_RIGHT	TCCCGTAGAAGAGGCTTGCA	1b_24_RIGHT	TCTCTGACTCCACTCGGGTGAT
1b_25_LEFT	TCTTCAGCCAGCCAGTTGTCT	1b_26_LEFT	CAGAGATCCTGCGGAAATCCAG
1b_25_RIGHT	TCTTTCTCCGTGGAGGTGGT	1b_26_RIGHT	GACCAAGACCCGTCGCTAAGAT
1b_27_LEFT	CGATCTTAGCGACGGGTCTT	1b_28_LEFT	CCACAGTTAAGGCCAAGCTTCT
1b_27_RIGHT	GTCCTTTGCCCCATAGCCAAA	1b_28_RIGHT	CTTTGATTTCCAGGCATTACCA
1b_29_LEFT	AAGCCAGCTCGCCTTATCGT	1b_30_LEFT	AGACAGGCCATAAGGTGCGCT
1b_29_RIGHT	GGTATTACCGCAGCTAGTCGTC	1b_30_RIGHT	AGTACACCCTTTTGCCAGATGC
1b_31_LEFT	CCGGAATACGACTTGGAGTTGA	1b_32_LEFT	CCATTGAGCCACTTGACCTACC
1b_31_RIGHT	AGCCACCCTATTGATCTCACCT	1b_32_RIGHT	AGATGCCTACCCCTACAGAAAGT

## 7. Secuencia utilizada para el análisis bioinformático



### Figura 8 Flujo de trabajo para el tratamiento de datos crudos de secuenciación.

Los archivos crudos se analizan con Genome Detective, la secuencia es ensamblada mediante el software BWA, los alineamientos se realizan mediante el servidor MAFFT en su versión 7 y visualizados en MEGA X. El árbol filogenético de máxima similitud se realiza mediante IQ-Tree y es visualizado en FigTree. El análisis de TempEst requiere de los archivos generados por IQ-Tree y finalmente se crean archivos .beauti en el software beauti con parámetros de reloj molecular estricto, modelo de sustitución G4, población constante coalescente y el largo de la cadena de Monte Carlo de  $40 \times 10^6$  generaciones. Éste archivo es analizado por BEAST 2.0 y posteriormente se utilizan los archivos .trees generados para ser comprimidos en el software Tree Annotador, el árbol final se visualiza nuevamente en FigTree.

### 7.1.1. Secuenciación y ensamblaje

La preparación de la biblioteca se realizó con el Kit de Ligación SQK-LSK109 y el barcoding se realizó con el kit Native Barcoding Kit para posteriormente cargar la celda de flujo del MinION FLO-MIN106 de acuerdo con las instrucciones del fabricante. El tiempo de corrida fue de 20hrs Se utilizó el software MinkNOW en su versión 22.03.5 y Guppy 6.0.6 para el *basecalling*. La secuencia de referencia utilizada en el software fue EU256084.1. Se visualizó la cobertura del genoma con el servidor web de Genome Detective de los archivos crudos de secuenciación y se obtuvieron las secuencias de los contigs en formato .fasta. Para el ensamblaje se utilizó el software Burrows-Wheeler Alignment tool (BWA) con parámetros -x ont (Li & Durbin, 2009), el código se obtuvo mediante Github y el script para la limpieza de secuencias se utilizó Porechop (Anexo XI).

### 7.1.2. Análisis filogenético de máxima verosimilitud

Se realizó el control de calidad de las secuencias genómicas. A partir de la base de datos Los Alamos (Kuiken C. et al., 2015) se descargaron las secuencias parciales de los genes NS3 (519pb), NS3/4(111pb) y NS5B (308pb) del VHC de la República Mexicana, secuencias completas (~9400pb) de Estados Unidos de América y secuencias internacionales. Se construyeron 5 sets de datos: set 1, secuencias internacionales de Asia, Europa y América (n=263); set 2, secuencias de Estados Unidos de América (n=103); set 3, secuencias parciales de la República Mexicana del gen NS3 del VHC (n=6); set 4, secuencias parciales de la República Mexicana del gen NS5B del VHC (n=11); set 5, secuencias parciales de la República Mexicana de los genes NS3/4A (n=118). Se descargaron las secuencias parciales de 4 virus obtenidos y un VHC completo.

Las secuencias parciales de los VHC se alinearon con las secuencias parciales de los genes NS3, NS3/4A y NS5B. La secuencia completa del VHC obtenido se alineó con las de Estados Unidos de América y las internacionales. El alineamiento se realizó por medio del servidor web de MAFFT (Kato et al., 2019) y para cada set de datos, se realizó un filtrado de secuencias. Aquellas que mostraban estar incompletas en más de un 90% se eliminaban y se eliminaron los gaps conservados para evitar altas penalizaciones durante la construcción del árbol filogenético de máxima verosimilitud. Posteriormente los alineamientos se utilizaron para construir un árbol de máxima verosimilitud empleando el software IQ-tree (Minh B. et al., 2020) con un modelo de sustitución TPM2+F+G4 y GTR+G4 mediante el método Bootstrap Ultra Fast (Hoang D. et al., 2018) en línea de comandos con parámetros -bb 1000 para el Bootstrap, -m TEST para la selección del mejor modelo y -msub viral para indicar el tipo de secuencias de entrada. El árbol se visualiza mediante Fig tree y finalmente se realiza el análisis filogenético contemporáneo en TempEst.

Los comandos utilizados fueron:

```
1 bin\iqtree -s todas_1b_alineadas.fasta -m TEST -msub viral -bb 1000
```



### 7.1.3. Análisis evolutivo bayesiano

La reconstrucción de la filogenia por métodos Bayesianos se realizó utilizando los alineamientos múltiples previamente realizados para generar un archivo con extensión .xml, a partir del programa BEAUTi (Drummond et al., 2012). Los parámetros incluidos para generar el archivo fueron los modelos de sustitución GTR+G4, el reloj molecular estricto y la población de crecimiento coalescente constante. Se utilizaron además cadenas de 40,000,000 de estados, El archivo generado se llevó al software BEAST v1.10.4 (Drummond et al., 2012) que fue utilizado con los parámetros establecidos. Los archivos resultantes se combinaron utilizando el programa LogCombiner v1.10.4 (Drummond et al., 2012) y se determinó que hubiera alcanzado un tamaño de muestra efectivo mayor a 200 para cada archivo .xml generado de cada set de datos. Únicamente se realizaron filogenias evolutivas de los análisis TempEst cuyas relaciones mostraron una pendiente positiva bajo el método de la media residual heurística al cuadrado los cuales fueron NS3 (Figura 32(A)) y NS5B(Figura 32(D)).

Finalmente, se generó un consenso de los árboles predichos utilizando el software TreeAnnotator 1.8.4 descartando el 10% de los árboles. El consenso de árboles se visualizó en FigTree v1.4.3. Se determinó el ancestro común más reciente.

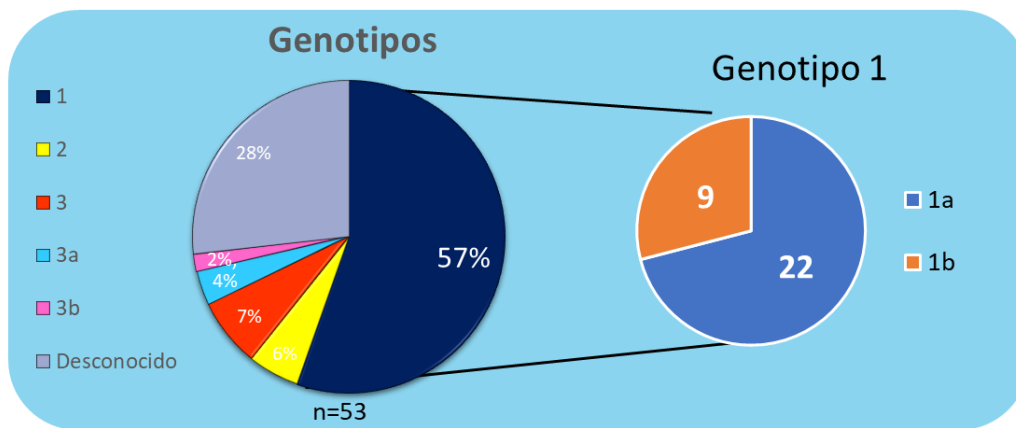
#### **7.1.4. Identificación de sustituciones asociadas a resistencia**

Las RAS se identificaron mediante el software en línea HCV-Glue (<http://hcv-glue.cvr.gla.ac.uk/>). Para ello se descargaron las secuencias .fasta obtenidas previamente con Genome Detective de cada VHC secuenciado. La identificación de RAS en la secuencia completa del VHC obtenido se llevó a cabo con la secuencia ensamblada mediante el software BWA.

## 8. Resultados

### 8.1. Muestras de pacientes

Se recolectaron 53 muestras de sujetos positivos para infección con el VHC. Del total de sueros, 31 fueron identificadas como pertenecientes al GT 1. Además, 9 fueron GT 1a y 22 GT 1b. Las muestras de suero fueron de pacientes que residen en Monterrey, Nuevo León. Los sujetos presentaban un cuadro clínico crónico y al momento de la toma de muestra no habían recibido ningún tratamiento. Cinco muestras cumplieron con los criterios de inclusión y fueron utilizadas para la secuenciación (Figura 9).



**Figura 9** Muestras de pacientes infectados con VHC. De un total de 53 muestras, el 57% perteneció al GT 1 del cual, 9 muestras eran GT 1b y 22 GT 1a

**Tabla 15** Muestras seleccionadas de pacientes infectados con VHC

Desconocido			GT 1a			GT 1b		
ID Muestra	Carga [UI/ml]	<sup>1</sup> Año	ID Muestra	Carga [UI/ml]	<sup>1</sup> Año	ID Muestra	Carga [UI/ml]	<sup>1</sup> Año
180	76x10 <sup>6</sup>	2016	157	8x10 <sup>6</sup>	2016	336	1,156,919	2017
RAPL	NA	2016	364	1.6x10 <sup>6</sup>	2018	182	142,205	2016
			213	2.4x10 <sup>6</sup>	2016	429	2,134,082	2020
						261	685,225	2017

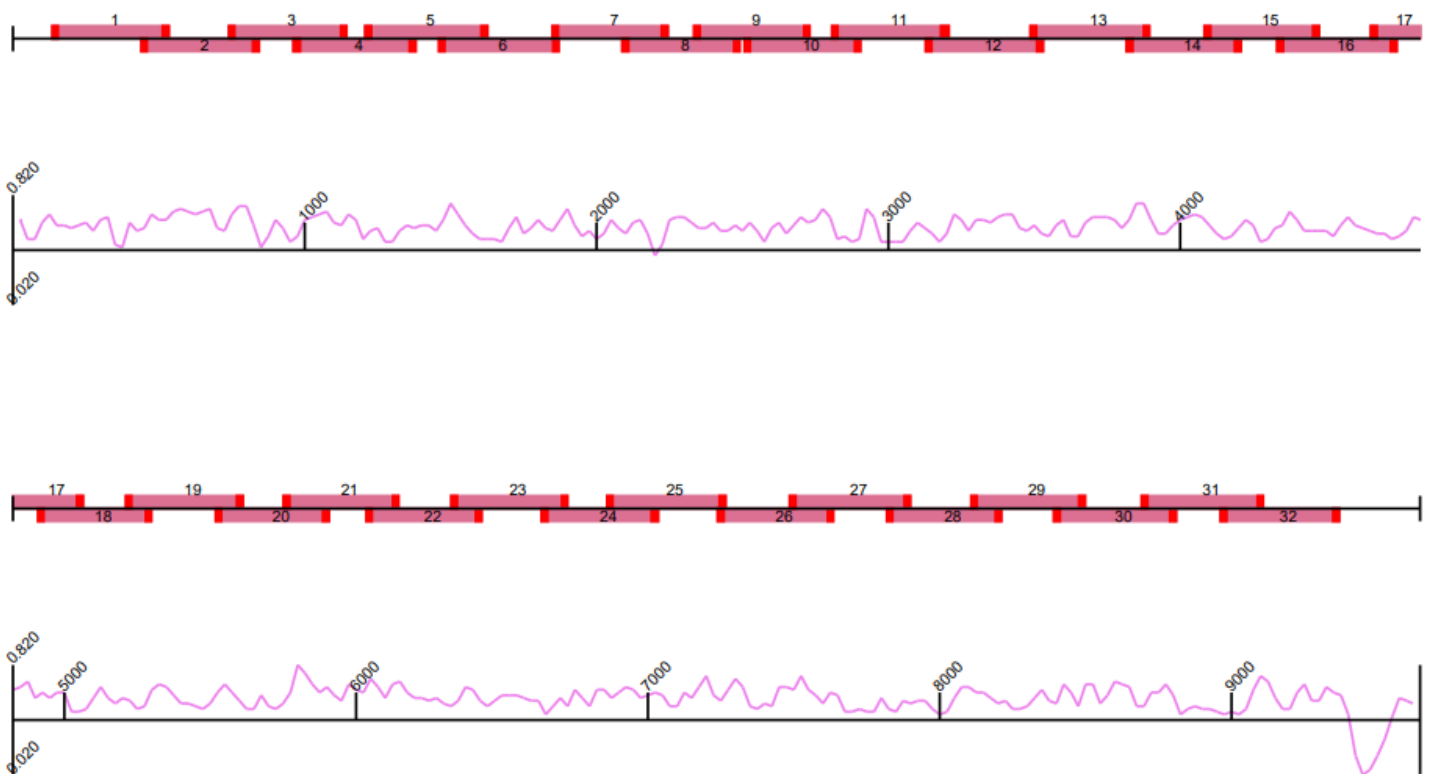
<sup>1</sup> Año de muestreo

## 8.2. Diseño *in silico*

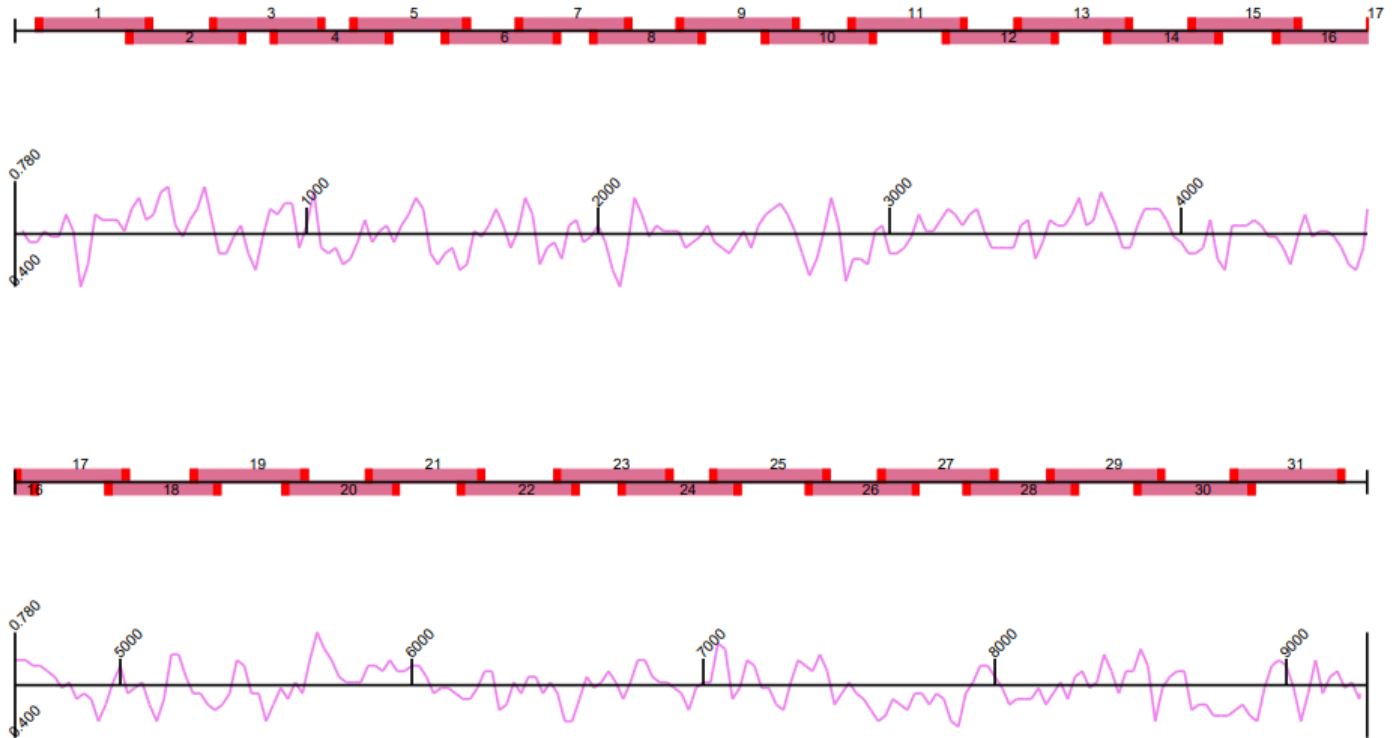
El diseño *in silico* de oligonucleótidos con Primal Scheme se realizó dos veces para los oligonucleótidos 1a, uno fue el oligonucleótido diseño y el segundo fue el rediseño. No se logró la generación de un segundo set oligonucleótidos *in silico* para el GT 1b debido a las limitantes del software Primal Scheme.

### 8.2.1. Síntesis *in silico* de oligonucleótidos 1a obtenidos a partir de Primal Scheme

A)

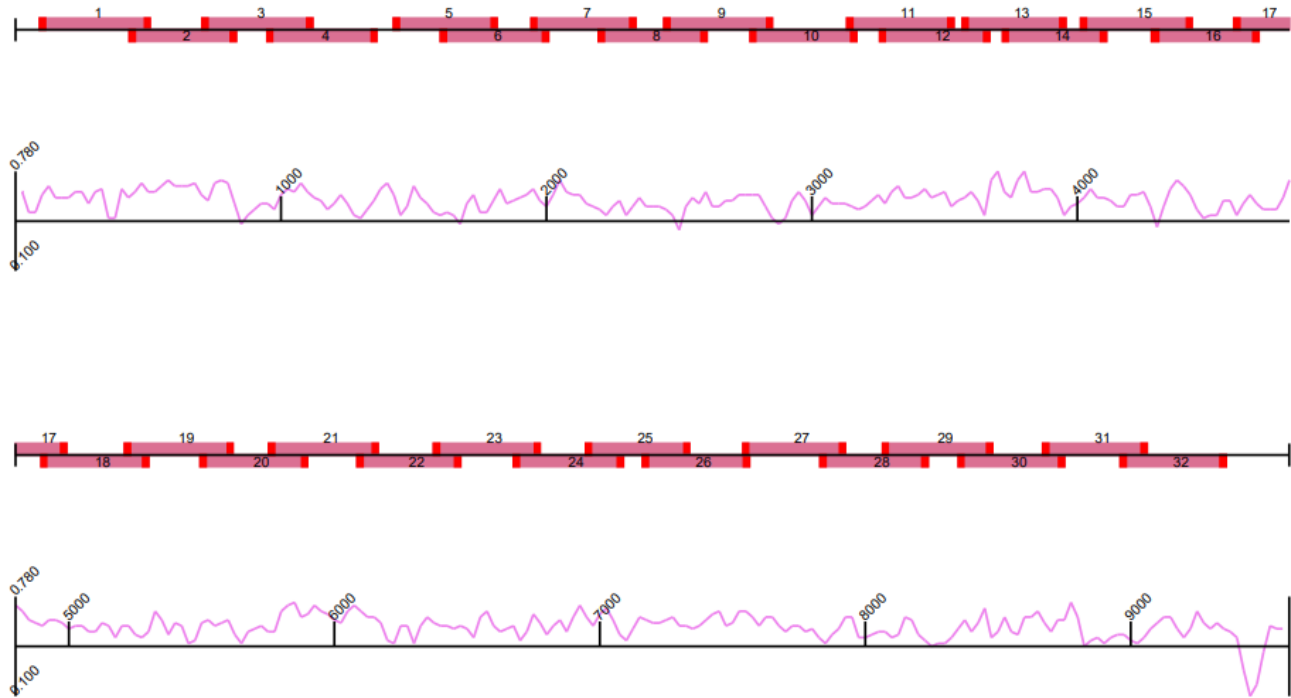


B)



**Figura 10** Diseño *in silico* de Primal Scheme para los GT 1a y 1b . A) Primer diseño de oligonucleótidos 1<sup>a</sup> utilizando 10 secuencias de NCBI. B) Segundo diseño de oligonucleótidos 1a utilizando 35 secuencias de “Los Alamos”. Los rectángulos enumerados son los oligonucleótidos solapados (mostrados en color rojo). La línea horizontal en el eje “x” es la cobertura del genoma (línea negra), las líneas verticales al eje “x” representan la posición en el genoma y se muestran en intervalos de cada 1000 nucleótidos (1-9300pb). La gráfica sobre el eje “x” (línea rosa) muestra la variabilidad genética que tiene el set de datos.

## 8.2.2. Síntesis in silico de oligonucleótidos 1b obtenidos a partir de Primal Scheme



**Figura 11 Primer diseño de oligonucleótidos 1b con 10 secuencias de NCBI.** Los rectángulos enumerados son los oligonucleótidos solapados y la gráfica en el eje “x” es la cobertura del genoma. Los rectángulos enumerados son los oligonucleótidos solapados (mostrados en color rojo). La línea horizontal en el eje “x” es la cobertura del genoma (línea negra), las líneas verticales al eje “x” representan la posición en el genoma y se muestran en intervalos de cada 1000 nucleótidos (1-9300pb). La gráfica sobre el eje “x” (línea rosa) muestra la variabilidad genética que tiene el set de datos.

### **8.2.3. Primer diseño de oligonucleótidos**

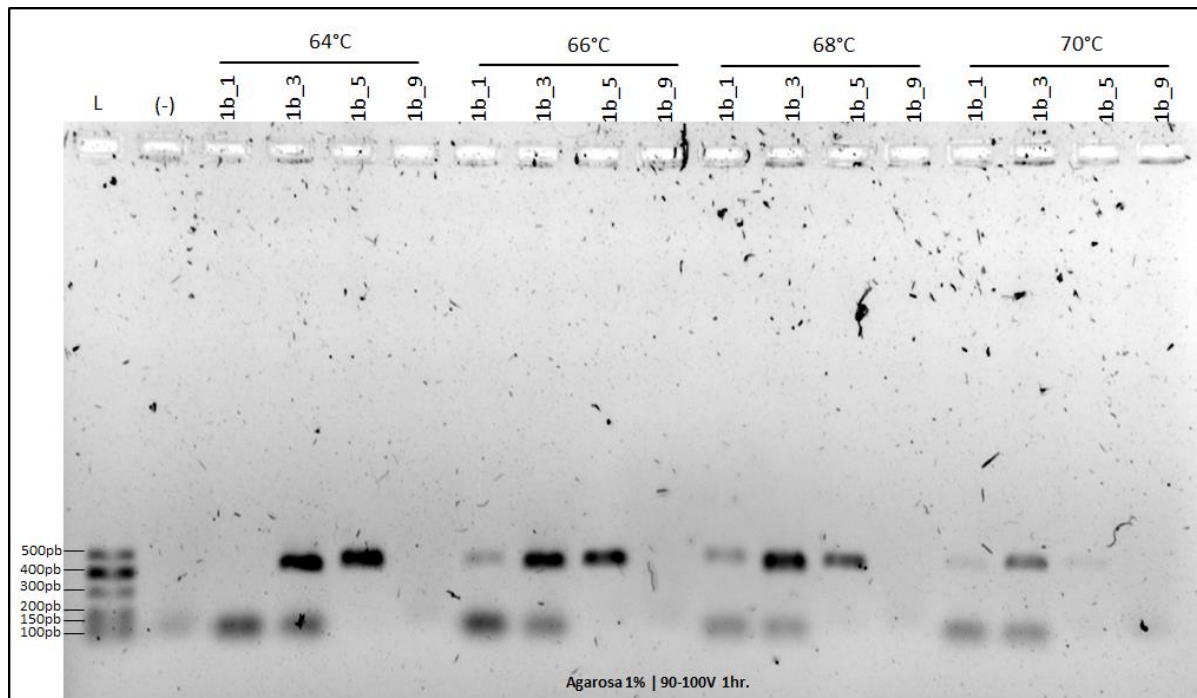
#### **8.2.3.1. Estandarización de PCR punto final mediante gradiente de temperatura y amplificación del genotipo 1b**

Se utilizó una de la muestra de suero ID 429 con una carga viral de 2,134,082UI/ml del genotipo 1b para llevar a cabo la estandarización de las etapas de la secuenciación ya que la carga viral era alta.

Las temperaturas de alineamiento a las cuales se obtuvieron la mayoría de los productos amplificados fueron a 64°C y 66°C (Figura 12). Los oligonucleótidos que amplificaron en todas las temperaturas fueron: 1b\_3, 1b\_19, 1b\_29. Los oligonucleótidos de que lograron amplificar un producto muy visible fueron: 1b\_3, 1b\_5, 1b\_13, 1b\_19, 1b\_25, 1b\_31, 1b\_12, 1b\_14 y 1b\_18. En la Figura 14 se muestra un ejemplo de uno de los geles con bandas tenues y bandas más concentradas. Las temperaturas experimentales y óptimas para cada par de oligonucleótidos se muestran en la Figura 15. Más del 50% de los oligonucleótidos amplificaron fragmentos de 400pb a las temperaturas de 64°C y 66°C. Estas temperaturas se utilizaron como condiciones térmicas experimentales para



estandarizar la PCR Múltiplex. La mejor amplificación de fragmentos se logró a una temperatura de 64°C (Figura 13).



**Figura 12** Amplificación de los fragmentos de 400pb GT 1b. Geles de electroforesis. Agarosa 1.5-2%. 90V 5 min 100V 30min Marcador Biorange 25-500pb. Cada par de oligonucleótidos fue amplificado en un gradiente de temperatura de 64°C-70°C.

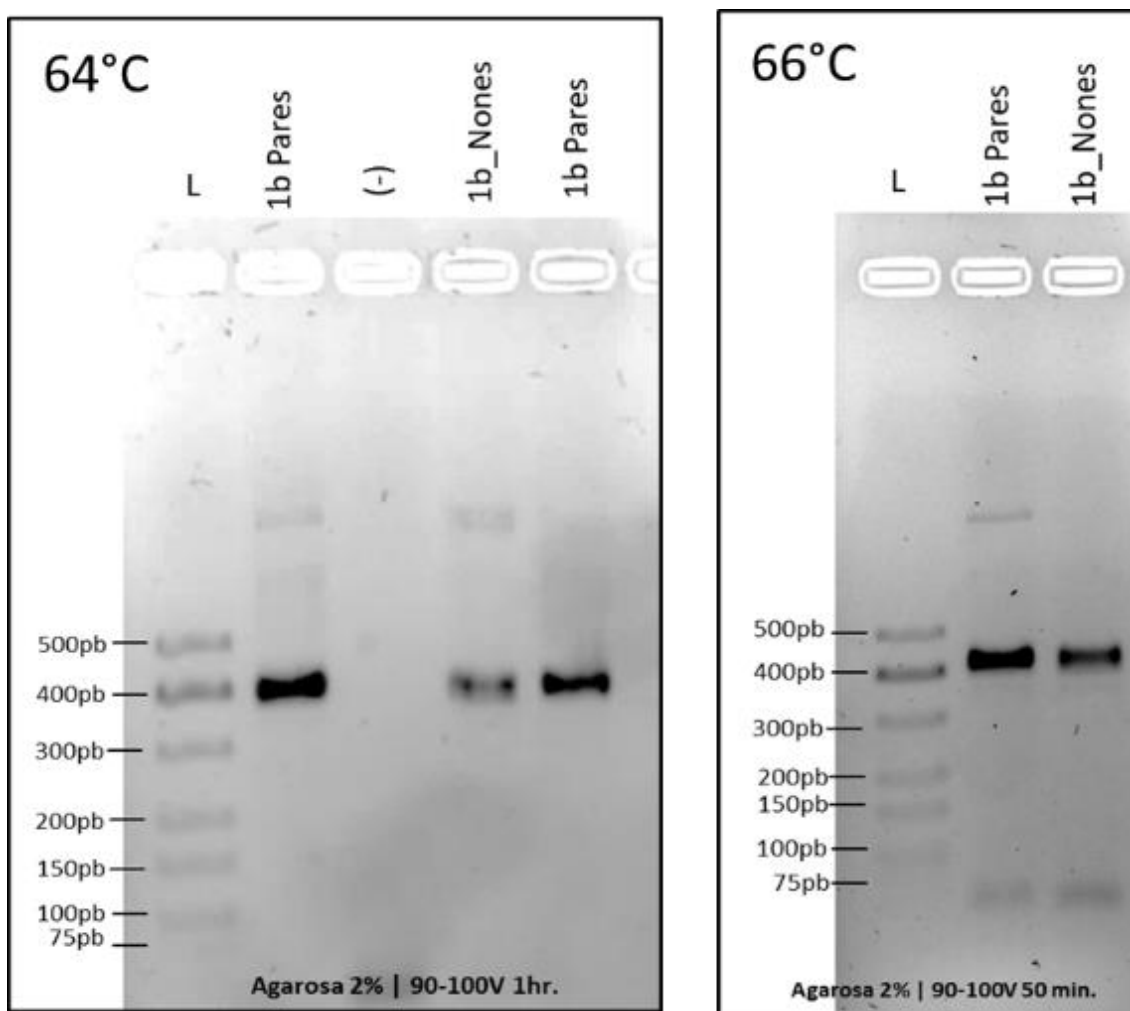
Primer	Tm	Tm seleccionada	Temperatura (°C)			
1b_1_LEFT	64.1	66	64	66	68	70
1b_1_RIGHT	65.4					
1b_2_LEFT	64.8		64	66	68	70
1b_2_RIGHT	69.8					
1b_3_LEFT	67.9	66	64	66	68	70
1b_3_RIGHT	68.2					
1b_4_LEFT	66.9	64	64	66	68	70
1b_4_RIGHT	65.6					
1b_5_LEFT	70.8	66	64	66	68	70
1b_5_RIGHT	67					
1b_6_LEFT	67.7		64	66	68	70
1b_6_RIGHT	68.2					
1b_7_LEFT	67		64	66	68	70
1b_7_RIGHT	65.2					
1b_8_LEFT	64.1	68	64	66	68	70
1b_8_RIGHT	67.6					
1b_9_LEFT	66.5		64	66	68	70
1b_9_RIGHT	64.9					
1b_10_LEFT	69.7		64	66	68	70
1b_10_RIGHT	69.8					
1b_11_LEFT	69.6	70	64	66	68	70
1b_11_RIGHT	68.5					
1b_12_LEFT	62.4	64	64	66	68	70
1b_12_RIGHT	63.9					
1b_13_LEFT	68.3	66	64	66	68	70
1b_13_RIGHT	67.2					
1b_14_LEFT	68.3	66	64	66	68	70
1b_14_RIGHT	64.5					
1b_15_LEFT	66.3		64	66	68	70
1b_15_RIGHT	70					
1b_16_LEFT	64.2		64	66	68	70
1b_16_RIGHT	67.7					

Primer	Tm	Tm seleccionada	Temperatura (°C)			
1b_17_LEFT	66.5	66	64	66	68	70
1b_17_RIGHT	65.1					
1b_18_LEFT	66.7	64	64	66	68	70
1b_18_RIGHT	69.2					
1b_19_LEFT	63.6	66	64	66	68	70
1b_19_RIGHT	67.3					
1b_20_LEFT	68.4	64	64	66	68	70
1b_20_RIGHT	69.2					
1b_21_LEFT	65.6	64	64	66	68	70
1b_21_RIGHT	66.7					
1b_22_LEFT	68		64	66	68	70
1b_22_RIGHT	69.5					
1b_23_LEFT	67.2	64	64	66	68	70
1b_23_RIGHT	67.4					
1b_24_LEFT	70.7		64	66	68	70
1b_24_RIGHT	67.7					
1b_25_LEFT	66.7	64	64	66	68	70
1b_25_RIGHT	66.3					
1b_26_LEFT	69.3		64	66	68	70
1b_26_RIGHT	67.7					
1b_27_LEFT	65.7	66	64	66	68	70
1b_27_RIGHT	69.1					
1b_28_LEFT	65.7		64	66	68	70
1b_28_RIGHT	69.1					
1b_29_LEFT	67.4	66	64	66	68	70
1b_29_RIGHT	64					
1b_30_LEFT	65.3		64	66	68	70
1b_30_RIGHT	66.5					
1b_31_LEFT	67.3		64	66	68	70
1b_31_RIGHT	65.2					
1b_32_LEFT	67		64	66	68	70
1b_32_RIGHT	63.3					

**Figura 13** Temperaturas óptimas de amplificación de los oligonucleótidos GT 1b. La columna “Tm” es la temperatura señalada por el fabricante. La columna “Tm seleccionada” muestra la temperatura óptima de amplificación de acuerdo a lo observado en el gel de electroforesis. La columna “Temperatura (°C)” es el gradiente de temperatura utilizado para cada par. Los oligonucleótidos con recuadros negros no mostraron amplificación en ninguna de esas temperaturas.

### 8.2.3.2. Estandarización de PCR Múltiplex

Para realizar la PCR Múltiplex, se realizaron dos amplificaciones, una a 64°C y otra a 66°C, cada una con oligonucleótidos pares y oligonucleótidos impares (Figura 16). Ambas temperaturas mostraron la amplificación de fragmentos de 400pb con pocas inespecificidades, Sin embargo se seleccionó la temperatura de 64°C para las próximas PCR Multiplex debido a que ésta fue óptima para la mayoría de los oligonucleótidos.

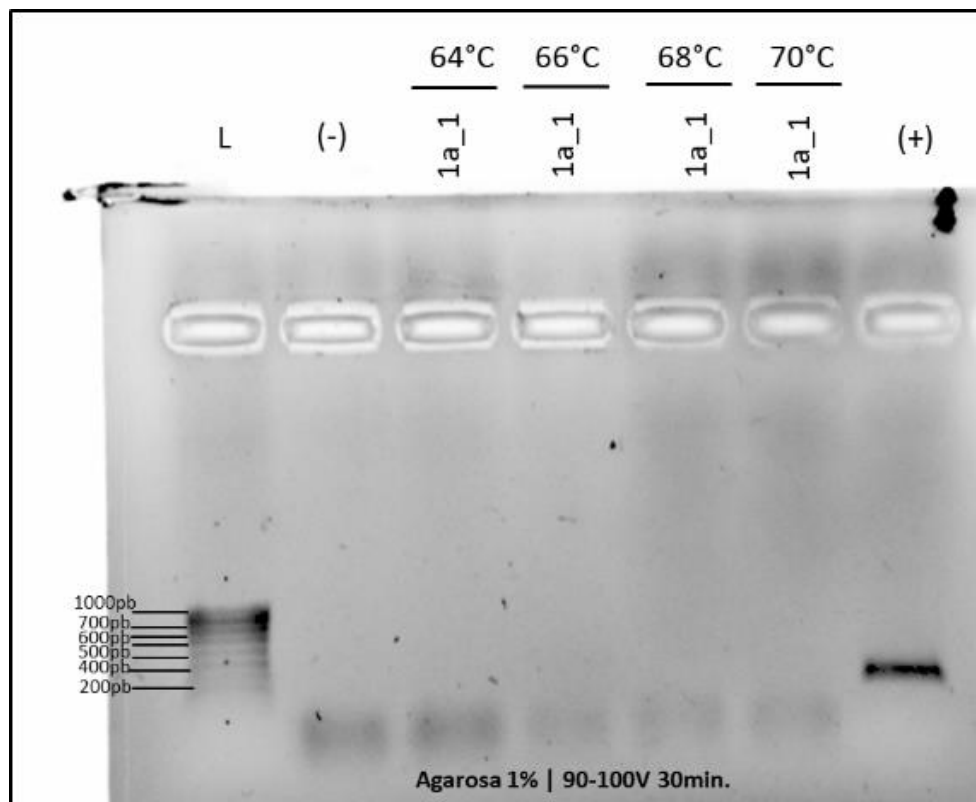


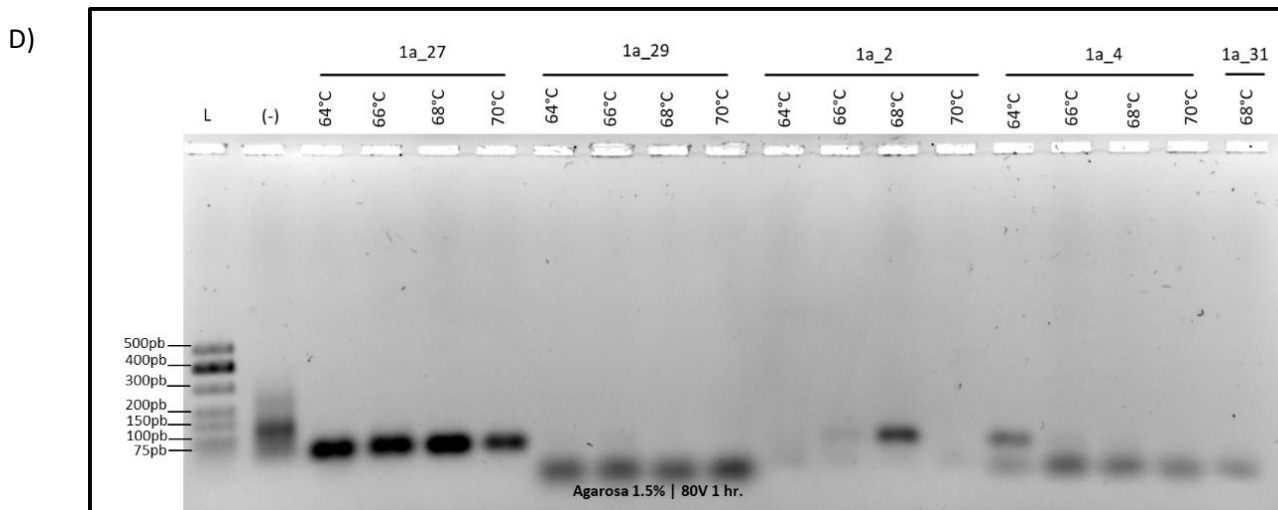
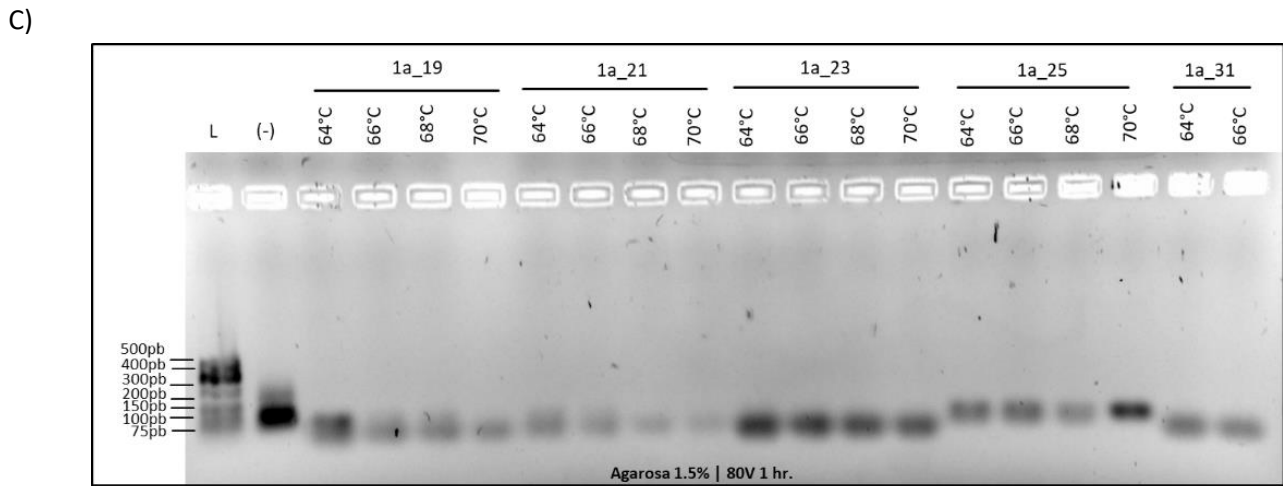
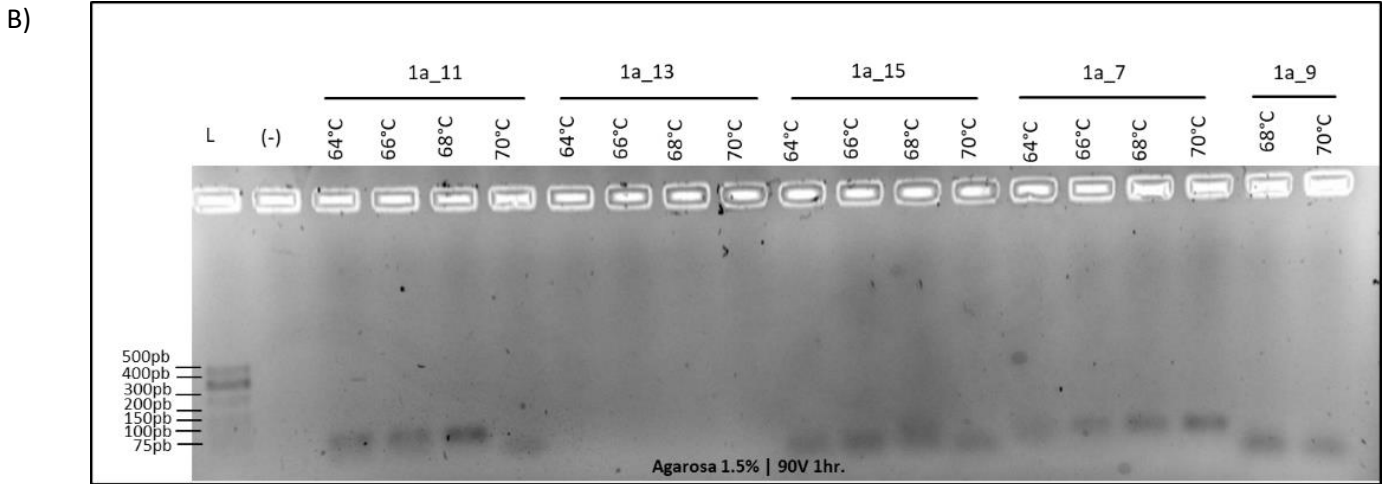
**Figura 14** Primera PCR Múltiplex GT 1b. Gel de agarosa al 2%. 90V durante 5min y posteriormente 100v 50 min, marcador Biorange 25-500pb. Las temperaturas de la fase de alineamiento se señalan en la esquina superior izquierda.

### 8.2.3.3. Amplificación del genotipo 1a

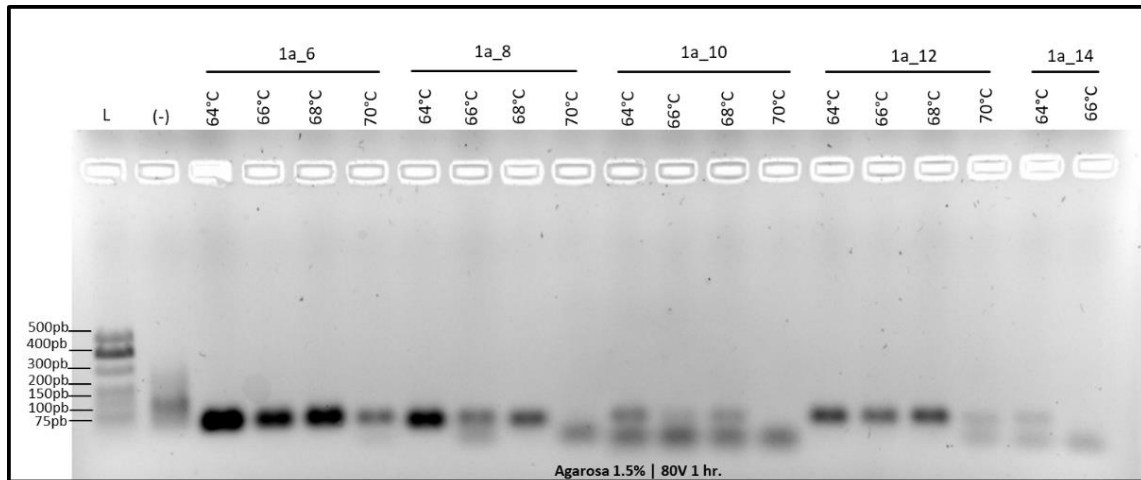
Para la amplificación de los fragmentos del genotipo 1a se utilizó la muestra de suero ID 157. No se observó amplificación de ningún fragmento de 400pb. Se utilizó en un gel un control positivo (fragmento amplificado del genotipo 1b con el oligonucleótido número 5 del primer diseño) para descartar fallos ocasionados por reactivos (Figura 15).

A)





E)

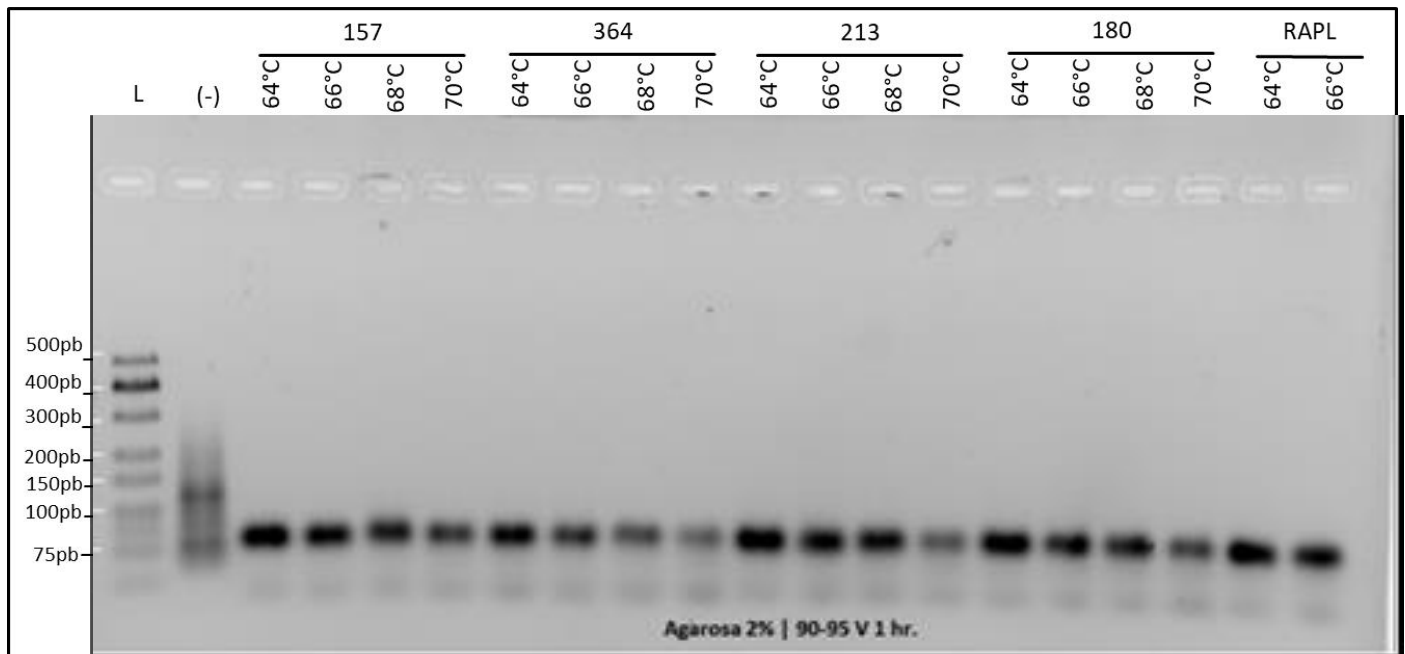


**Figura 15** Amplificación de fragmentos GT 1<sup>a</sup>. Agarosa 1-1.5%, 90V. 1hr. Marcador Biorange 25-500pb. Gel Red A) Marcador 1000pb Gradiente de temperatura 64°-70°C. B) Oligonucleótidos nones 11,13,15,7,9. C) Oligonucleótidos nones 19,21,23,25,31. D) Oligonucleótidos nones 27,29 y 31, oligonucleótidos pares, 2,4. E) Oligonucleótidos pares 6,8,10,12,14. Se observan dímeros de oligonucleótidos.

## 8.2.4. Segundo diseño de oligonucleótidos

### 8.2.4.1. Amplificación del genotipo 1a

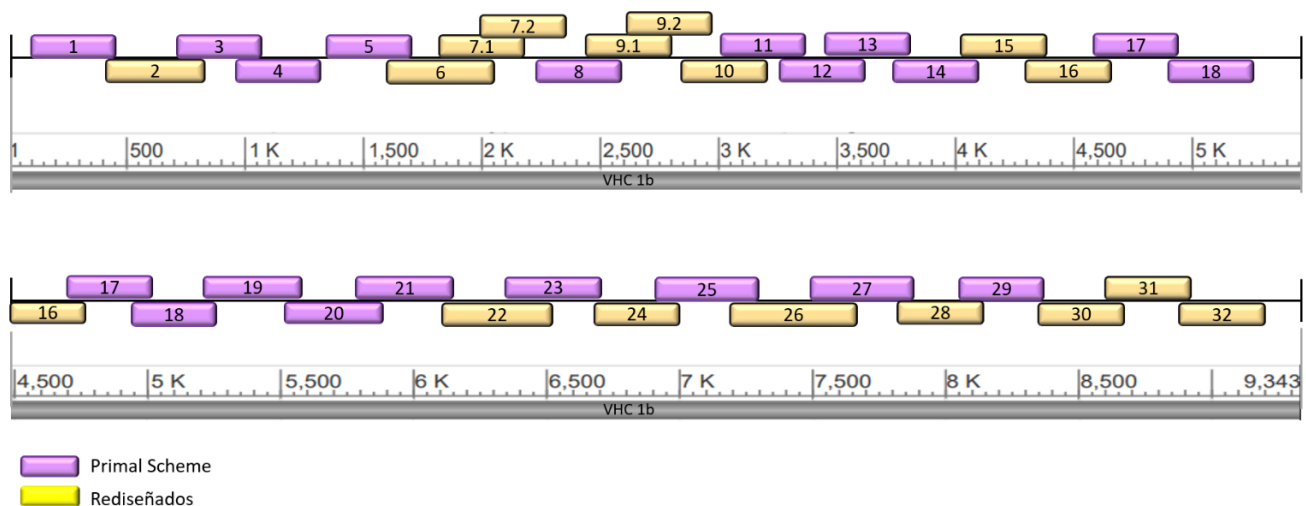
Para descartar que el problema de amplificación se debiera a una potencial degradación del material genético por un prolongado periodo de almacenamiento de las muestras, se realizó la PCR con el oligonucleótido 1a\_2 con distintas muestras de GT 1a y 2 muestras sin genotipificación (Figura 16). Ninguna muestra amplificó independientemente de la carga viral y la antigüedad.



**Figura 16** Segundo diseño de oligonucleótidos GT 1a. Gel electroforesis Agarosa 2%, 90V 5 min 95V 55 min Marcador Biorange 25-500pb. Gradiente de temperatura de 64°C-70°C. Amplificación con oligonucleótido 1a\_2. Muestras 157, 264 y 213 GT 1a, muestras 180 y RAPL no genotipadas.

### 8.2.5. Rediseño y amplificación del genotipo 1b

Una vez visualizados los oligonucleótidos que no lograron amplificar los fragmentos de 400pb, se realizó un rediseño, observando que algunos de éstos se encontraban en regiones hipervariables (sitios con una alta variabilidad en la secuencia de sus nucleótidos) del VHC, en consecuencia, el rediseño se enfocó en las regiones mayormente conservadas obteniendo nuevos amplicones *in silico* mayores y menores de 400pb. Para algunos pares se realizaron dos diseños (Figura 17 oligonucleótidos 7.1,7.2, 9.1,9.2).



**Figura 17** Rediseño de oligonucleótidos GT 1b. Los oligonucleótidos en color lila son los oligonucleótidos que amplificaron diseñados con Primal Scheme, los oligonucleótidos en color amarillo son los oligonucleótidos que no amplificaron. La línea gris en el eje "x" es la cobertura del genoma del VHC GT 1b.

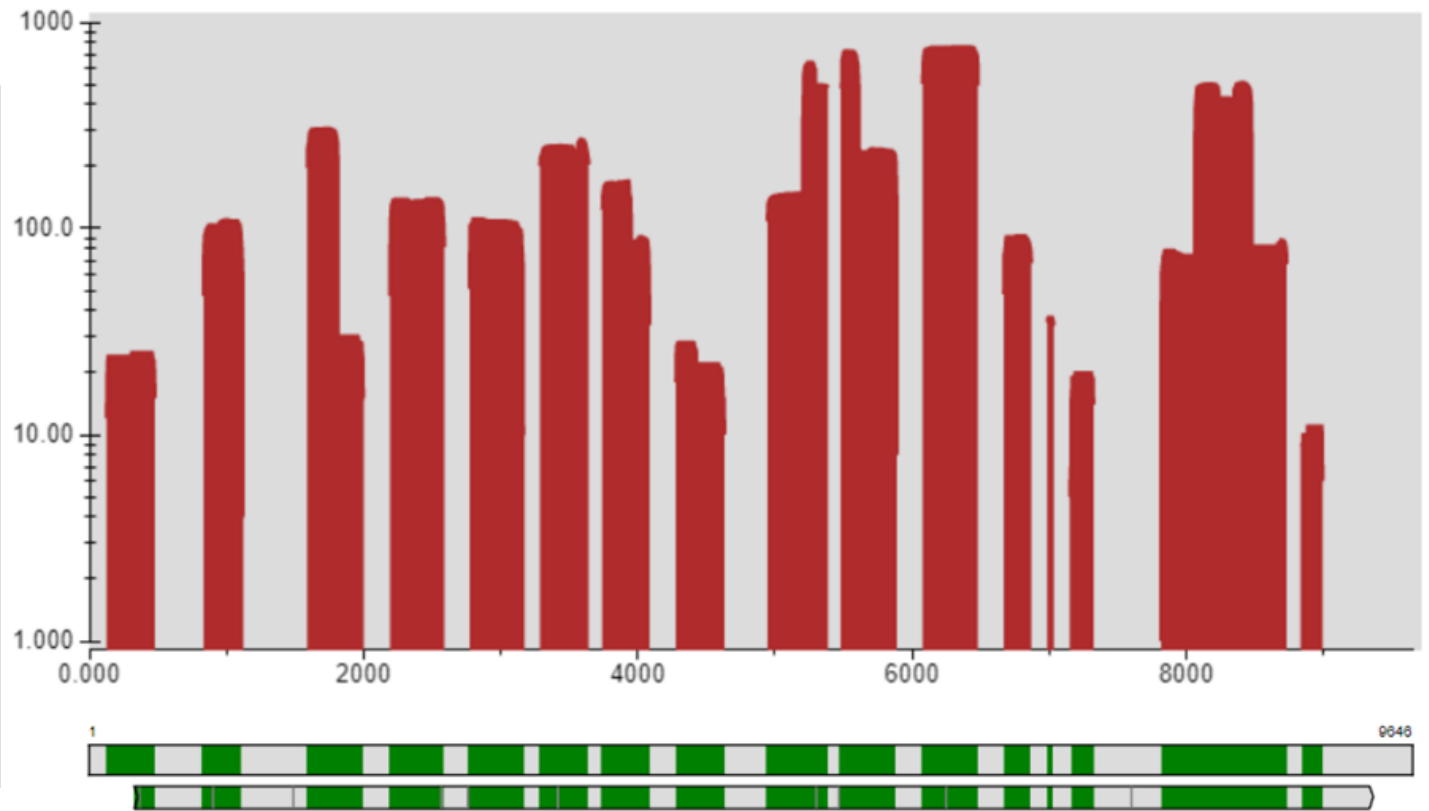
La primera secuenciación de la muestra de suero ID 429, mostró las regiones que faltaban por amplificar y las regiones que habían sido amplificadas. Se logró una cobertura total del 58% con un total de 16 contigs y se anotó la posición exacta inicial y final y el tamaño (pb) de cada contig secuenciado para poder enfocar el rediseño en las posiciones exactas faltantes de cobertura (Figura 18). Los contigs, en promedio, muestran una profundidad de lectura mayor a 100X. La anotación de las posiciones se realizó en referencia a la cepa NC\_004102.1.



En la Figura 19 se observa el alineamiento de secuencias de Estados Unidos extraídas de Los Alamos en colores. Se realizó este alineamiento con la finalidad de conocer la variación de las bases de los oligonucleótidos sintetizados respecto a los genomas virales de Estados Unidos. Las regiones en blanco muestran el dominio conservado de cada oligonucleótido en referencia al conjunto de genomas virales. En la Figura 19.A se observa un ejemplo de alineamiento de la proteína core de las secuencias del VHC con el oligonucleótido sintetizado. En este alineamiento se observa en blanco las bases que están conservadas, las que no muestran el 100% de conservación, no se resaltan en blanco.

Las bases que no mostraban una posición conservada se cambiaron por una base degenerada. Para determinar la base que tenía que ser modificada se estableció que la base no tenía que tener un porcentaje de variación igual o mayor al 60%. Para poder determinar el porcentaje de cambio de cada base en el alineamiento, se utilizó el software Jalview (Figura 19.B). Las bases que fueron identificadas con una alta variabilidad se sustituyeron por bases degeneradas.

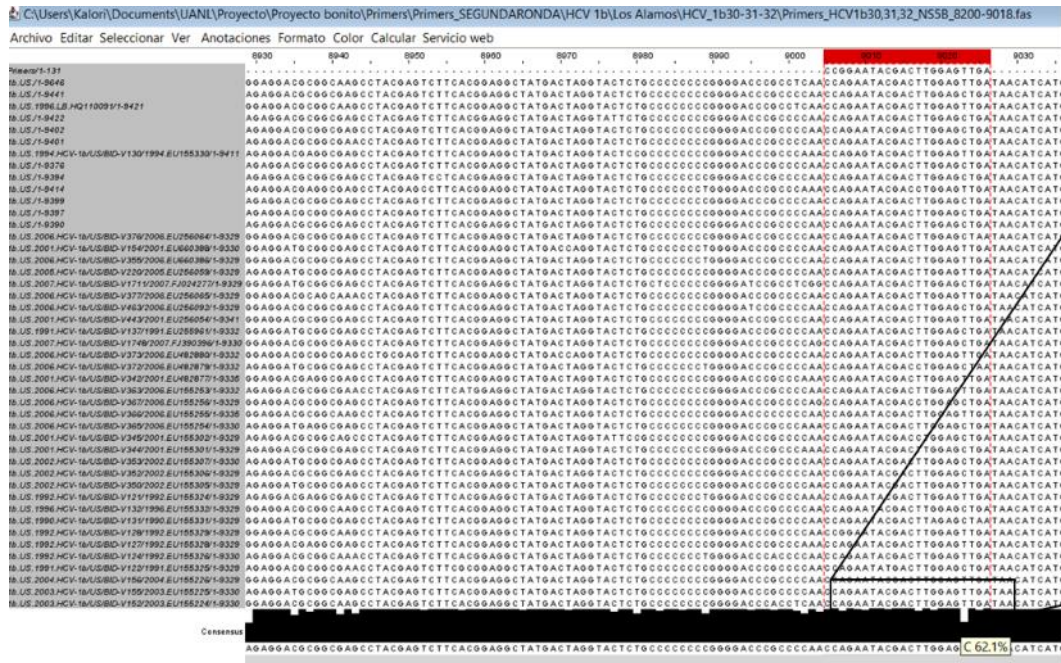
Contig	Posición inicial	Posición final	Tamaño (pb)
1	128	480	352
2	834	1120	286
3	1590	2002	412
4	2194	2590	396
5	2773	3173	400
6	3287	3642	355
7	3742	4086	344
8	4279	4632	353
9	4946	5388	442
10	5479	5888	409
11	6077	6482	405
12	6670	6869	199
13	6987	7035	48
14	7161	7325	164
15	7818	8733	915
16	8851	8991	140



**Figura 18** Primera secuenciación. La tabla muestra las posiciones iniciales y finales de cada contig (barra en rojo a la derecha) en la secuenciación. En la ilustración se muestra en el eje "y" la profundidad de la secuenciación mostrando lecturas >10.00. En el eje "x" se muestra la longitud del genoma (pb). Los rectángulos en verdes representan la cobertura del genoma, la primera línea es un genoma de referencia y la segunda línea es el virus secuenciado.

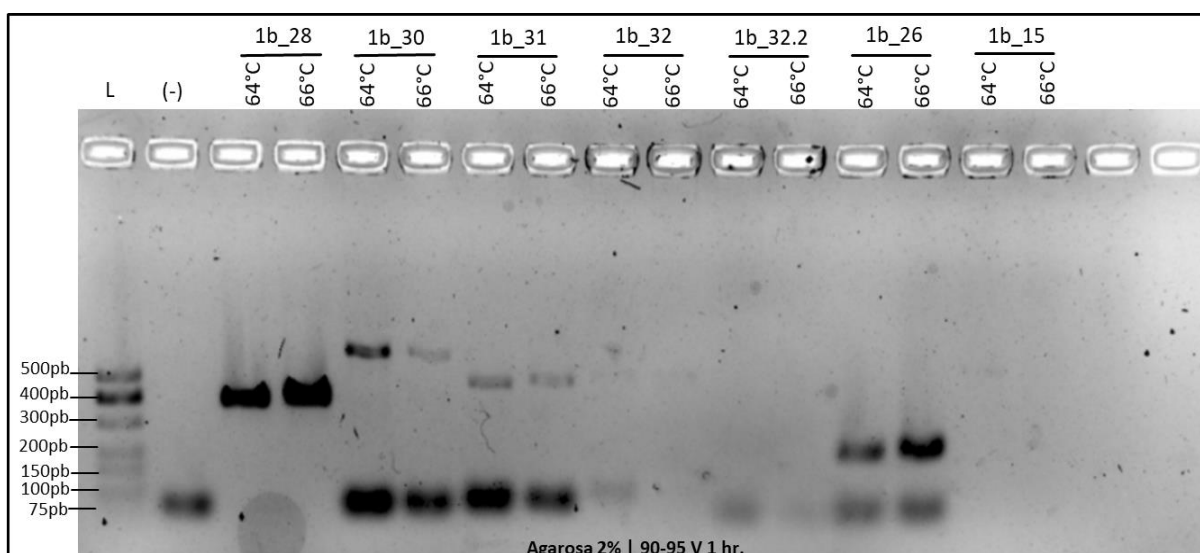
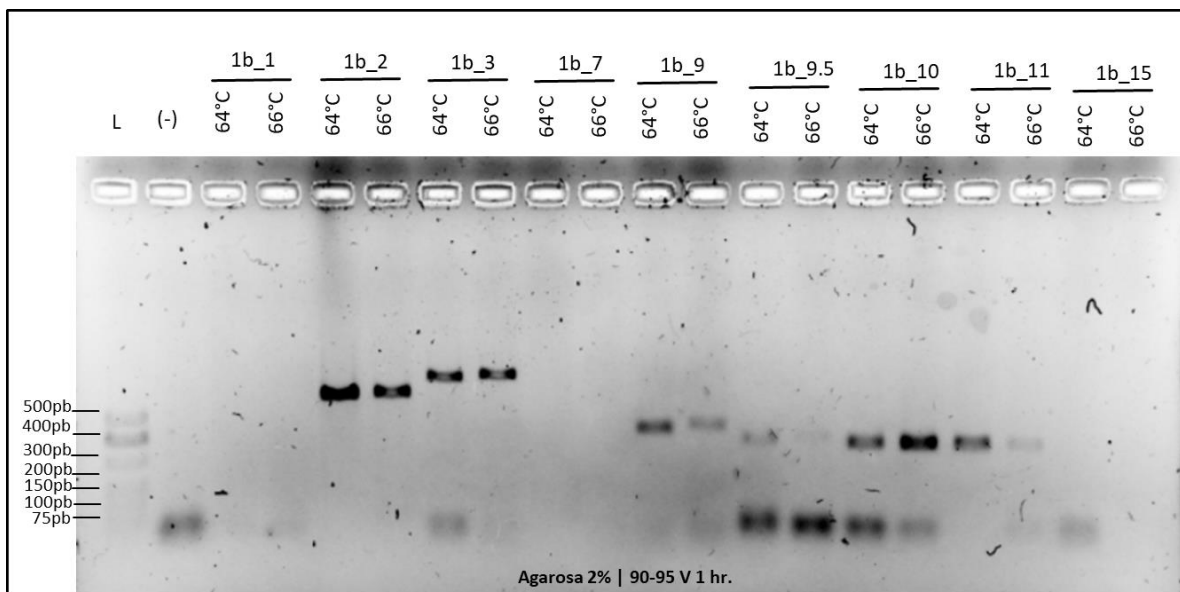


B)



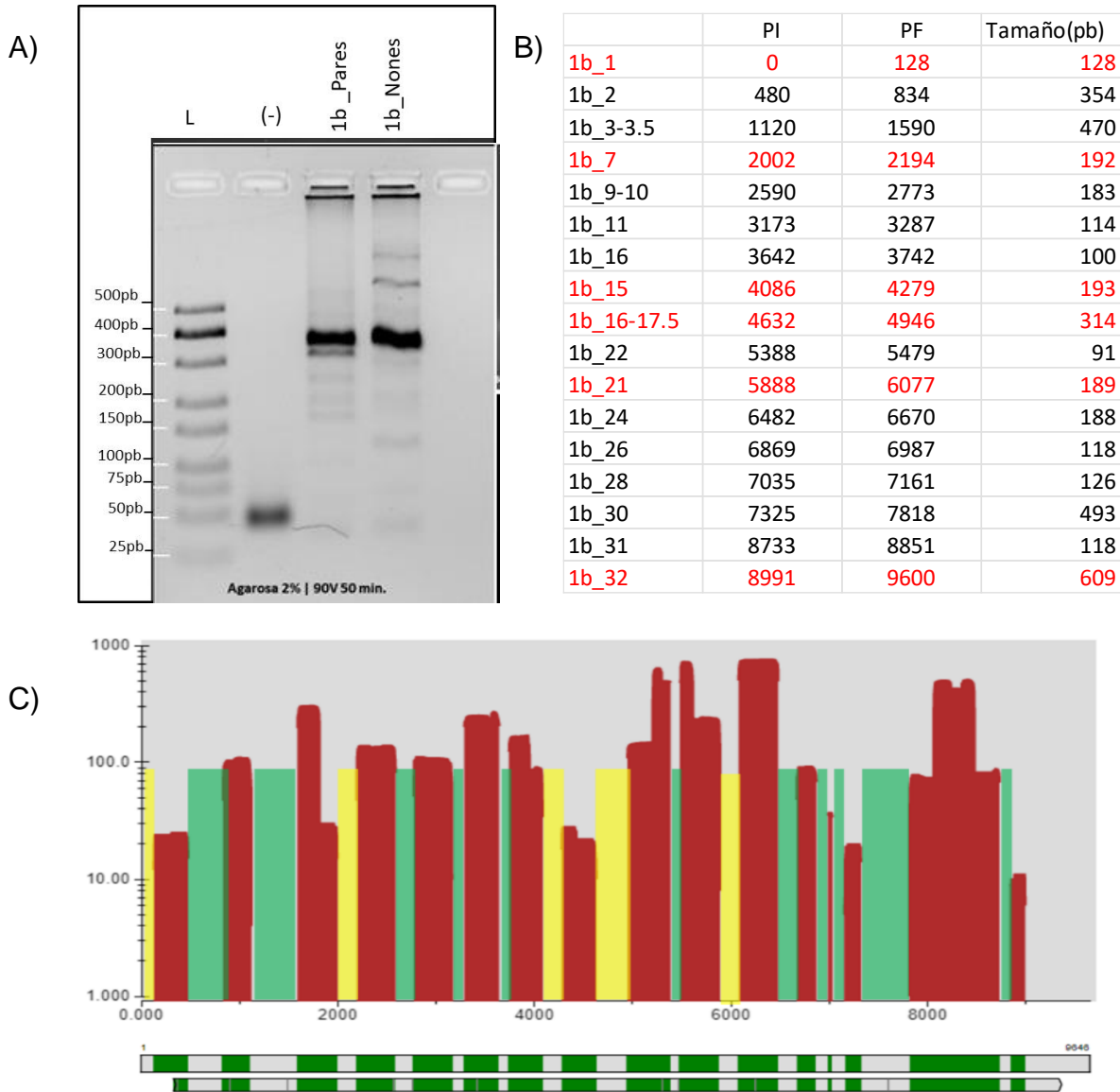
**Figura 19 .** Ejemplo del flujo de trabajo para el rediseño de oligonucleótidos GT1b. Alineamiento previamente realizado con MAFFT servidor 7. A) Visualización del alineamiento de la proteína Core con el oligonucleótido 1b\_2Right mediante Mega X. Se observan resaltadas las regiones conservadas del primer con las secuencias de Estados Unidos. B) Visualización de la misma región y el mismo oligonucleótido mediante Jalview resaltado en rojo, el acercamiento muestra el porcentaje de homogeneidad de ese nucleótido

Una vez analizados, *in silico* con Amplify4, se sintetizaron. La PCR punto final se realizó a 64°C y 66°C amplificando exitosamente los nuevos fragmentos rediseñados (Figura 21). Algunos oligonucleótidos mostraron tener Tm teóricas de hasta 70°C, por lo tanto esos oligonucleótidos no amplificaron a las temperaturas de 64°C ni 66°C y fueron descartados para realizar los pools de oligonucleótidos nones y pares para la PCR Múltiplex.



**Figura 20** Segunda amplificación de fragmentos GT 1b. Gel electroforesis agarosa 2%, 90V 5 min y posteriormente 95V 55 min Marcador Biorange 25-500pb. Los fragmentos que se observan mayores y menores a 400pb corresponden con el segundo diseño previamente confirmado *in silico*.

En esta PCR Multiplex (Figura 21) se utilizaron los oligonucleótidos de la primera etapa y los oligonucleótidos que si amplificaron de la segunda etapa, es decir, del rediseño. En la tabla se muestran los pares de oligonucleótidos sintetizados con sus posiciones en el genoma del VHC.



**Figura 21** PCR Multiplex y resumen de resultados del segundo diseño GT 1b. A) PCR Multiplex. Gel de agarosa 2%. 90v 50 min Marcador BioRange 25-500pb. Se amplificaron los fragmentos con oligonucleótidos nones y oligonucleótidos pares a 64°C. B) Tabla de posiciones en el genoma del VHC GT 1b de los oligonucleótidos del segundo diseño. PI=Posición Inicial, PF=Posición Final. En rojo se resaltan los oligonucleótidos que no amplificaron ni a 64° ni 66°C. C) Gráfico de cobertura de Genome Detective editado. En amarillo se muestran las regiones faltantes de los oligonucleótidos que no amplificaron y en verde las regiones que si amplificaron con el segundo diseño.

### 8.2.6. Secuenciación

Se logró la secuenciación de 5 muestras, obteniendo más de 200k lecturas y aproximadamente 500 lecturas N50. Se obtuvo la secuenciación completa de una muestra correspondiente al etiquetado 429 con genotipo 1b y del año 2020, y se obtuvieron 4 secuencias parciales, 3 previamente genotipadas como 1b y una genotipada como 1a. La profundidad media de secuenciación de las muestras fue de 600x. Uno de los más grandes hallazgos de este trabajo es que la muestra genotipada **previamente como 1a (resaltada en la Tabla 16) se identificó mediante el análisis bioinformático como otro genotipo el cual fue 1b.**

Se realizó la secuenciación con 3 enfoques para comparar la secuenciación con y sin PCR Multiplex.

Primer enfoque: secuenciación de la amplificación del genoma por PCR Multiplex a partir de la síntesis de ADNc con oligonucleótidos random.

Segundo enfoque: secuenciación de la retro transcripción a partir de oligonucleótidos específicos cercanos a la región 3'UTR (primer reverse 28, posición ~8300 ). La muestra de suero ID 213 fue sintetizada con el primer 1a\_28 anti sentido.

Tercer enfoque: secuenciación de los productos de la reacción de retro transcripción a partir de los oligonucleótidos específicos a la mitad del genoma (primer reverse 11, posición ~3000).

**Tabla 16** Secuenciación GT 1b

ID Muestra	Año de muestreo	Carga viral (UI/ml)	#Barcode <sup>1</sup>	# Contigs	Cobertura (%)
336	2017	1,156,919	Barcode 1_28	9	75
261	2017	685,225	Barcode 2_28	8	82.4
			Barcode 9_11	9	63.2
213	2016	2,426,335	Barcode 3_28	7	76.9
182	2016	142,205	Barcode 4_28	4	85.9
			Barcode 7_11	8	66.1
			Barcode 8_28	9	68.1
			Barcode 10_28	7	72
429	2020	2,134,082	Barcode 5_11	9	81.1
			Barcode 6_28	7	83.7
			Barcode 11 <sup>2</sup>	9	100

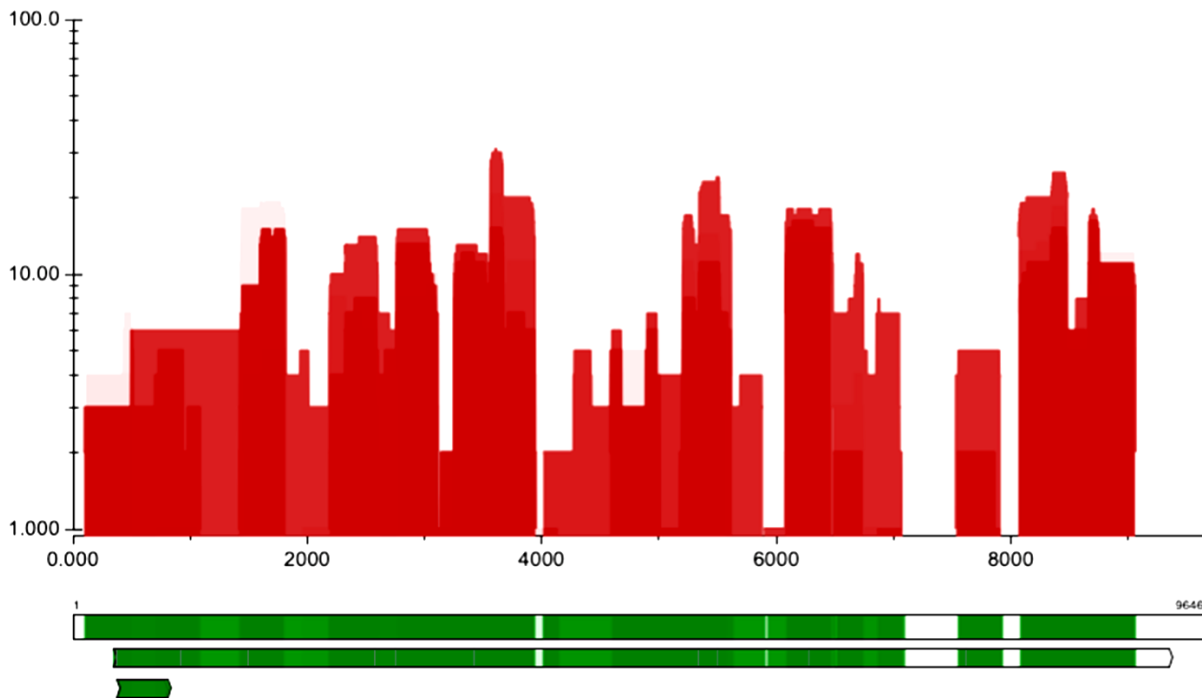
<sup>1</sup>El primer número corresponde al Barcode utilizado en el protocolo y el segundo número corresponde al primer específico utilizado para la amplificación del VHC de esa muestra.

<sup>2</sup>En esta muestra se utilizaron los oligonucleótidos nones y pares.

La secuenciación completa de la muestra de suero ID 429 comienza en el nucleótido 53 y termina en el nucleótido 9319, en referencia a la cepa NC\_004102.1 (Figura 22), con una longitud de 9264 nucleótidos. El resultado se registró en **GenBank** con el número de acceso **ON676559**.



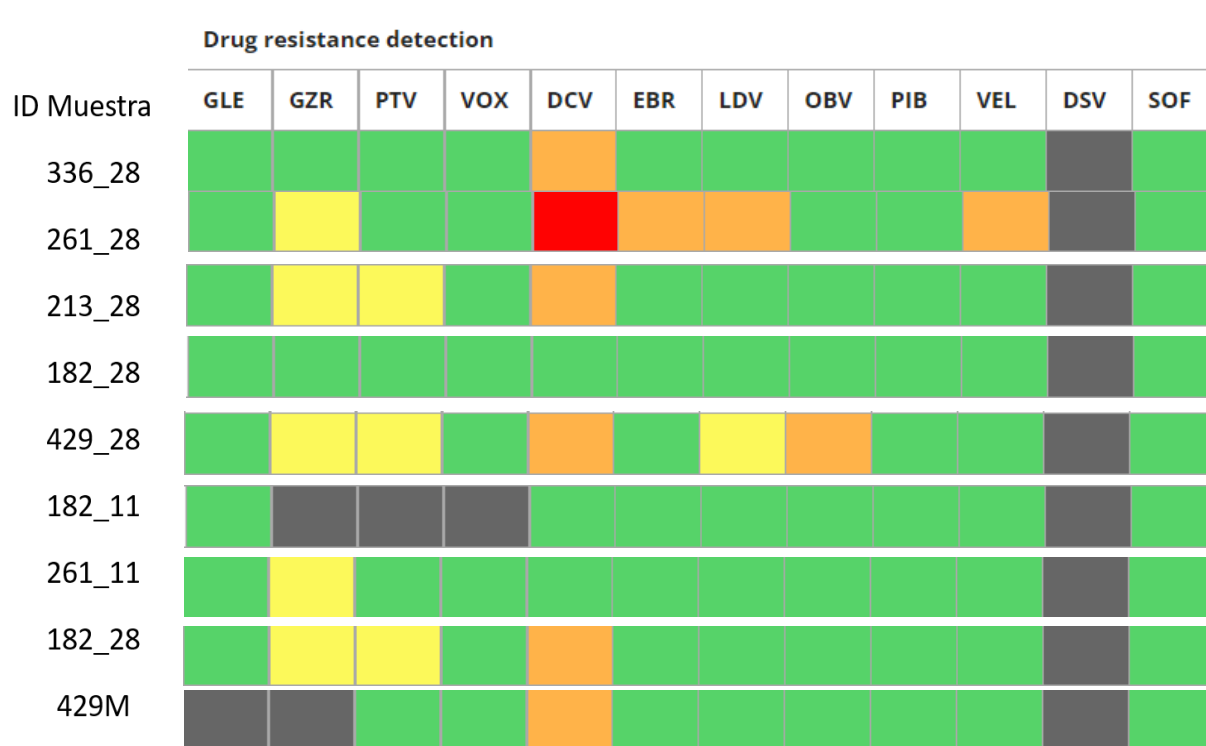
Las regiones que mostraron una mayor falta de cobertura en las secuencias parciales corresponden a los genes E1(907-1487), E2(1487-2570) y NS5A (6252-7602).



**Figura 22** Gráficos de Genome Detective sobrelapados de la segunda secuenciación del GT1b. En el eje "x" se muestra la longitud del genoma en pb y en el eje "y" se muestra la profundidad de la secuenciación. El gráfico en color verde muestra la cobertura del genoma del VHC secuenciado de referencia NC\_004102.1.

La identificación de resistencias a antivirales se realizó mediante el software en línea HCV-Glue(B Singer et al., 2018). Ésta es una base de datos que contiene las secuencias de VHC reportadas en conjunto con los metadatos de NCBI. Estos metadatos incluyen los polimorfismos asociados a resistencia a antivirales de acción directa. A través de las herramientas de análisis que incluye este software, es posible realizar el genotipado y el análisis de resistencia a DAA y su visualización a través de las secuencias FASTA.

Se observa como las secuencias del VHC de las muestras de suero de los pacientes con ID 261, 213, 429 y 182 demuestran una posible resistencia a grazoprevir y peritapretavir. Se detectó una resistencia al daclatasvir en la secuencia del VHC de la muestra de suero ID 261.



**Figura 23** Detección de resistencia a DAAs. Verde: no hay resistencia. Amarillo: posible resistencia. Naranja: probable resistencia. Rojo: se detectó resistencia. Gris: no existe suficiente investigación o cobertura para ese fármaco.

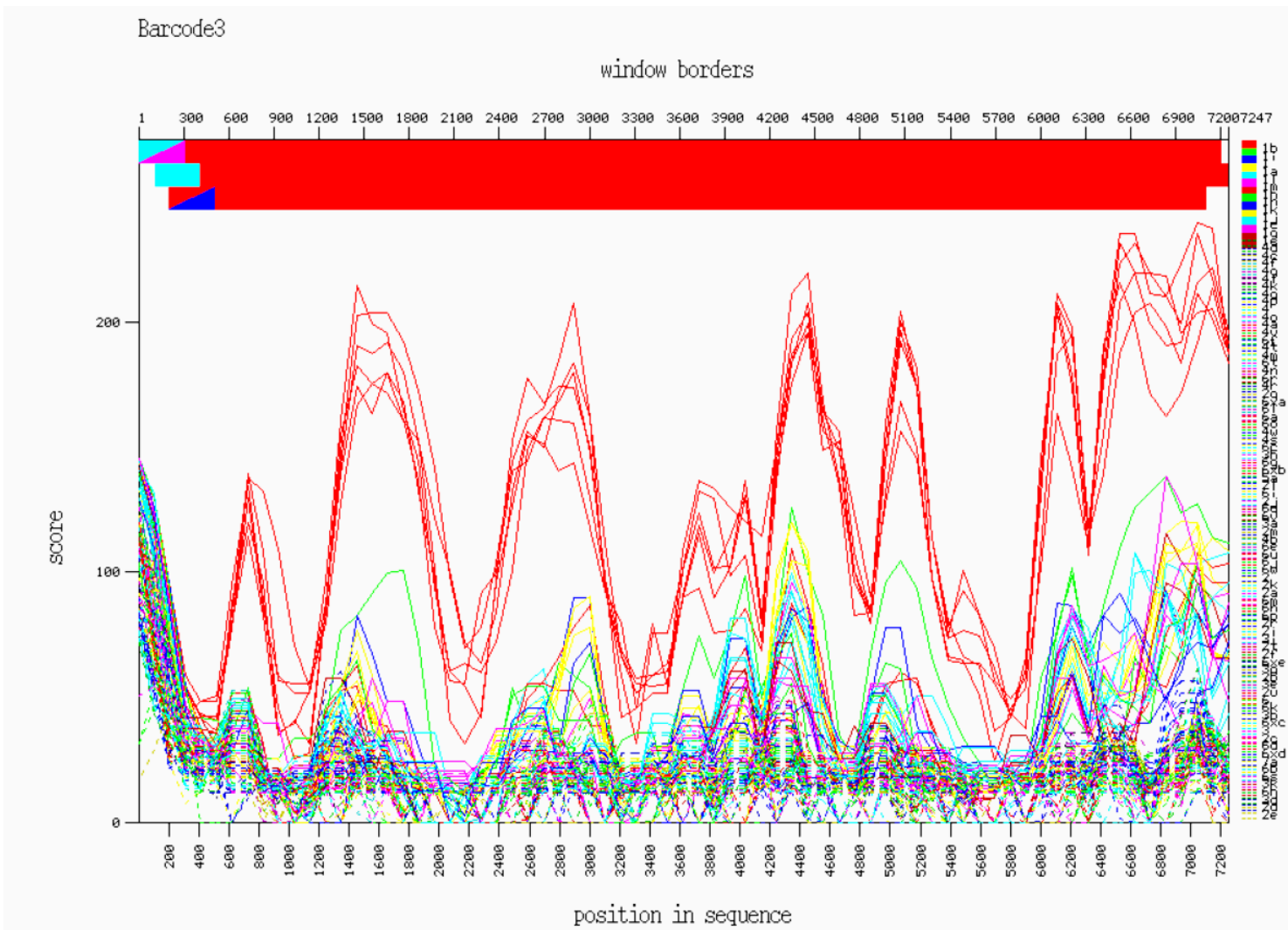
Las sustituciones asociadas a resistencia se describen en la siguiente tabla de acuerdo con los resultados obtenidos por HCV-Glue.

**Tabla 17** RAS de las secuencias VHC GT 1b

RAS							
ID Muestra	Inhibidor NS5A					Inhibidor NS3/4A	
	DCV	EBR	LDV	VEL	OBV	GZR	PTV
336_28	F/L37L						
261_28	32del	32del	32del	32del		D168G	
213_28	F/L37L					R155G	D168T
429_28	F/L37L		Y93C		Y93C	R155G	D168T
261_11						D168G	
182_28						A156G	
429M	F/L37L						

DCV: Daclastavir. EBR: Elbasvir. LDV: Ledipasvir. VEL: Velpatasvir. GZR: Grazoprevir. PTV: Paritaprevir. OBV: Ombitasvir

Al ser identificada la muestra de suero ID 213 como genotipo 1b y no como genotipo 1a como se había identificado previamente por la Unidad de Hígado, mediante la secuenciación, se decidió realizar un segundo análisis bioinformático capaz de confirmar este hallazgo y así poder tratar los datos de esa secuencia como genotipo 1b. Para ello se utilizó la herramienta de genotipado de NCBI, Genotyping Tool (<https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>.)



**Figura 24** Genotipado muestra 213. NCBI Genotyping Tool. Eje "x" posición en secuencia del VHC. Eje "y" lado izquierdo el score para cada posición y lado derecho el genotipo ordenado por puntaje máximo obtenido al comparar esa región con las secuencias en BLAST. La barra superior muestra los puntajes para cada ventana pintándose del color del genotipo que haya pasado el limite del score en coincidencia con la secuencias en BLAST. Genotipo 1b se muestra en rojo

## **8.2.7. Análisis filogenético**

### **8.2.7.1. Análisis filogenéticos en México**

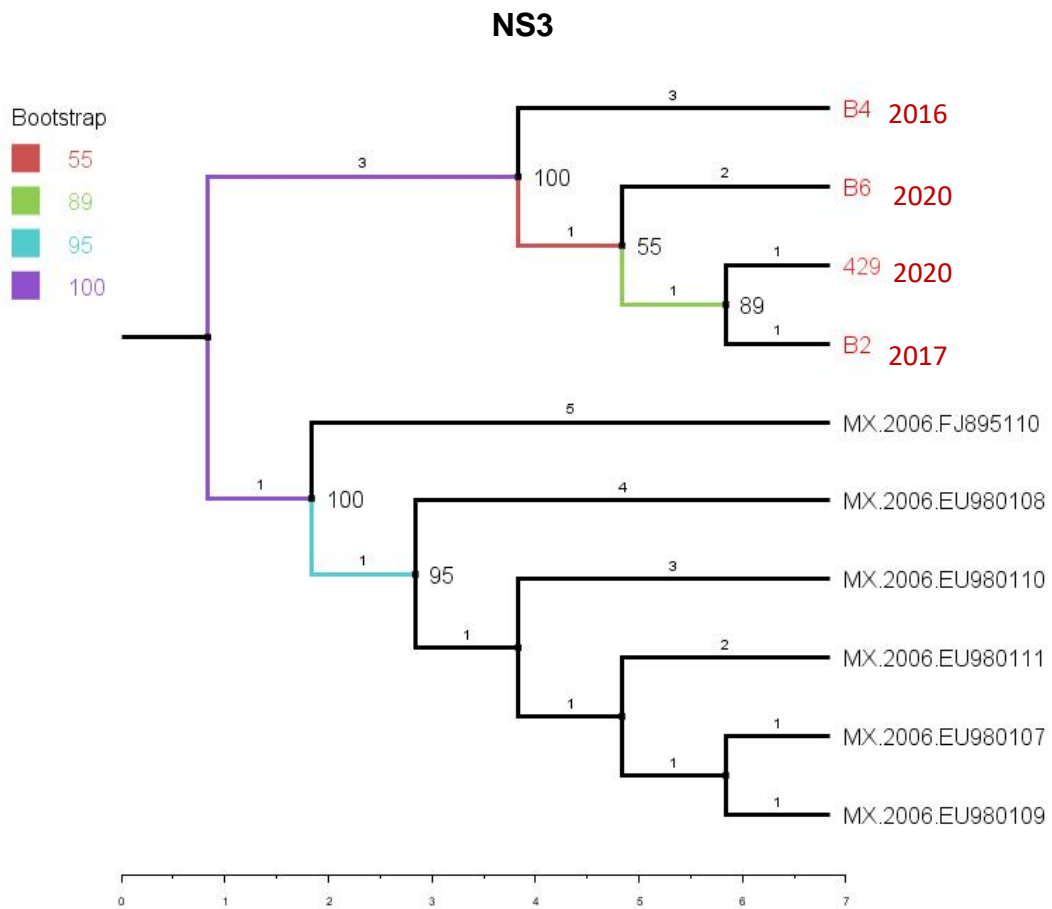
Para este análisis únicamente se muestran filogenias de la nueva secuencia con los genes del VHC previamente reportados en México, éstos corresponden a los genes NS3, NS4A y NS5A, obtenidos de la base de datos “Los Alamos”. Algunas secuencias parciales mostraron una cobertura completa de los genes NS3 y NS4A, por lo que se realizaron dos filogenias para el gen NS3: una filogenia única de secuencias NS3 (Figura 25) y otro análisis filogenético de ambos genes (Figura 26).

El árbol filogenético de máxima verosimilitud del gen NS3 (Figura 25) de los virus secuenciados en Nuevo León, no muestra un agrupamiento con el gen NS3 de los virus previamente reportados secuenciados en México. El clado que no se agrupa muestra tener un valor de confiabilidad Bootstrap igual a 100.

El árbol filogenético de máxima verosimilitud de los genes NS3/4A (Figura 26) de los virus obtenidos de Nuevo León, muestra que estos genes se encuentran dentro de un clado con los virus previamente reportados en México en el 2006. Los dos cladros muestran tener un valor de confiabilidad Bootstrap igual a 94, mientras que el sub-clado en donde se agrupan los virus de Monterrey, señalados en color naranja, muestran tener un valor inferior de confiabilidad Bootstrap igual a 42.

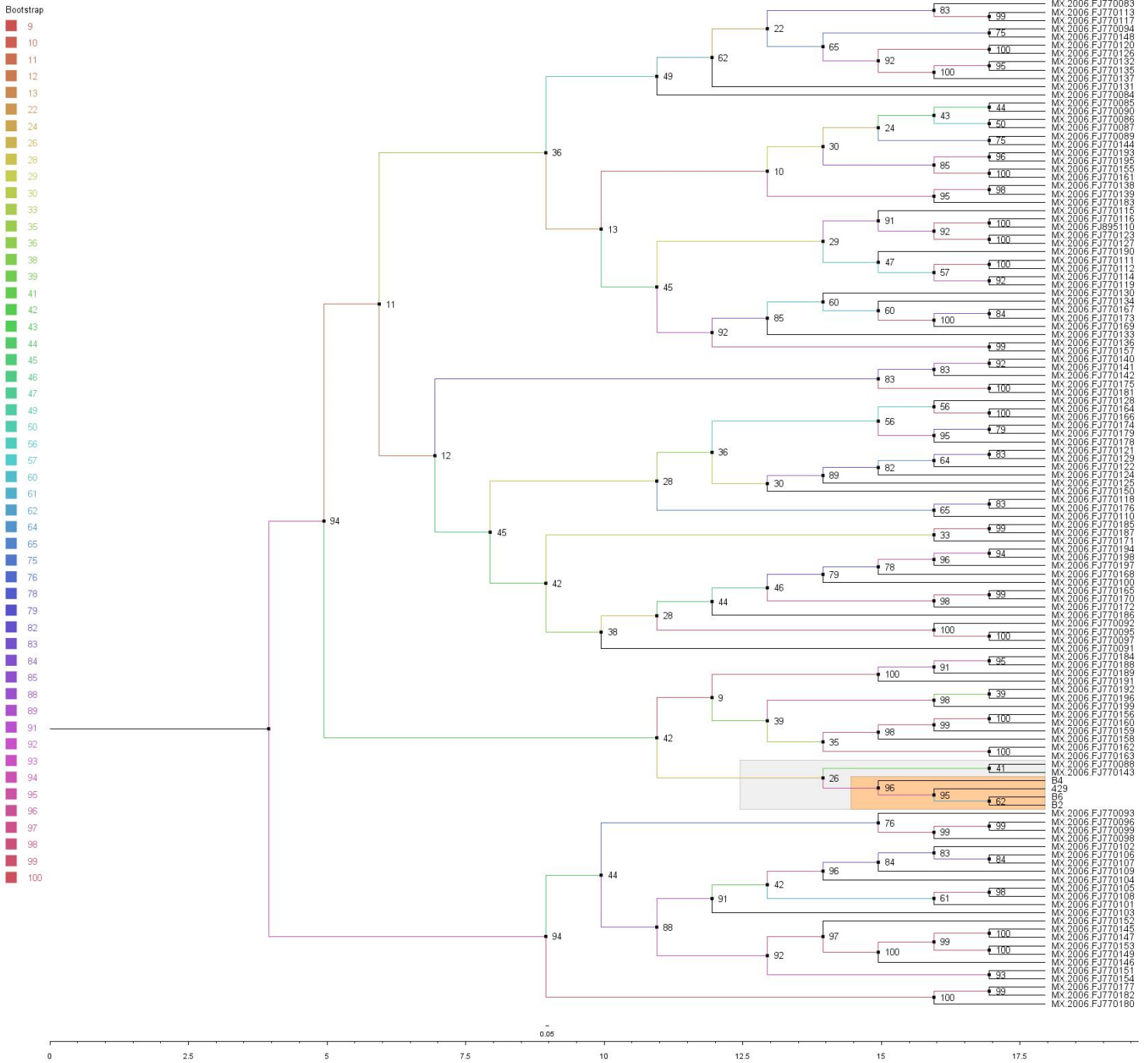
El árbol filogenético de máxima verosimilitud del gen NS4A (Figura 27) muestra el agrupamiento del gen NS4A de los virus secuenciados con los virus circulantes reportados en México en el 2010 con un valor de confiabilidad Bootstrap igual a 100. Además, se observa que los genes de los virus circulantes en Nuevo León se agrupan en un clúster con la secuencia AB472149.

El árbol filogenético del gen NS5B (Figura 28) muestra un agrupamiento de este gen de los virus secuenciados de Nuevo León, con los virus reportados circulantes en México en el 2010. Se observa un valor de confiabilidad de Bootstrap igual a 92 en el clúster donde se agrupa la secuencia AB472188 con los virus secuenciados.



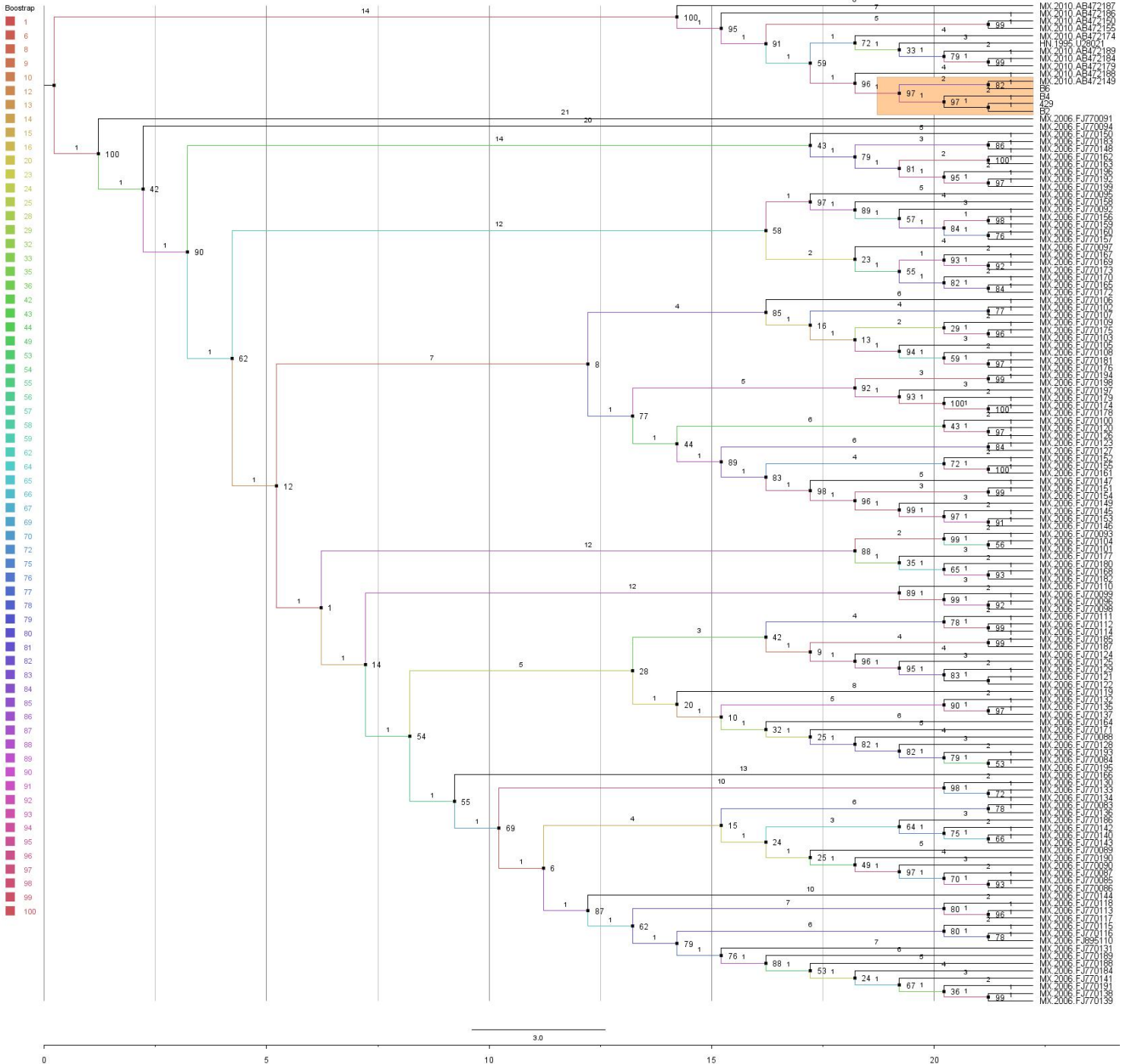
**Figura 25** Árbol filogenético de máxima verosimilitud con el modelo TPM2+F+G4 del gen NS3 en México. Los virus secuenciados se muestran en un color rojo. Los nodos representan el valor del Bootstrap. Las cepas se indican por las iniciales del país, seguido por el año y número de acceso.

# NS3-4A



**Figura 26** Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 de los genes NS3 y NS4A en México. Los virus secuenciados se muestran en un clado señalado en color naranja. Los nodos representan el valor del Bootstrap. Las cepas se indican por las iniciales del país, seguido por el año y número de acceso.

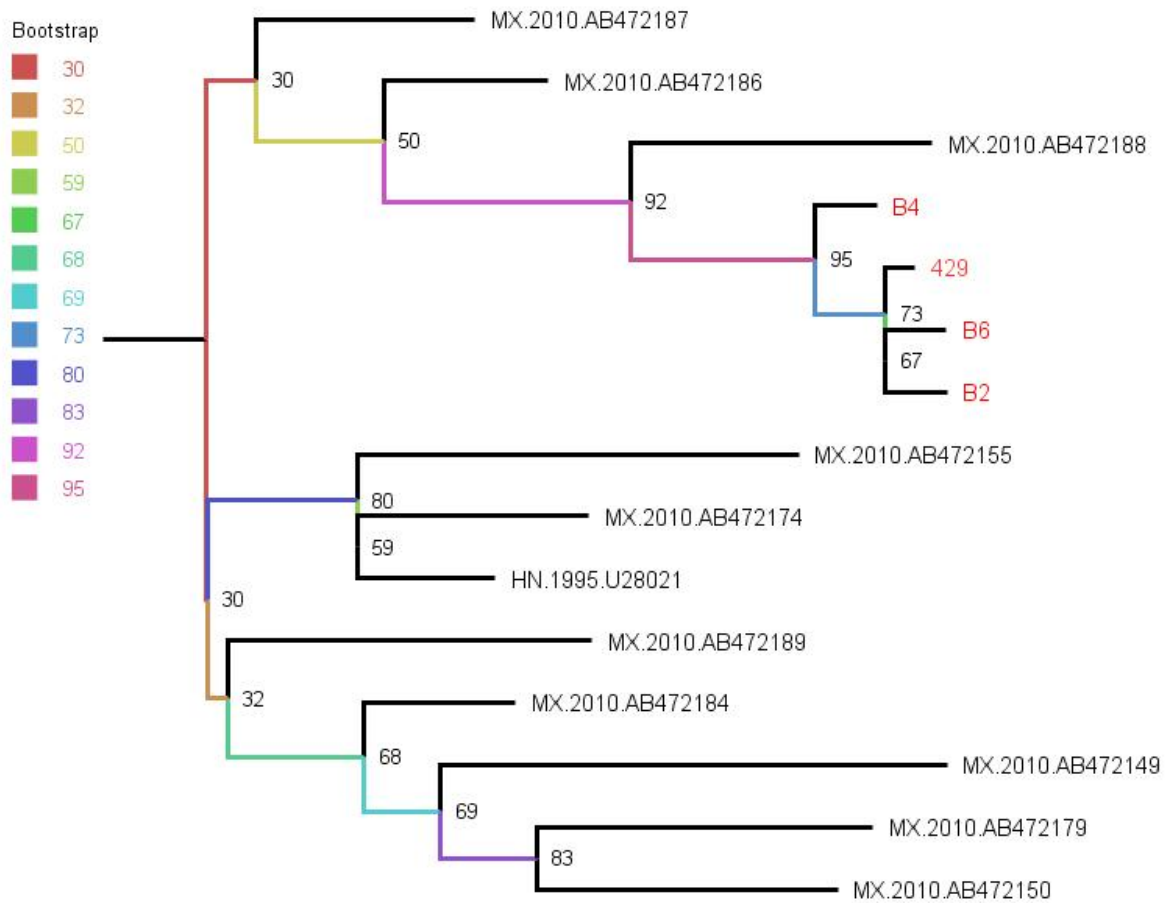
# NS4A



**Figura 27** Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del gen NS4A en México. Los virus secuenciados se muestran en un clado en color naranja. Los nodos representan el valor del Bootstrap. Las cepas se indican por las iniciales del país, seguido por el año y número de acceso.



## NS5B



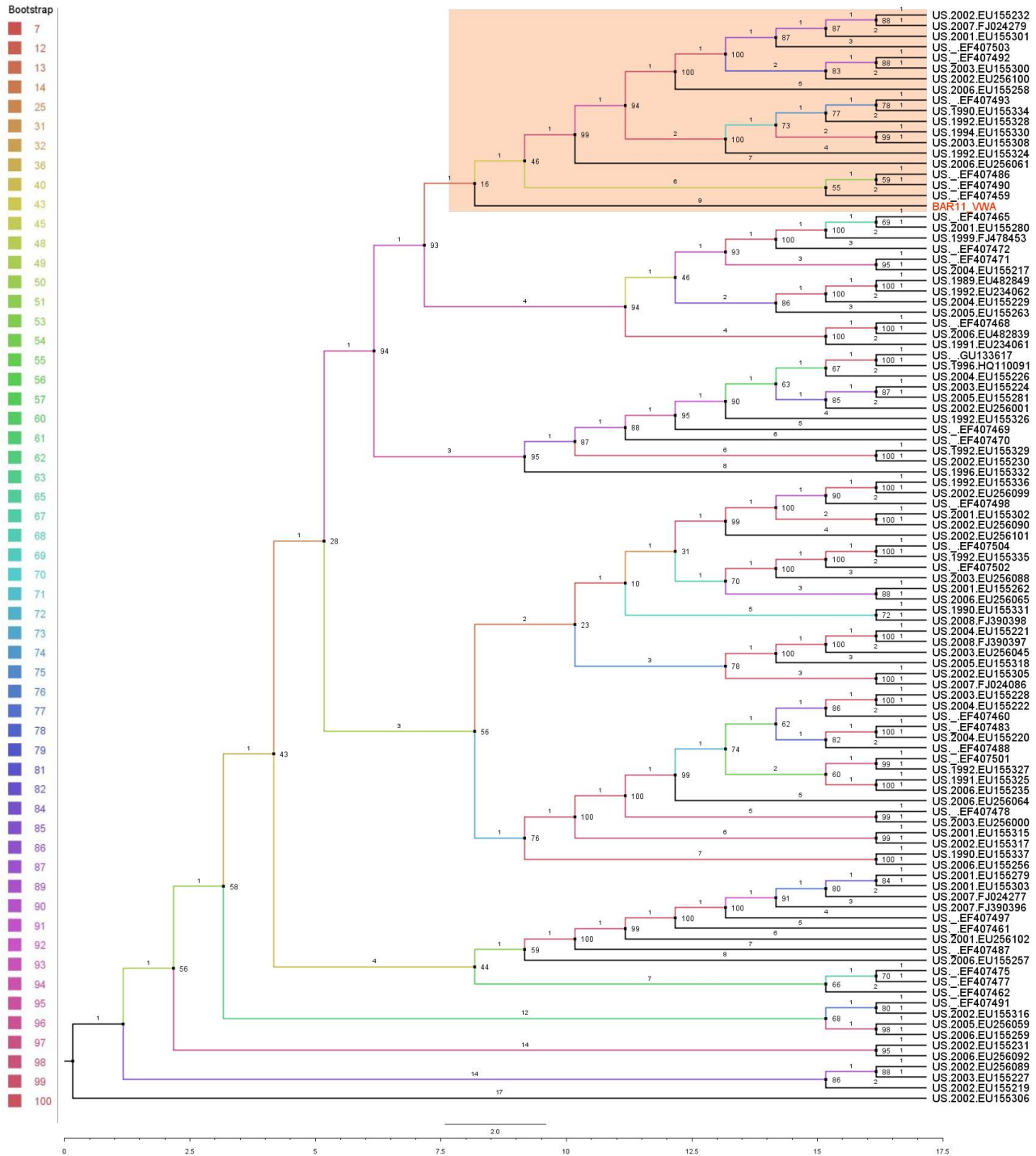
**Figura 28** Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del gen NS5B en México. Los virus secuenciados se muestran señalados en color rojo. Los nodos representan el valor del Bootstrap. Las cepas se indican por las iniciales del país, seguido por el año y número de acceso.

### **8.2.7.2. Análisis filogenéticos internacionales**

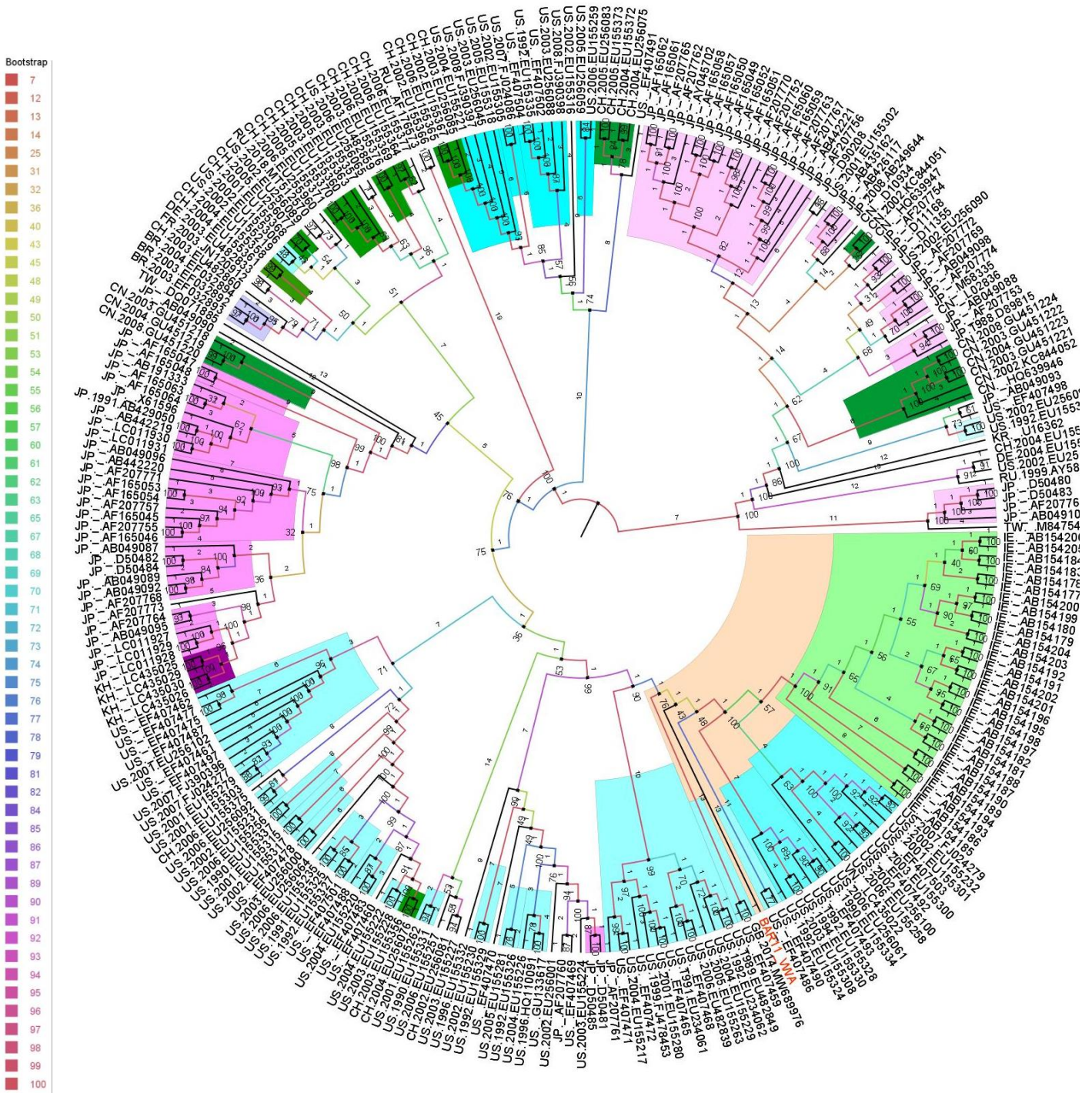
En el árbol filogenético de máxima verosimilitud de las secuencias completas de los VHC de Estados Unidos de América y la secuencia del VHC de Nuevo León (Figura 29, BAR11 VWA) se observa un clúster suportado con un valor de confiabilidad Bootstrap igual a 93 cuyos virus circulantes pertenecen a los años 90's, 2001, 2002 y 2007. Sin embargo, no se agrupa con las secuencias pertenecientes a ese clúster. Se observa una similitud de éste con los virus circulantes EF407486, EF407490 y EF407459 en el año 2007.

En el árbol filogenético de máxima verosimilitud de las secuencias internacionales y la secuencia de Nuevo León (Figura 30) se observa que la secuencia de VHC neolonesa está agrupada en un clúster con secuencias de los virus circulantes en Estados Unidos de América de los años 2006 y 2014 con un valor de confiabilidad Bootstrap igual a 90. No obstante, el VHC neolonesa dentro de este clúster no se agrupa con ningún otro virus. Además, en el clúster señalado en color naranja, donde se agrupa el virus de Nuevo León, se observa s agrupamiento con los virus de Estados Unidos de América y los virus circulantes en Irlanda en el año 2003.

## Análisis filogenético con secuencias de Estados Unidos de América



**Figura 29** Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del virus secuenciado en México con las secuencias de Estados Unidos de América. La secuencia del VHC está señalada en rojo y se resalta su clado en color naranja. Los nodos representan el valor del Bootstrap. Las cepas se indican por las iniciales del país, seguido por el año y número de acceso.



**Figura 30** Árbol filogenético de máxima verosimilitud con el modelo GTR + G4 del virus secuenciado en México con las secuencias a nivel mundial. La secuencia del VHC está señalada en rojo y se resalta su clado en color naranja. Los nodos representan el valor del Bootstrap. Las cepas se indican por las iniciales del país, seguido por el año y número de acceso. Los clados señalados representan el país. Azul: Estados Unidos de América, Verde claro: Irlanda, Verde oscuro: China, Rosa: Japón, Morado: Cambodia o Vietnam, Lila: Brazil. Los clados en blanco son secuencias similares pero de distintos países.

### **8.2.8. Análisis evolutivo con TempEST**

Para encontrar secuencias problemáticas que interfieren con la construcción de filogenias se utilizó TempEST. Las secuencias identificadas como problemáticas (las que aparecían lejanas a la pendiente y sin agrupamiento en los residuales) fueron eliminadas de los alineamientos.

#### **8.2.8.1. Análisis nacional**

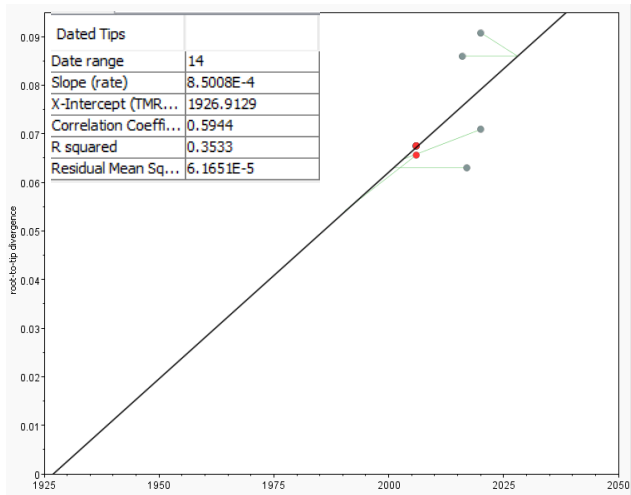
El enraizamiento se realizó con el método de la media residual heurística al cuadrado para las secuencias de los genes NS3, NS3-4A y NS5B (Figura 31).

Para el gen NS4A (Figura 31.C) se utilizó el método de correlación al cuadrado para obtener una pendiente positiva.

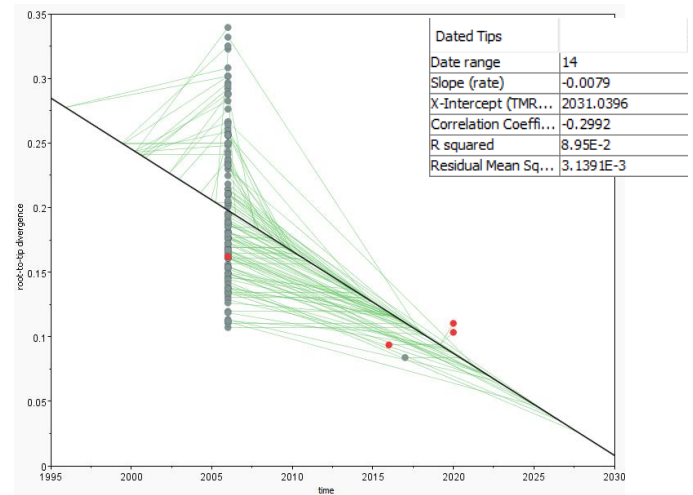
Únicamente los genes NS3 y NS5B (Figura 31.A y D) mostraron una correlación positiva, por lo que solamente éstos se utilizaron para el análisis evolutivo a través del tiempo. El ancestro en común data en el año 1926 y 1983 de cada gen respectivamente. Sus tasas de sustitución por sitio/por año fueron de  $8.50 \times 10^{-4}$  y  $2.3 \times 10^{-3}$ .

Los genes NS3/4A ((Figura 31.B)) no mostraron una correlación positiva con el método de la media residual heurística al cuadrado ni con el método de correlación al cuadrado, implicando así que estos genes no lograron cumplir los criterios para poder ser analizados mediante una filogenia evolutiva a través de TempEST.

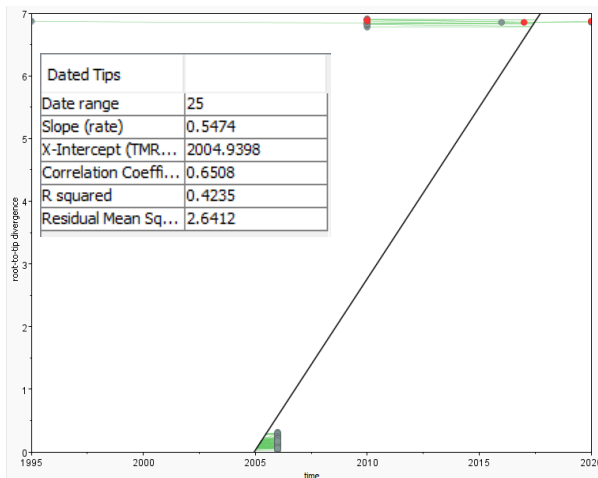
A)



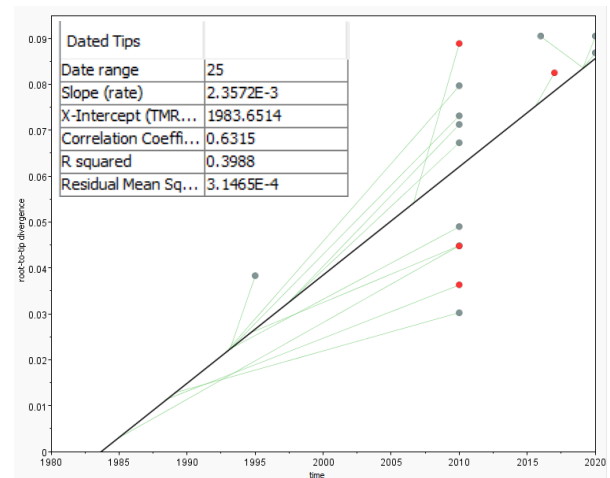
B)



C)



D)



**Figura 31** Análisis evolutivo TempEst con el método media residual heurística al cuadrado. Los virus secuenciados se señalan en rojo. Las líneas verdes relacionan a la secuencia con la pendiente que representa el tiempo en donde se encuentra el ancestro común. A) Gen NS3 B) Genes NS3 y NS4A C) Gen NS4A D) Gen NS5B.

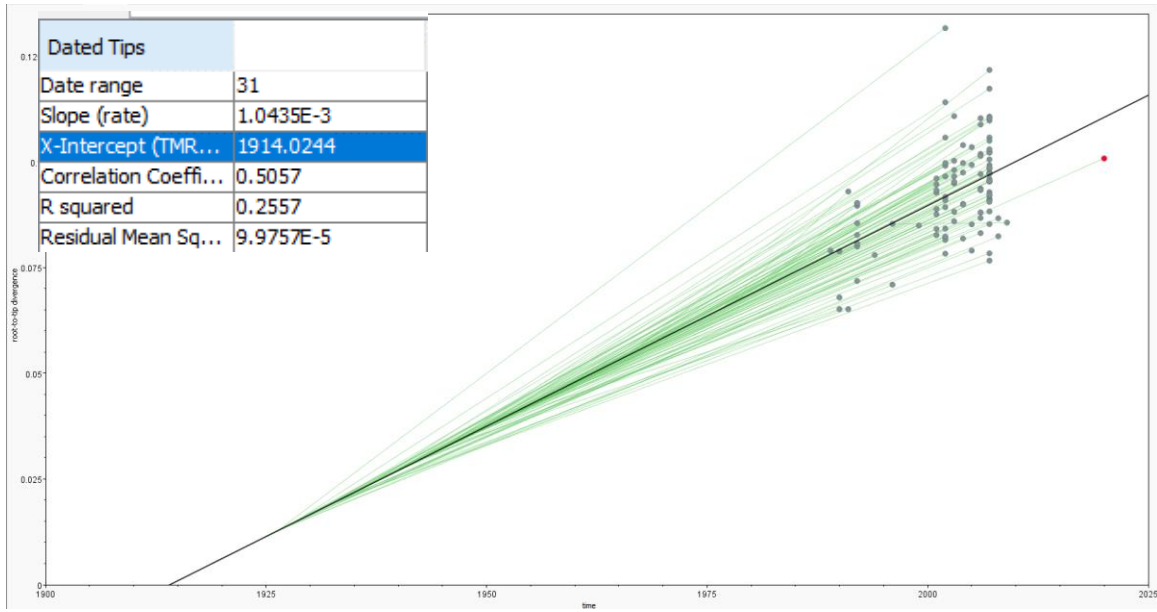
### **8.2.8.2. Análisis internacional**

Análisis en TempEst, el enraizamiento se realizó con el método de la media residual heurística al cuadrado para las secuencias de Estados Unidos de América y las globales. Los análisis muestran una tasa de  $1.043 \times 10^{-3}$  y  $7.9216 \times 10^{-4}$  sustituciones por sitio/año respectivamente. El ancestro en común en el año 1914 en el caso de las secuencias con Estados Unidos y 1865 con las secuencias globales. Ambos análisis realizados por TempEst mostraron una correlación positiva mediante el método de la media residual heurística al cuadrado.

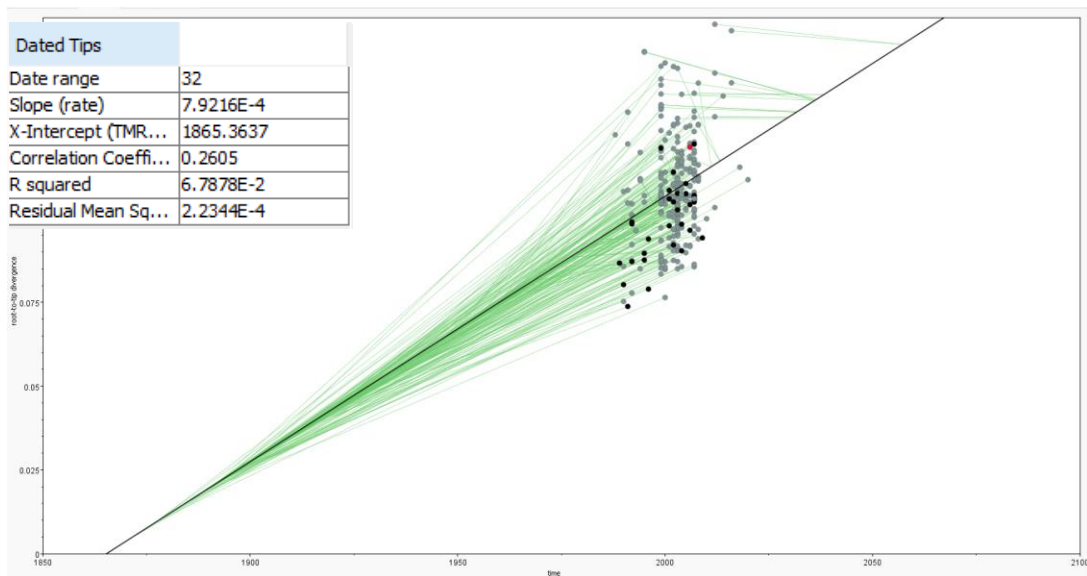
Realizando el análisis con las secuencias de Estados Unidos de América, se observa que el virus mexicano aparece 10 años después que los reportados en el 2010 de Estados Unidos de América, en la gráfica (Figura 32.A) se observa la falta de secuenciación del VHC durante esta década.

El análisis del VHC mexicano con las secuencias globales muestra que a nivel mundial se han estado reportando secuencias del VHC durante los años 90's, 2000, 2010 y 2020.

A)



B)



**Figura 32** Análisis evolutivo TempEst. El virus secuenciado se señala en rojo. Las líneas verdes relacionan a la secuencia con un punto en el tiempo donde debería estar el ancestro común. A) Secuencia de Estados Unidos B) Secuencias globales.



### **8.2.9. Análisis filogenético evolutivo con BEAST 2.0**

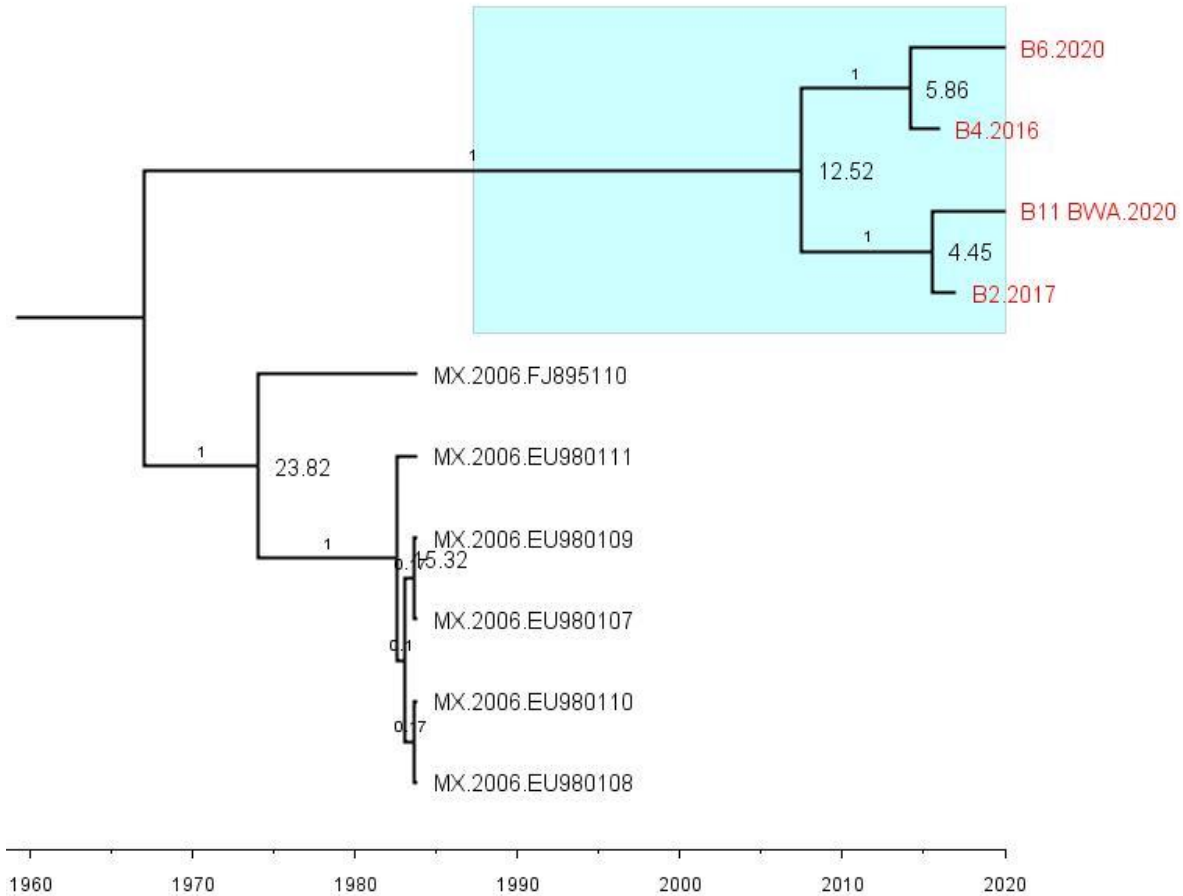
Se realizaron análisis filogenéticos evolutivos de los conjuntos de datos que no mostraron secuencias problemáticas que pudieran impedir el análisis temporal realizado en TempEST. Estos conjuntos de datos pertenecieron a los genes NS3 y NS5B y los análisis internacionales de las secuencias de Estados Unidos de América y las secuencias globales.

El archivo .beauti se generó bajo los siguientes parámetros: modelo de sustitución GTR Gamma 4 obtenido previamente con IQ-Tree, reloj molecular estricto, modelo coalescente de tamaño constante para la generación del árbol posterior y el largo de la cadena Markov Monte Carlo (MCMC) de  $40 \times 10^6$ . Para la unificar los árboles generados por BEAST 2.0 se utilizó TreeAnnotator con un límite de probabilidad a posteriori de 0.9.

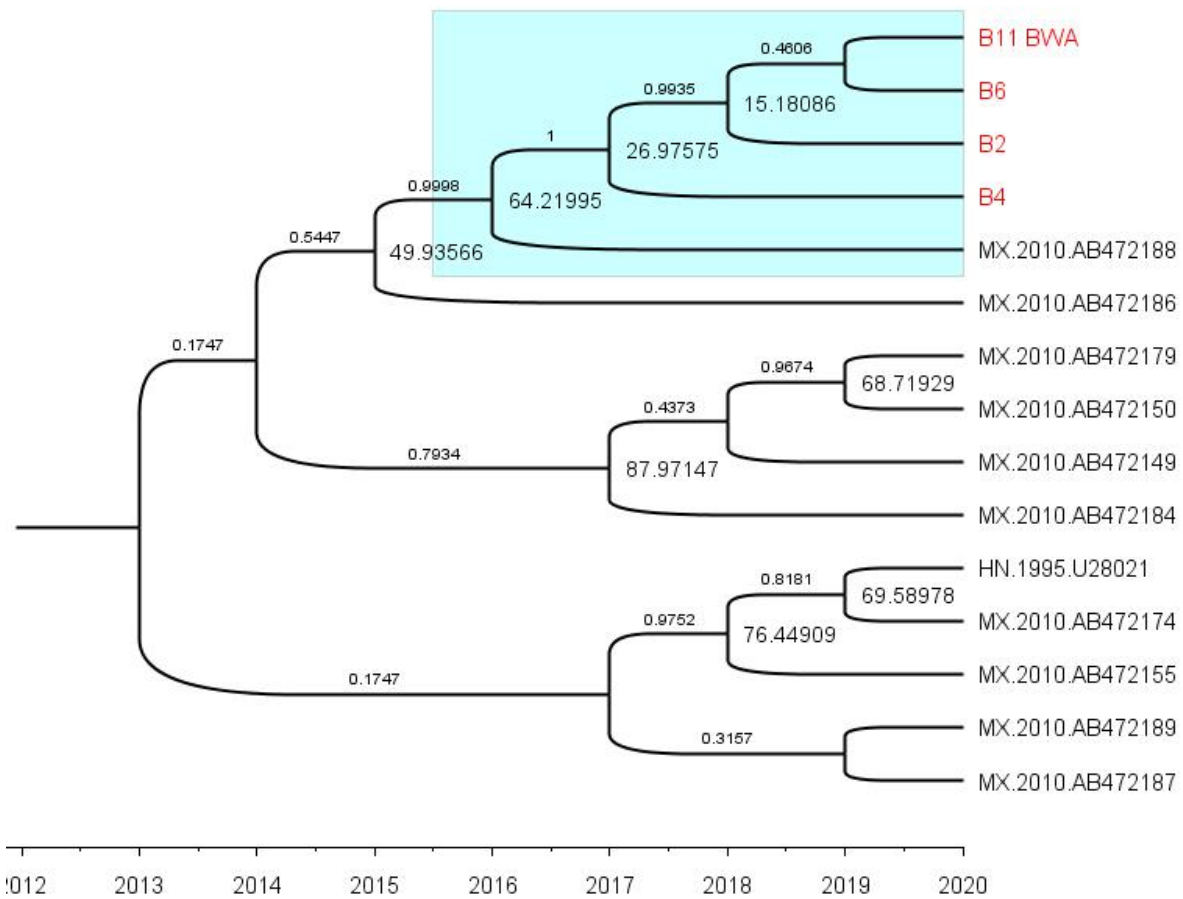
El análisis filogenético evolutivo a nivel nacional del gen NS3 (Figura 33) muestra una distribución de clados igual a la filogenia de máxima verosimilitud. Los nodos muestran una alta posibilidad (igual a 1) de que el clado que agrupa las secuencias mexicanas se hayan originado hace 12.5 años.

El análisis filogenético evolutivo a nivel nacional del gen NS5B (Figura 34) muestra a la par una distribución de clados igual a la filogenia de máxima verosimilitud. Las secuencias de los VHC de Nuevo León se agrupan en un clado con una secuencia del gen NS5B de un virus circulante en México en el 2010, AB472188. Estos virus pertenecientes a este clado tienen una alta posibilidad de origen de hace 64 años.

### 8.2.9.1. Análisis nacional



**Figura 33** Filogenia evolutiva del Gen NS3 en México. Los virus secuenciados se señalan en rojo agrupados en un clado resaltador en azul. Los nodos representan los años que han pasado desde su origen. Las ramas representan la posibilidad a posteriori. Kingman JFC (1982) Stoch Proc Appl 13, 235-248 [Constant Coalescent]. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Genetics 161, 1307-1320 [Serially Sampled Data].



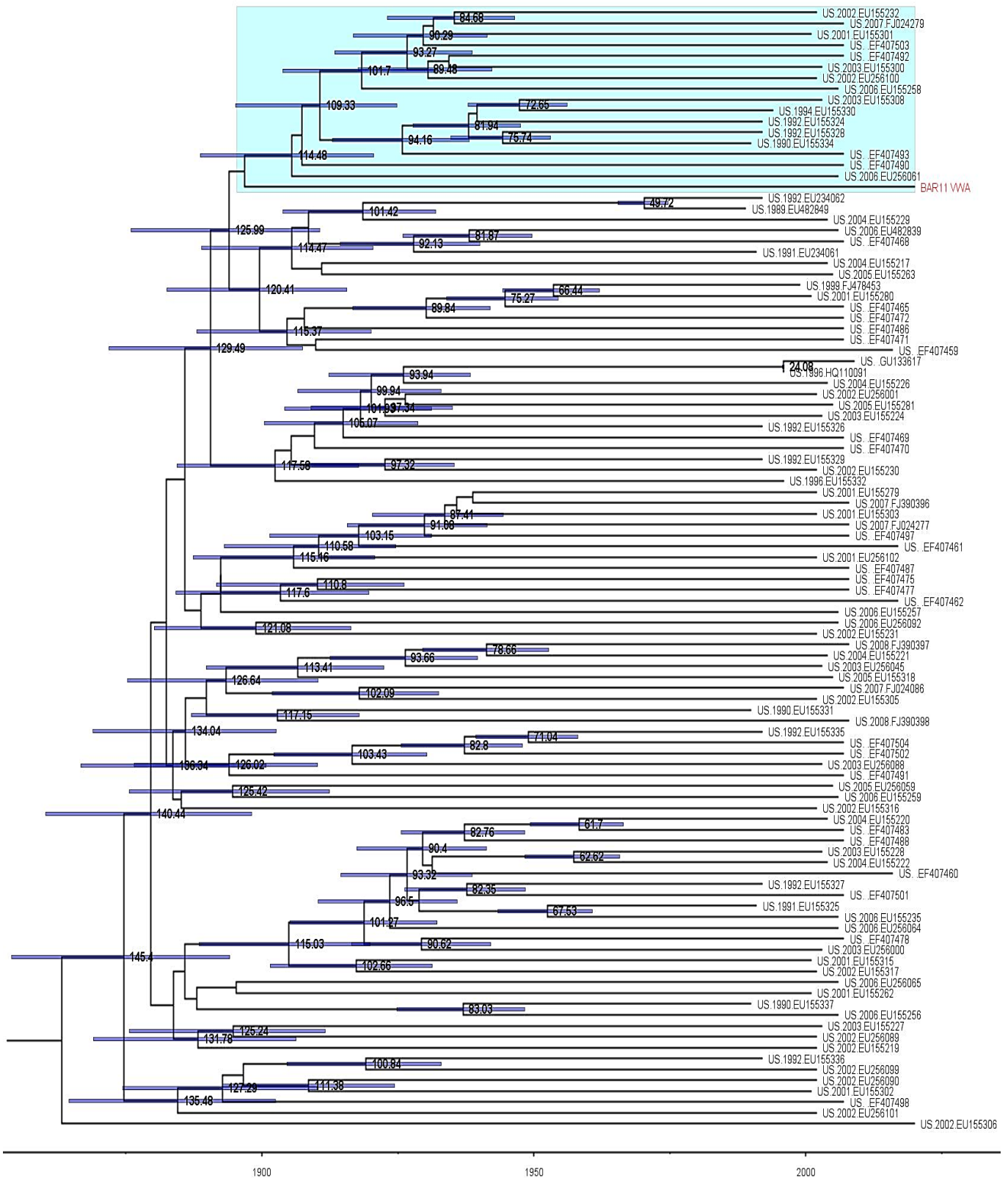
**Figura 34** Filogenia evolutiva del Gen NS5B en México. Los virus secuenciados se señalan en rojo agrupados en un clado resaltador en azul. Los nodos representan los años que han pasado desde su origen. Las ramas representan la posibilidad a posteriori. Kingman JFC (1982) Stoch Proc Appl 13, 235-248 [Constant Coalescent]. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Genetics 161, 1307-1320 [Serially Sampled Data].

### 8.2.9.2. Análisis internacional

El archivo .beauti se generó bajo los siguientes parámetros: modelo de sustitución GTR Gamma 4 obtenido previamente con IQ-Tree, reloj molecular estricto, modelo coalescente de tamaño constante para la generación del árbol posterior y el largo de la cadena Markov Monte Carlo (MCMC) de  $40 \times 10^6$ . Para la unificación de los árboles generados por BEAST 2.0 se utilizó TreeAnnotator con un límite de probabilidad a posteriori de 0.9.

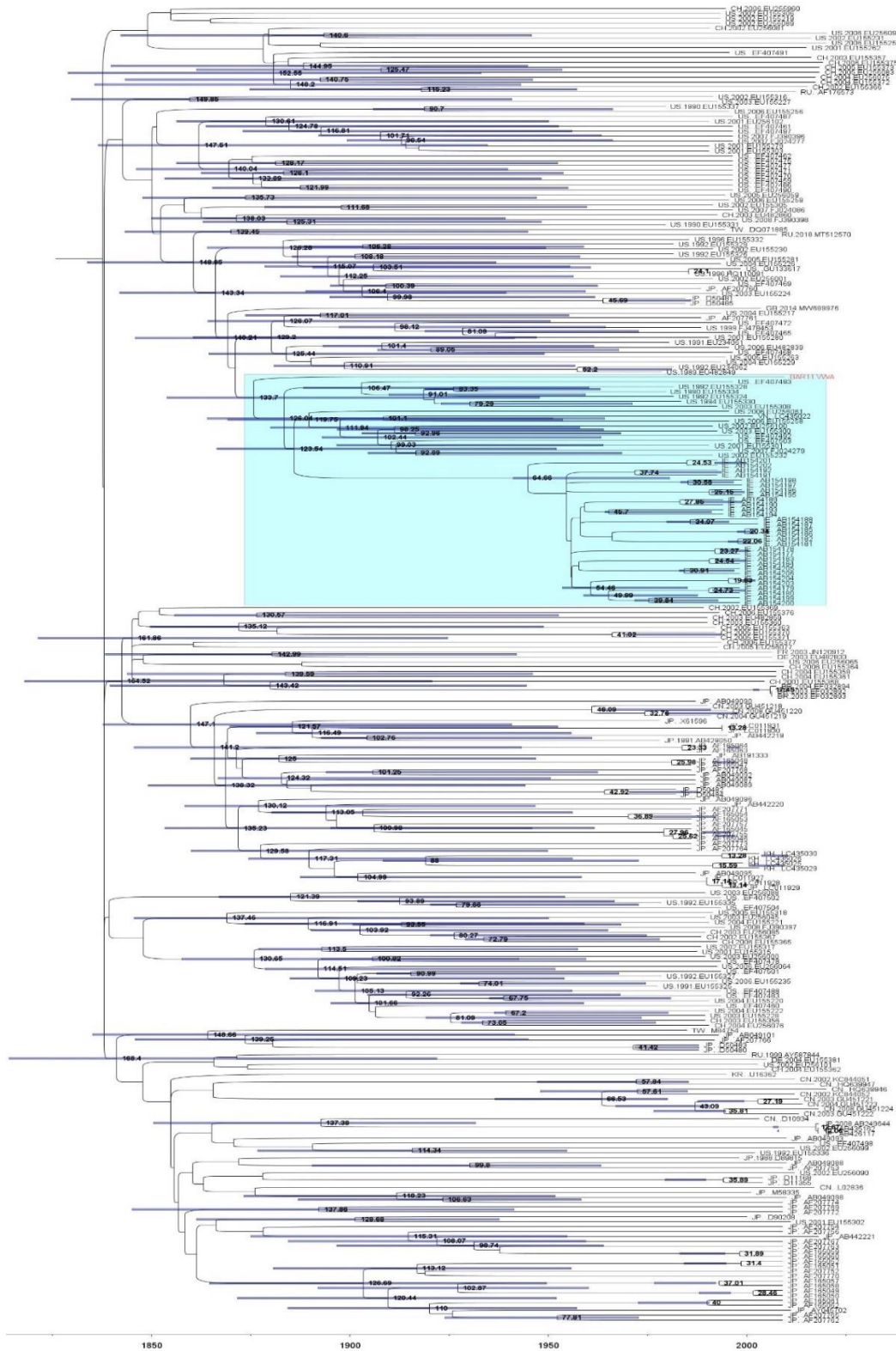
El análisis filogenético evolutivo internacional únicamente con las secuencias de Estados Unidos de América y el VHC mexicano (Figura 35) muestra una distribución de clados igual a la filogenia de máxima verosimilitud. El virus mexicano se agrupa en un clado cuyo origen tiene una posibilidad  $>0.9$  de haber aparecido hace 125 años. No obstante, el VHC mexicano no se agrupa en el mismo clúster que los virus estadounidenses.

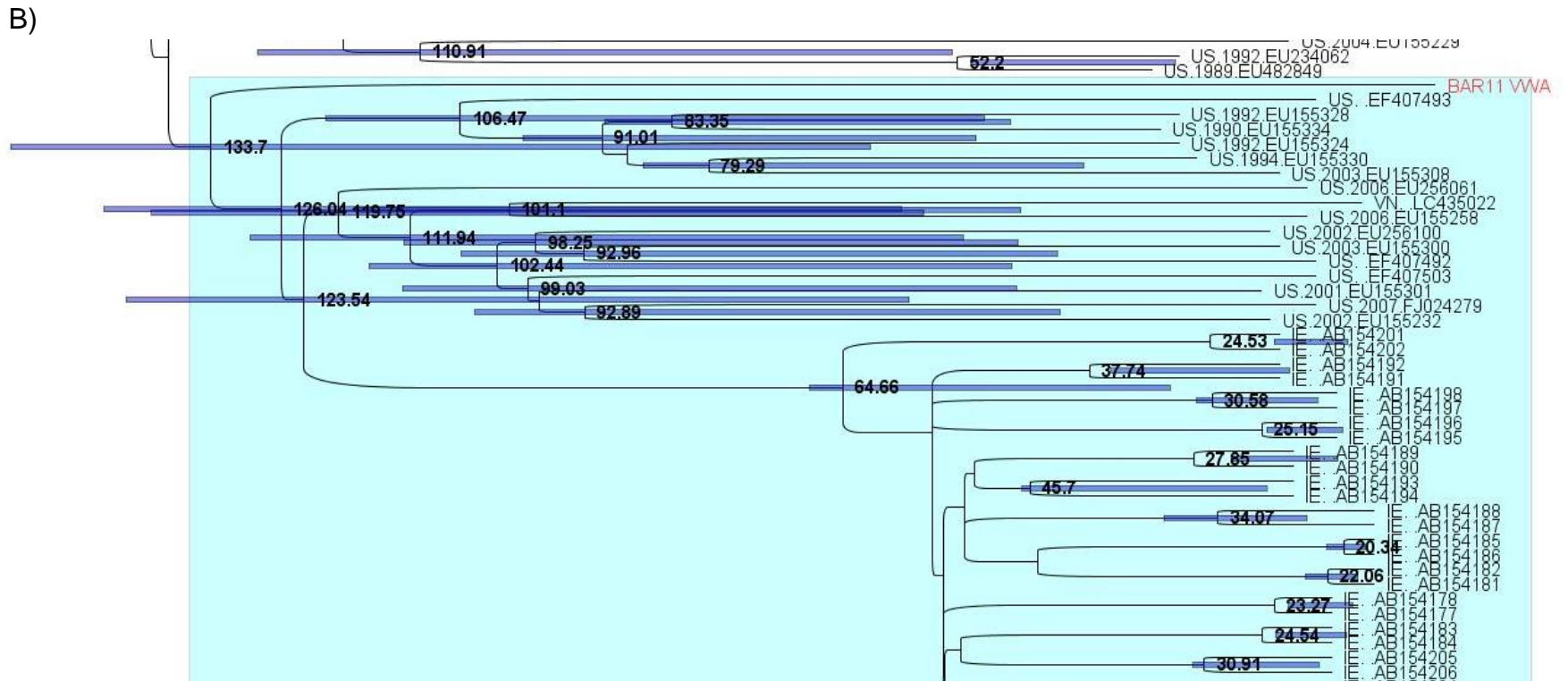
El análisis filogenético evolutivo realizado a partir de las secuencias globales y el virus mexicano secuenciado (Figura 36) muestra una distribución de clados igual a la filogenia de máxima verosimilitud. El virus mexicano no se agrupa dentro del mismo clúster de los virus estadounidenses, sin embargo, existe una posibilidad  $>0.9$  de que el VHC mexicano y los virus estadounidenses e irlandeses se hayan originado hace 133 años.



**Figura 35** Análisis evolutivo de la secuencia completa del VHC en México con las de Estados Unidos. El virus secuenciado se señalan en rojo agrupado en un clado resaltado en azul. Los nodos representan los años que han pasado desde su origen. Las barras representan los nodos cuya probabilidad a posteriori es  $>0.9$ . Kingman JFC (1982) *Stoch Proc Appl* 13, 235-248 [Constant Coalescent]. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) *Genetics* 161, 1307-1320 [Serially Sampled Data].

A)





**Figura 36 Análisis evolutivo a nivel global de la secuencia completa del VHC en México.** El virus secuenciado se señala en rojo agrupado en un clado resaltado en azul. Los nodos representan los años que han pasado desde su origen. Las barras representan los nodos cuya probabilidad a posteriori es >0.9. A) Filogenia complete. B) Acercamiento del clado en donde se agrupa el virus secuenciado Kingman JFC (1982) *Stoch Proc Appl* 13, 235-248 [Constant Coalescent]. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) *Genetics* 161, 1307-1320 [Serially Sampled Data].

## 9. Discusión

El diseño de oligonucleótidos se realizó en dos etapas. La primera etapa consistió en la generación de oligonucleótidos con únicamente 10 secuencias del VHC a nivel global. Debido a su alta variabilidad genética, lo que se buscaba fue hallar una secuencia consenso que contemplara esta característica intrínseca del virus. Sin embargo, en la PCR punto final realizada con los gradientes de temperatura, no se observó la amplificación de todos los 32 pares de oligonucleótidos.

Este primer acercamiento mostró que existían oligonucleótidos que no se encontraban dentro de regiones conservadas del virus y éstos tuvieron que ser rediseñados (Figura 19). Para el rediseño se tomaron en cuenta los huecos de la primera secuenciación que mostraba la gráfica de cobertura de Genome Detective (Figura 20). El pre-análisis filogenético de la secuencia parcial obtenida mostró una similitud con una secuencia de Estados Unidos, por lo tanto, el rediseño se realizó con nuevas secuencias procedentes únicamente de Estados Unidos de la base de datos Los Alamos, tomando como premisa que, a pesar de la heterogeneidad nucleotídica, las secuencias de los virus de los pacientes suelen mostrar clústers o similitud entre sujetos ubicados geográficamente cercanos (Minosse et al., 2021).

En el rediseño (Figura 19) se contemplaron oligonucleótidos que amplificaran regiones mayores a 400pb, con un tamaño máximo de 600pb, otros fragmentos fueron divididos en dos secciones ya que la región en donde se encontraba el primer diseño era altamente variable y, para poder amplificar completamente ese fragmento, se rediseñaron oligonucleótidos solapados que abarcaban 200pb antes y después del hueco. Finalmente, se obtuvo la amplificación de la mayoría los fragmentos rediseñados (Figura 22).

De acuerdo con los resultados experimentales, los obtenidos a partir del software Amplify 4 y su documentación, para la evaluación de la PCR *in silico*, se deben de considerar puntajes en el criterio de calidad “good”, “ok” y “moderate” para afirmar



que el diseño tiene una mayor probabilidad de tener éxito, además, realizar las evaluaciones en un mínimo de 10 secuencias aleatorias aumenta las posibilidades de que el diseño de los oligonucleótidos sea el adecuado.

En el primer diseño solo se realizó la evaluación con 3 secuencias y los oligonucleótidos en una categoría “weak” fueron modificados, aún cuando todas las evaluaciones obtuvieron el puntaje adecuado, no se observó la amplificación de todos los fragmentos de manera experimental. Un ejemplo de estas evaluaciones realizada con el segundo diseño de oligonucleótidos del genotipo 1b se observa en Anexo X.

No obstante, es importante resaltar que no debe ignorarse el criterio biológico y las propiedades intrínsecas del virus. Como ya se mencionó anteriormente, al tratarse de un virus con una alta heterogeneidad genética, cuya agrupación de clústers en diferentes individuos está relacionado con la cercanía geográfica de los portadores, tanto para el diseño como para la evaluación *in silico* deben de incluirse secuencias de la misma zona geográfica de interés, en este caso, se escogieron las secuencias de Estados Unidos. Este enfoque, en conjunto con el número de secuencias para la evaluación y los criterios de calidad, garantizaron el éxito de amplificación en el segundo diseño.

En comparación con lo que Spitz y colaboradores (2019), reportan en su trabajo de secuenciación del VHC GT 2b en Brasil, ellos realizaron PCR anidada, la cual es la metodología común utilizada para la secuenciación de este virus en otros países (Riaz et al., 2021; Uchida et al., 2019). La PCR Multiplex, utilizada en este trabajo permite la amplificación simultánea de varios fragmentos del mismo tamaño, los cuales pueden ser posteriormente ensamblados. El solapamiento de fragmentos de hasta 400pb asegura la cobertura total del genoma. Por otro lado, este método es utilizado cuando se tienen recursos limitados en el laboratorio y se necesita de un resultado rápido y efectivo (Rebelo et al., 2018).

El genotipo 1a no fue posible observarlo en el gel de electroforesis, solo se observaban dímeros. Es por ello que se decidió utilizar oligonucleótidos anti sentido para llevar a cabo la retro transcripción del ADNc como se ha reportado previamente en la literatura (Nátalia et al., 2019). Por lo tanto, se decidió utilizar el primer anti sentido que se encontraba aproximadamente en la posición nucleotídica 8300, con referencia en el esquema de la Figura 12.B, logrando la síntesis de ADNc del virus (Figura 24.C).

Al realizar el análisis bioinformático con Genome Detective, se encontró que la muestra 213 pertenece en realidad al genotipo 1b y no al genotipo 1a. Esto se confirmó mediante la herramienta de NCBI Genotyping Tool (<https://www.ncbi.nlm.nih.gov/projects/genotyping/genotype.cgi>) (Figura 26).

El principio de la prueba clínica INNO-LiPA para la genotipificación es que éste es capaz de hacer una distinción de los polimorfismos específicos de la región 5'UTR de cada genotipo (Maertens & Stuyver, 2003), sin embargo, en la Figura 26, en la región 5'UTR de la secuencia se observa la identificación del genotipo 1a, 1b y 1m, lo que puede conducir al genotipado incorrecto de la muestra. Otros autores han reportado que es más confiable subgenotipar una muestra de VHC con los genes NS3, NS5A y NS5B ya que al realizarlo con la región 5'UTR puede resultar en un genotipado distinto (Bagaglio et al., 2019). Yun et al., (1996) mencionan una buena concordancia en el subgenotipado cuando se utilizan el gen core y NS5.

Como se observa en la Figura 24, la profundidad de la cobertura aumenta si se realiza previamente la amplificación del genoma. A pesar de que la profundidad de la cobertura es más baja para las secuencias virales a las cuales no se les realizaron PCR Multiplex, éstas secuencias alcanzan un mínimo de 10x de cobertura, esta cobertura es similar a la reportada por Uchida et. al., (2019) en donde se ha utilizado PCR anidada para su amplificación.

La variabilidad genética del virus ha impulsado la investigación en medicamentos pangenotípicos que pueden brindar un mejor esquema de tratamiento

independiente al genotipo diagnosticado. Sin embargo, la frecuencia de las RAS ha aumentado con el paso de los años debido a la exposición del virus a los DAA específicamente en la proteína NS5A en el caso de genotipos como 1a, 3 y 4 (Howe et al., 2022).

En México, el tratamiento con antivirales de acción directa comenzó con el Boceprevir en el 2012, posteriormente en el 2015 con Simeprevir en combinación con Sofosbuvir o Ledipasvir y el Viekira (Muñoz, 2015), y algunas dependencias del Instituto Mexicano del Seguro Social incorporaron esquemas de tratamiento que incluían la combinación de Ombitasvir, Paritaprevir y Ritonavir.

Las secuencias virales de las muestras utilizadas en esta tesis provenían de sujetos que no habían sido sometidos a ningún esquema de tratamiento al momento de la toma de muestra. No obstante, se identificó una resistencia al antiviral Daclatasvir (inhibidor de la NS5A) en una de las secuencias (ID 261\_28, Figura 25). Debido a que se desconocen las relaciones interpersonales de los pacientes no puede asegurarse que el hallazgo de esta posible resistencia al Ombitasvir, Paritaprevir y Ledipasvir presente en las secuencias virales de Monterrey, Nuevo León, se deba al contacto con alguna persona portadora del VHC en tratamiento.

Además, se sabe que la variabilidad genética del virus está influenciada por la forma de transmisión, por ejemplo, encontrándose el genotipo 3 más frecuentemente entre sujetos que utilizan drogas inyectables (Minosse et al., 2021) y el genotipo 1b entre donadores de sangre (Zhang et al., 2020). En México, aún cuando desde 1994 se incorporaron los análisis de las donaciones de sanguíneas como medida de prevención para la transmisión de la hepatitis C, no fue hasta el 2012 que se implementó la NOM253-SSA1-2012. Esta norma únicamente establece que ante los signos y síntomas de una reacción transfusional el médico tratante o personal de salud deberá interrumpir la transfusión y no existe actualmente una unificación en las medidas para la hemovigilancia a nivel nacional (Bravo-Lindoro, 2018). Esta falta de vigilancia, la manifestación tardía de la enfermedad, la forma de transmisión y el hecho de que el paciente no haya recibido tratamiento en el momento de la toma de

muestra, sugieren que las mutaciones que llevan a una posible resistencia a los antivirales se esté dando de forma natural (Ceccherini-Silberstein et al., 2018).

La diferencia entre los parámetros “posible” y “probable” de un resultado de HCV-Glue radica en la evidencia de resistencia a ese fármaco que se ha reportado. Un resultado con una fuerte evidencia indica que la resistencia *in vitro* es un nivel  $\geq 5$  y se ha encontrado en la línea base o emergente al tratamiento *in vivo*, o en ambos. Un “probable” resultado indica un nivel *in vitro*  $\geq 50$  o encontrado en la línea de base o en el tratamiento emergente y, finalmente, un “posible” resultado indica un nivel *in vitro*  $\geq 5$  (Singer et al., 2018). Las mutaciones originadas por DAAs, como 32Del, Y93C, D168G, R155G y D168T concuerdan con lo reportado por Ceccherini-Silberstein y colaboradores. La mutación A156G al Grazoprevir no está reportada, pero si se identifica con variantes A156T/V por este mismo grupo de investigación.

Otro aspecto de importancia a considerar son las características moleculares que conllevan las infecciones por un virus de ARN, por ejemplo, el “cuello de botella”, el efecto transmitido/fundador y la deriva genética para así comprender mejor la transmisión y biología del VHC. Después de un cuello de botella los virus de ARN experimentan una rápida evolución, lo que lleva a la acumulación de mutaciones deletéreas y a la aparición de variantes raras adaptativas, las cuales pueden dominar la siguiente población (Bull et al., 2011). Estos cambios presentes en las cuasiespecies pueden explicar por qué una misma muestra secuenciada presenta discrepancias durante la identificación de RAS.

En México, se registró la prevalencia del genotipo 1b hace 10 años por Jiménez-Méndez y colaboradores. Las secuencias virales obtenidas pertenecen al genotipo 1b, se realizó el análisis filogenético de máximo verosimilitud con las secuencias parciales y la secuencia completa recién obtenidas con las secuencias mexicanas previamente reportadas, las cuales correspondían a los genes NS3, NS3/4 y NS5B.

El análisis filogenético del gen NS3 se construyó con el modelo TPM2+F+G4, lo que significa que existe una frecuencia equitativa en la aparición de las bases (AC=AT,

AG=CT, CG=GT) (TPM2); la frecuencia de las bases es empírica(F), es decir, la cuenta de las frecuencias fue proporcionada por el alineamiento y se utilizó el modelo Gamma para la tasa de heterogeneidad de los sitios nucleotídicos. En la Figura 27 se observan dos clados, uno que agrupa los virus recién secuenciados y el otro que agrupa los reportados en el 2006. Ambos clados están fuertemente soportados (Bootstrap 100). El virus B4 del 2016 muestra tener más cambios que los virus del 2020 y 2017. Un resultado interesante sobre esta filogenia es que la secuencia de los virus B6 y 429 proceden del mismo paciente y, sin embargo, de todas las iteraciones realizadas, aproximadamente en el 50% de éstas resultó en la separación en distintos grupos demostrando una baja similitud en sus secuencias. Esto podría estar relacionado con la evolución de las cuasiespecies reportada en la literatura y discutido anteriormente.

En la Figura 35 se repite el resultado con una alta probabilidad a posteriori. Además, se observa que primero se originó la secuencia B6 y 1 año y medio después se originó la secuencia B11. Estas diferencias nucleotídicas también se observan en la Tabla 11, en donde la secuencia B6 muestra tener más mutaciones con una probable y posible resistencia a los DAAs grazoprevir y paritaprevir, inhibidores de la NS3. Los virus B6 y B4 tuvieron cambios nucleotídicos aproximadamente en el año 2014 y los virus B11 y B2 en el año 2016. Ambos grupos se originaron en el 2007, es decir, 12.5 años atrás a partir del virus más reciente. El clado más antiguo parece haberse originado en 24 años atrás del 2006, es decir en los años 80's. Esto concuerda con los primeros hallazgos sobre el virus de la hepatitis C, en ese entonces conocido como hepatitis no-A no-B (Choo, 1989).

El análisis filogenético del gen NS5B se construyó con un modelo general de tiempo reversible (GTR) y se utilizó el modelo Gamma para la tasa de heterogeneidad de los sitios nucleotídicos (Figura 30). En la filogenia se observa la misma distribución que presenta el gen NS3. La diferencia es que la similitud de los VHC secuenciados está mejor suportada. Además, hay una alta confiabilidad en que los virus

secuenciados y la secuencia AB472188 sean similares o compartan el mismo origen.

En la Figura 36, se observa como hace 49 años se originaron las secuencias de los virus recién secuenciados y el virus AB472186 y hace 64 años (año 1960) se originó o se dieron los cambios en el gen NS5B que tienen los virus secuenciados y este mismo virus, siendo este resultado más confiable debido a la alta probabilidad a posteriori que muestra la rama (0.99).

El análisis filogenético de los genes NS3-4A (Figura 28) no presentan buenos valores de Bootstrap, es decir, inferiores a 0.9, (Sitnikova et al., 1995) que sustenten la similitud con los genes previamente reportados en México. Solo se observa un resultado confiable (Bootstrap de 96 y 95) en el mismo clado que agrupa a los VHC secuenciados.

El análisis del gen NS4A (Figura 29) muestra que los VHC secuenciados se agrupan con virus previamente reportados en el 2010, formando así dos clados principales, separados por el año de divergencia.

Esto puede deberse a que los genes NS3/4 y el gen NS4A presenten posibles eventos de recombinación o tienen una historia evolutiva diferente debido a un polimorfismo ancestral (Kubatko & Degnan, 2007; Mirarab et al., 2016).

Los distintos valores de Bootstrap varían dependiendo de la región del VHC que se esté analizando, especialmente en el caso los genes NS3-4a. Esta diferencia en los análisis filogenéticos se ha reportado anteriormente, sugiriendo que la región del VHC tiene una influencia en la topología del árbol (Lamoury et al., 2015).

El análisis internacional (Figura 37) muestra clúster que comprende secuencias desde los años 90's hasta el 2006, siendo éstas últimas las más cercanas al virus secuenciado. Existe una alta posibilidad de que ese grupo se haya originado hace 126 años, sin embargo, existe una baja probabilidad que este VHC comparta el mismo tiempo de origen que los VHC de Estados Unidos.

Por otro lado, en el análisis global (Figura 38) se observa una alta probabilidad de que las secuencias de Estados Unidos y la del VHC secuenciado tengan un origen en común hace 133 años. Además, se observa un sub-clado que se origina 65 años antes el cual pertenece a las secuencias de Irlanda.

## Conclusiones

- Se estandarizó la metodología y las condiciones óptimas para la secuenciación de aislados del virus de hepatitis C de pacientes infectados para mediante la tecnología de tercera generación Oxford Nanopore. Esta estandarización de la metodología permite dar el primer paso para realizar análisis filogenéticos futuros del VHC predominante en México y además puede extrapolarse para la identificación y análisis de otros genotipos circulantes.
- Se secuenció por primera vez el genoma completo circulante del VHC GT 1b aislado de sujetos infectados de Nuevo León en México. Su secuencia se registró en GenBank bajo el número de acceso ON676559.
- Se detectaron probables y posibles mutaciones asociadas resistencias a DAA de las proteínas no estructurales del VHC NS3-4A y NS5A en secuencias muestreadas del 2016 al 2020. Estas secuencias pertenecieron a las muestras de suero ID 213, ID 182, ID 261, ID 336 y ID 429. En la muestra de suero ID 213 se detectó una resistencia al daclatasvir.
- El análisis filogenético de las secuencias VHC GT 1b muestra una similitud, mayormente, con secuencias mexicanas del 2010 de la proteína NS5B y con secuencias completas de Estados Unidos del 2007 a un nivel internacional y global.



## 10. Perspectivas

Para metas a corto y mediano plazo es necesario incluir la secuenciación como parte de la investigación del VHC circulante en México para poder detectar las mutaciones asociadas a resistencia a antivirales de acción directa y de este modo brindar un mejor esquema de tratamiento al paciente.

Conjuntamente, la secuenciación de tercera generación aporta información molecular relevante para futuras investigaciones que puedan brindar conocimiento *de novo* sobre la biología molecular del virus, como lo es la evolución de las cuasiespecies.

## 11. Bibliografía

- Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K., & Miyano, M. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure*, 7(11), 1417–1426. [https://doi.org/10.1016/S0969-2126\(00\)80031-3](https://doi.org/10.1016/S0969-2126(00)80031-3)
- Albecka, A., Montserret, R., Krey, T., Tarr, A. W., Diesis, E., Ball, J. K., Descamps, V., Duverlie, G., Rey, F., Penin, F., & Dubuisson, J. (2011). Identification of new functional regions in hepatitis C virus envelope glycoprotein E2. *Journal of Virology*, 85(4), 1777–1792. <https://doi.org/10.1128/JVI.02170-10>
- Appel, N., Zayas, M., Miller, S., Krijnse-Locker, J., & Schaller, T. (2008). Essential Role of Domain III of Nonstructural Protein 5A for Hepatitis C Virus Infectious Particle Assembly. *PLoS Pathog*, 4(3), 1000035. <https://doi.org/10.1371/journal.ppat.1000035>
- Appleby, T. C., Perry, J. K., Murakami, E., Barauskas, O., Feng, J., Cho, A., Fox, D., Wetmore, D. R., McGrath, M. E., Ray, A. S., Sofia, M. J., Swaminathan, S., & Edwards, T. E. (2015). Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. *Science (New York, N.Y.)*, 347(6223), 771–775. <https://doi.org/10.1126/SCIENCE.1259210>
- Atoom, A. M., Taylor, N. G. A., & Russell, R. S. (2014). The elusive function of the hepatitis C virus p7 protein. *Virology*, 462–463(1), 377–387. <https://doi.org/10.1016/J.VIROL.2014.04.018>
- Bacon, B. R., Gordon, S. C., Lawitz, E., Marcellin, P., Vierling, J. M., Zeuzem, S., Poordad, F., Goodman, Z. D., Sings, H. L., Boparai, N., Burroughs, M., Brass, C. A., Albrecht, J. K., & Esteban, R. (2011). Boceprevir for previously treated chronic HCV genotype 1 infection. *The New England Journal of Medicine*, 364(13), 1207–1217. <https://doi.org/10.1056/NEJMOA1009482>
- Bagaglio, S., Messina, E., Hasson, H., Galli, A., Uberti-Foppa, C., & Morsica, G. (2019). Geographic Distribution of HCV-GT3 Subtypes and Naturally Occurring Resistance Associated Substitutions. *Viruses 2019, Vol. 11, Page 148, 11(2)*, 148. <https://doi.org/10.3390/V11020148>
- Bartenschlager, R., Penin, F., Lohmann, V., & André, P. (2011). Assembly of infectious hepatitis C virus particles. *Trends in Microbiology*, 19(2), 95–103. <https://doi.org/10.1016/J.TIM.2010.11.005>
- Berger, C., Romero-Brey, I., Radujkovic, D., Terreux, R., Zayas, M., Paul, D., Harak, C., Hoppe, S., Gao, M., Penin, F., Lohmann, V., & Bartenschlager, R. (2014). Daclatasvir-like inhibitors of NS5A block early biogenesis of hepatitis C virus-induced membranous replication factories, independent of RNA replication. *Gastroenterology*, 147(5), 1094–1105.e25. <https://doi.org/10.1053/J.GASTRO.2014.07.019>
- Blanchard, E., Belouzard, S., Goueslain, L., Wakita, T., Dubuisson, J., Wychowski, C., & Rouillé, Y. (2006). Hepatitis C Virus Entry Depends on Clathrin-Mediated Endocytosis. *Journal of Virology*, 80(14), 6964–6972. <https://doi.org/10.1128/JVI.00024-06>

- Borgia, S. M., Hedskog, C., Parhy, B., Hyland, R. H., Stamm, L. M., Brainard, D. M., Subramanian, M. G., McHutchison, J. G., Mo, H., Svarovskaia, E., Shafran, S. D., & Mani Subramanian, G. (2018). *Identification of a Novel Hepatitis C Virus Genotype from Punjab, India-Expanding Classification of Hepatitis C Virus into 8 Genotypes*.  
<https://doi.org/10.1093/infdis/jiy401/5047409>
- Boyce da, S. E., Tirunagari, N., Niedziela-Majka, A., Perry, J., Wong, M., Kan, E., Lagpacan, L., Barauskas, O., Hung, M., Fenaux, M., Appleby, T., Watkins, W. J., Schmitz, U., & Sakowicz, R. (n.d.). *Structural and Regulatory Elements of HCV NS5B Polymerase-b-Loop and C-Terminal Tail-Are Required for Activity of Allosteric Thumb Site II Inhibitors*.  
<https://doi.org/10.1371/journal.pone.0084808>
- Brass, V., Berke, J. M., Montserret, R., Blum, H. E., Penin, F., & Moradpour, D. (2008a). Structural determinants for membrane association and dynamic organization of the hepatitis C virus NS3-4A complex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14545–14550. <https://doi.org/10.1073/PNAS.0807298105>
- Brass, V., Berke, J. M., Montserret, R., Blum, H. E., Penin, F., & Moradpour, D. (2008b). Structural determinants for membrane association and dynamic organization of the hepatitis C virus NS3-4A complex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14545–14550. <https://doi.org/10.1073/PNAS.0807298105>
- Bravo-Lindoro, A. G. (2018, June). *Hemovigilancia y transfusión en México – Revista de Hematología*. <https://revistadehematologia.org.mx/article/hemovigilancia-y-transfusion-en-mexico/>
- Bryan-Marrugo, O. L., Ramos-Jiménez, J., Barrera-Saldaña, H., Rojas-Martínez, A., Vidaltamayo, R., & Rivas-Estilla, A. M. (2015). History and progress of antiviral drugs: From acyclovir to direct-acting antiviral agents (DAAs) for Hepatitis C. *Medicina Universitaria*, *17*(68), 165–174.  
<https://doi.org/10.1016/J.RMU.2015.05.003>
- Bukh, J. (2016). The history of hepatitis C virus (HCV): Basic research reveals unique features in phylogeny, evolution and the viral life cycle with new perspectives for epidemic control. *Journal of Hepatology*, *65*(1 Suppl), S2–S21. <https://doi.org/10.1016/J.JHEP.2016.07.035>
- Bukh, J., Purcell, R. H., & Miller, R. H. (1992). Sequence analysis of the 5' noncoding region of hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(11), 4942. <https://doi.org/10.1073/PNAS.89.11.4942>
- Bull, R. A., Luciani, F., McElroy, K., Gaudieri, S., Pham, S. T., Chopra, A., Cameron, B., Maher, L., Dore, G. J., White, P. A., & Lloyd, A. R. (2011). Sequential Bottlenecks Drive Viral Evolution in Early Acute Hepatitis C Virus Infection. *PLOS Pathogens*, *7*(9), e1002243.  
<https://doi.org/10.1371/JOURNAL.PPAT.1002243>
- C, W., P, S., & A, S. (1993). Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *Journal of Virology*, *67*(6), 3338–3344.  
<https://doi.org/10.1128/JVI.67.6.3338-3344.1993>

- Carretero Colomer, M. (2006). PEG-interferón | *Offarm. Elsevier*, 25, 116–118.  
<https://www.elsevier.es/es-revista-offarm-4-articulo-peg-interferon-13085797>
- Carvajal Carvaja, C. (2019, March). *Bilirrubina: metabolismo, pruebas de laboratorio e hiperbilirrubinemia*. [https://www.scielo.sa.cr/scielo.php?script=sci\\_arttext&pid=S1409-00152019000100073](https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1409-00152019000100073)
- Ceccherini-Silberstein, F., Cento, V., di Maio, V. C., Perno, C. F., & Craxì, A. (2018). Viral resistance in HCV infection. *Current Opinion in Virology*, 32, 115–127.  
<https://doi.org/10.1016/J.COVIRO.2018.10.005>
- Chandler, D. E., Penin, F., Schulten, K., & Chipot, C. (2012). *The p7 Protein of Hepatitis C Virus Forms Structurally Plastic, Minimalist Ion Channels*.  
<https://doi.org/10.1371/journal.pcbi.1002702>
- Chen, J. H., Zeng, Z., Zhang, X. X., Zhang, Y., Zhang, R. W., Wang, S., Wu, C. H., Yu, M., Liu, D., Xi, H. L., Zhou, X. Y., An, Y. Y., & Xu, X. Y. (2017). Efficacy and safety of combined directly acting antivirals for treatment of Chinese chronic hepatitis C patients in a real-world setting. *World Journal of Gastroenterology*, 23(22), 4072–4079. <https://doi.org/10.3748/WJG.V23.I22.4072>
- Childs-Kean, L. M., & Hong, J. (2017). *Detectable Viremia at the End of Treatment With Direct-Acting Antivirals Can Be Associated With Subsequent Clinical Cure in Patients With Chronic Hepatitis C: A Case Series*.  
<https://doi.org/10.1053/j.gastro.2017.06.062>
- Choo, Q. , K. G. , W. A. , O. L. , B. D. y H. M. (1989). Isolation of a cDNA Clone Derived from a BloodBorne Non-A, Non-B Viral Hepatitis Genome. *Proc. Natl. Acad. Sci. U.S.A*, 241(9), 13135. [www.sciencemag.org](http://www.sciencemag.org)
- Choo, Q. L., Richman, K. H., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P. J., Weiner, A. J., Bradley, D. W., Kuo, G., & Houghton, M. (1991). Genetic organization and diversity of the hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 88(6), 2451.  
<https://doi.org/10.1073/PNAS.88.6.2451>
- Coller, K. E., Heaton, N. S., Berger, K. L., Cooper, J. D., Saunders, J. L., & Randall, G. (2012a). Molecular determinants and dynamics of hepatitis C virus secretion. *PLoS Pathogens*, 8(1).  
<https://doi.org/10.1371/JOURNAL.PPAT.1002466>
- Coller, K. E., Heaton, N. S., Berger, K. L., Cooper, J. D., Saunders, J. L., & Randall, G. (2012b). Molecular determinants and dynamics of hepatitis C virus secretion. *PLoS Pathogens*, 8(1).  
<https://doi.org/10.1371/JOURNAL.PPAT.1002466>
- Corless, L., Crump, C. M., Griffin, S. D. C., & Harris, M. (2010). Vps4 and the ESCRT-III complex are required for the release of infectious hepatitis C virus particles. *The Journal of General Virology*, 91(Pt 2), 362–372. <https://doi.org/10.1099/VIR.0.017285-0>

- Crouchet, E., Baumert, T. F., & Schuster, C. (2017). Hepatitis C virus-apolipoprotein interactions: molecular mechanisms and clinical impact. *Expert Review of Proteomics*, 14(7), 593–606. <https://doi.org/10.1080/14789450.2017.1344102>
- Cummings, M. D., Lindberg, J., Lin, T. I., de Kock, H., Lenz, O., Lilja, E., Feiländer, S., Baraznenok, V., Nyström, S., Nilsson, M., Vrang, L., Edlund, M., Rosenquist, Å., Samuelsson, B., Raboisson, P., & Simmen, K. (2010). Induced-fit binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target. *Angewandte Chemie (International Ed. in English)*, 49(9), 1652–1655. <https://doi.org/10.1002/ANIE.200906696>
- Dawood, R. M., El-Meguid, M. A., Elrobe, W., Salum, G. M., Zayed, N., Mousa, S., & Medhat, E. (2021). Significance of Hereditary Hemochromatosis Gene (HFE) Mutations in Chronic Hepatitis C and Hepatocellular Carcinoma Patients in Egypt: A Pilot Study. *Asian Pacific Journal of Cancer Prevention*, 22(9), 2837–2845. <https://doi.org/10.31557/APJCP.2021.22.9.2837>
- de Beeck, A. O., Montserret, R., Duvet, S., Cocquerel, L., Cacan, R., Barberot, B., Maire, M. le, Penin, F., & Dubuisson, J. (2000). The Transmembrane Domains of Hepatitis C Virus Envelope Glycoproteins E1 and E2 Play a Major Role in Heterodimerization \*. *Journal of Biological Chemistry*, 275(40), 31428–31437. <https://doi.org/10.1074/JBC.M003003200>
- de Francesco, R., Pessi, A., & Steinkühler, C. (1999). Mechanisms of hepatitis C virus NS3 proteinase inhibitors. *Journal of Viral Hepatitis*, 6 Suppl 1(SUPPL. 1), 23–30. <https://doi.org/10.1046/J.1365-2893.1999.00002.X>
- del Carmen, M., Vela, M., & García García, F. (2015). *VARIABILIDAD GENÉTICA DEL VIRUS DE LA HEPATITIS C*.
- Denolly, S., Mialon, C., Bourlet, T., Amirache, F., Penin, F., Lindenbach, B., Boson, B., & Cosset, F. L. (2017). The amino-terminus of the hepatitis C virus (HCV) p7 viroporin and its cleavage from glycoprotein E2-p7 precursor determine specific infectivity and secretion levels of HCV particle types. *PLoS Pathogens*, 13(12). <https://doi.org/10.1371/JOURNAL.PPAT.1006774>
- DEV. (2021). *Structure - HCV Biology - Hepatitis C Online*. <https://www.hepatitisc.uw.edu/biology/structure>
- D.S. Guttman, & N.J. Provart. (2021). *Bioinformatic Methods*. <http://megasoftware.net>
- Dustin, L. B., & Rice, C. M. (2007). Flying under the radar: the immunobiology of hepatitis C. *Annual Review of Immunology*, 25, 71–99. <https://doi.org/10.1146/ANNUREV.IMMUNOL.25.022106.141602>
- el Omari, K., Iourin, O., Kadlec, J., Sutton, G., Harlos, K., Grimes, J. M., & Stuart, D. I. (2014). Unexpected structure for the N-terminal domain of hepatitis C virus envelope glycoprotein E1. *Nature Communications*, 5. <https://doi.org/10.1038/NCOMMS5874>

- Eng, F. J., El-Shamy, A., Doyle, E. H., Klepper, A., Muerhoff, A. S., & Branch, A. D. (2017). Newly discovered hepatitis C virus minicore circulate in human blood. *Hepatology Communications*, 2(1), 21–28. <https://doi.org/10.1002/HEP4.1125>
- Esser-Nobis, K., Romero-Brey, I., Ganten, T. M., Gouttenoire, J., Harak, C., Klein, R., Schemmer, P., Binder, M., Schnitzler, P., Moradpour, D., Bartenschlager, R., Polyak, S. J., Stremmel, W., Penin, F., Eisenbach, C., & Lohmann, V. (2013). Analysis of hepatitis C virus resistance to silibinin in vitro and in vivo points to a novel mechanism involving nonstructural protein 4B. *Hepatology (Baltimore, Md.)*, 57(3), 953–963. <https://doi.org/10.1002/HEP.26260>
- Evans, M. J., von Hahn, T., Tscherne, D. M., Syder, A. J., Panis, M., Wölk, B., Hatziioannou, T., McKeating, J. A., Bieniasz, P. D., & Rice, C. M. (2007). Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature*, 446(7137), 801–805. <https://doi.org/10.1038/NATURE05654>
- Farquhar, M. J., Hu, K., Harris, H. J., Davis, C., Brimacombe, C. L., Fletcher, S. J., Baumert, T. F., Rappoport, J. Z., Balfe, P., & McKeating, J. A. (2012). Hepatitis C Virus Induces CD81 and Claudin-1 Endocytosis. *Journal of Virology*, 86(8), 4305–4316. <https://doi.org/10.1128/JVI.06996-11>
- Fraser, C. S., & Doudna, J. A. (2007). Structural and mechanistic insights into hepatitis C viral translation initiation. *Nature Reviews. Microbiology*, 5(1), 29–38. <https://doi.org/10.1038/NRMICRO1558>
- Fried, M. W., Buti, M., Dore, G. J., Flisiak, R., Ferenci, P., Jacobson, I., Marcellin, P., Manns, M., Nikitin, I., Poordad, F., Sherman, M., Zeuzem, S., Scott, J., Gilles, L., Lenz, O., Peeters, M., Sekar, V., de Smedt, G., & Beumont-Mauviel, M. (2013). Once-daily simeprevir (TMC435) with pegylated interferon and ribavirin in treatment-naïve genotype 1 hepatitis C: the randomized PILLAR study. *Hepatology (Baltimore, Md.)*, 58(6), 1918–1929. <https://doi.org/10.1002/HEP.26641>
- Gastaminza, P., Cheng, G., Wieland, S., Zhong, J., Liao, W., & Chisari, F. v. (2008a). Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *Journal of Virology*, 82(5), 2120–2129. <https://doi.org/10.1128/JVI.02053-07>
- Gastaminza, P., Cheng, G., Wieland, S., Zhong, J., Liao, W., & Chisari, F. v. (2008b). Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *Journal of Virology*, 82(5), 2120–2129. <https://doi.org/10.1128/JVI.02053-07>
- Gastaminza, P., Dryden, K. A., Boyd, B., Wood, M. R., Law, M., Yeager, M., & Chisari, F. v. (2010). Ultrastructural and Biophysical Characterization of Hepatitis C Virus Particles Produced in Cell Culture. *Journal of Virology*, 84(21), 10999. <https://doi.org/10.1128/JVI.00526-10>
- Gawlik, K., & Gallay, P. A. (2014). HCV core protein and virus assembly: what we know without structures. *Immunologic Research*, 60(1), 1–10. <https://doi.org/10.1007/S12026-014-8494-3>

- Gouttenoire, J., Penin, F., & Moradpour, D. (2010). Hepatitis C virus nonstructural protein 4B: a journey into unexplored territory. *Reviews in Medical Virology*, 20(2), 117–129. <https://doi.org/10.1002/RMV.640>
- Gupta, E., Bajpai, M., & Choudhary, A. (2014). Hepatitis C virus: Screening, diagnosis, and interpretation of laboratory assays. *Asian Journal of Transfusion Science*, 8(1), 19. <https://doi.org/10.4103/0973-6247.126683>
- Haddad, J. G., Rouillé, Y., Hanouille, X., Descamps, V., Hamze, M., Dabboussi, F., Baumert, T. F., Duverlie, G., Lavie, M., & Dubuisson, J. (2017). Identification of Novel Functions for Hepatitis C Virus Envelope Glycoprotein E1 in Virus Entry and Assembly. *Journal of Virology*, 91(8). <https://doi.org/10.1128/JVI.00048-17>
- Hajime, T., Okamoto, H., Lizuka, H., Kishimoto, J., Tsuda, F., Miyawaka, Y., & Mayumi, M. (1998). The entire nucleotide sequences of three hepatitis C virus isolates in genetic groups 7–9 and comparison with those in the other eight genetic groups. *Journal of General Virology*, 79, 1847–1857.
- Han, J. H., Shyamala, V., Richman, K. H., Brauer, M. J., Irvine, B., Urdea, M. S., Tekamp-Ulson, P., Kuo, G., Choo, Q. L., & Houghton, M. (1991). Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5' untranslated region and poly(A) tails at the 3' end. *Proceedings of the National Academy of Sciences of the United States of America*, 88(5), 1711. <https://doi.org/10.1073/PNAS.88.5.1711>
- HHS, CDC, Ddid, NCHHSTP, & DVH. (2020). *Hepatitis C ¿Qué es la hepatitis?* [www.cdc.gov/hepatitis](http://www.cdc.gov/hepatitis)
- Howe, A. Y. M., Rodrigo, C., Cunningham, E. B., Douglas, M. W., Dietz, J., Grebely, J., Popping, S., Sfalcin, J. A., Parczewski, M., Sarrazin, C., de Salazar, A., Fuentes, A., Sayan, M., Quer, J., Kjellin, M., Kileng, H., Mor, O., Lennerstrand, J., Fourati, S., ... Figueruela, B. (2022). Characteristics of hepatitis C virus resistance in an international cohort after a decade of direct-acting antivirals. *JHEP Reports*, 4(5), 100462. <https://doi.org/10.1016/J.JHEPR.2022.100462>
- Jacobson, I. M., McHutchison, J. G., Dusheiko, G., di Bisceglie, A. M., Reddy, K. R., Bzowej, N. H., Marcellin, P., Muir, A. J., Ferenci, P., Flisiak, R., George, J., Rizzetto, M., Shouval, D., Sola, R., Terg, R. A., Yoshida, E. M., Adda, N., Bengtsson, L., Sankoh, A. J., ... Zeuzem, S. (2011). Telaprevir for previously untreated chronic hepatitis C virus infection. *The New England Journal of Medicine*, 364(25), 2405–2416. <https://doi.org/10.1056/NEJMOA1012912>
- Jang, J. Y., & Chung, R. T. (2010). New treatments for chronic hepatitis C. *The Korean Journal of Hepatology*, 16(3), 263–277. <https://doi.org/10.3350/KJHEP.2010.16.3.263>
- Ji, D., Chen, G. F., Wang, C., Wang, Y. D., Shao, Q., Li, B., Zhao, J., You, S. L., Hu, J. H., Liu, J. L., Niu, X. X., Chen, J., Lu, L., Wu, V., & Lau, G. (2016). Twelve-week ribavirin-free direct-acting antivirals for treatment-experienced Chinese with HCV genotype 1b infection including cirrhotic patients. *Hepatology International*, 10(5), 789–798. <https://doi.org/10.1007/S12072-016-9755-0>

- Jimenez-Mendez, R., Uribe-Salas, F., López-Guillen, P., Cisneros-Garza, L., & Castañeda-Hernandez, G. (2010). Distribution of HCV genotypes and HCV RNA viral load in different regions of Mexico. *Annals of Hepatology*, 9(1), 33–39. [https://doi.org/10.1016/S1665-2681\(19\)31676-X](https://doi.org/10.1016/S1665-2681(19)31676-X)
- Jones, C. T., Murray, C. L., Eastman, D. K., Tassello, J., & Rice, C. M. (2007). Hepatitis C Virus p7 and NS2 Proteins Are Essential for Production of Infectious Virus. *Journal of Virology*, 81(16), 8374–8383. <https://doi.org/10.1128/JVI.00690-07>
- Jones, D. M., & McLauchlan, J. (2010a). Hepatitis C virus: assembly and release of virus particles. *The Journal of Biological Chemistry*, 285(30), 22733–22739. <https://doi.org/10.1074/JBC.R110.133017>
- Jones, D. M., & McLauchlan, J. (2010b). Hepatitis C virus: assembly and release of virus particles. *The Journal of Biological Chemistry*, 285(30), 22733–22739. <https://doi.org/10.1074/JBC.R110.133017>
- K, T.-K., N, I., M, K., & A, N. (1992). Internal ribosome entry site within hepatitis C virus RNA. *Journal of Virology*, 66(3), 1476–1483. <https://doi.org/10.1128/JVI.66.3.1476-1483.1992>
- Kaito, M., Watanabe, S., Tsukiyama-Kohara, K., Yamaguchi, K., Kobayashi, Y., Konishi, M., Yokoi, M., Ishida, S., Suzuki, S., & Kohara, M. (1994). Hepatitis C virus particle detected by immunoelectron microscopic study. *The Journal of General Virology*, 75 ( Pt 7)(7), 1755–1760. <https://doi.org/10.1099/0022-1317-75-7-1755>
- Kao, C. C., Yi, G., & Huang, H. C. (2016). The core of hepatitis C virus pathogenesis. *Current Opinion in Virology*, 17, 66–73. <https://doi.org/10.1016/J.COVIRO.2016.01.009>
- Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066. <https://doi.org/10.1093/NAR/GKF436>
- Khaliq, S., Jahan, S., & Hassan, S. (2010). Hepatitis C virus p7: molecular function and importance in hepatitis C virus life cycle and potential antiviral target. *Liver International*. <https://doi.org/10.1111/j.1478-3231.2010.02442.x>
- Klein, K. C., Polyak, S. J., & Lingappa, J. R. (2004). Unique features of hepatitis C virus capsid formation revealed by de novo cell-free assembly. *Journal of Virology*, 78(17), 9257–9269. <https://doi.org/10.1128/JVI.78.17.9257-9269.2004>
- Kono, N., & Arakawa, K. (2019). Nanopore sequencing: Review of potential applications in functional genomics. *Development, Growth & Differentiation*, 61(5), 316–326. <https://doi.org/10.1111/DGD.12608>
- Koutsoudakis, G., Forns, X., & Pérez-del-Pulgar, S. (2013). Biología molecular aplicada del virus de la hepatitis C. *Gastroenterología y Hepatología*, 36(4), 280–293. <https://doi.org/10.1016/J.GASTROHEP.2012.11.005>



- Koutsoudakis, G., Kaul, A., Steinmann, E., Kallis, S., Lohmann, V., Pietschmann, T., & Bartenschlager, R. (2006). Characterization of the Early Steps of Hepatitis C Virus Infection by Using Luciferase Reporter Viruses. *Journal of Virology*, *80*(11), 5308–5320. <https://doi.org/10.1128/JVI.02460-05>
- Krüger, K., Krauth, C., Rossol, S., Mauss, S., Boeker, K. H. W., Müller, T., Klinker, H., Pathil, A., Heyne, R., Stahmeyer, J. T., Buggisch, P., Günther, R., Hüppe, D., Herold, C., Weber, A., Teuber, G. de, Ullrich, R., John, C., Link, R., ... Zeuzem, S. (2019). Outcomes and costs of treating hepatitis C patients with second-generation direct-acting antivirals: results from the German Hepatitis C-Registry. *European Journal of Gastroenterology & Hepatology*, *31*(2), 230–240. <https://doi.org/10.1097/MEG.0000000000001283>
- Kubatko, L. S., & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, *56*(1), 17–24. <https://doi.org/10.1080/10635150601146041>
- Kuiken, C., & Simmonds, P. (2009). Nomenclature and Numbering of the Hepatitis C Virus. *Methods in Molecular Biology (Clifton, N.J.)*, *510*, 33–53. [https://doi.org/10.1007/978-1-59745-394-3\\_4](https://doi.org/10.1007/978-1-59745-394-3_4)
- Lamoury, F. M. J., Jacka, B., Bartlett, S., Bull, R. A., Wong, A., Amin, J., Schinkel, J., Poon, A. F., Matthews, G. v., Grebely, J., Dore, G. J., & Applegate, T. L. (2015). The Influence of Hepatitis C Virus Genetic Region on Phylogenetic Clustering Analysis. *PLOS ONE*, *10*(7), e0131437. <https://doi.org/10.1371/JOURNAL.PONE.0131437>
- Lapa, D., Garbuglia, A. R., Capobianchi, M. R., & Porto, P. del. (2019). Hepatitis C Virus Genetic Variability, Human Immune Response, and Genome Polymorphisms: Which Is the Interplay? *Cells*, *8*(4), 305. <https://doi.org/10.3390/CELLS8040305>
- Lavillette, D., Bartosch, B., Nourrisson, D., Verney, G., Cosset, F. L., Penin, F., & Pécheur, E. I. (2006). Hepatitis C Virus Glycoproteins Mediate Low pH-dependent Membrane Fusion with Liposomes \*. *Journal of Biological Chemistry*, *281*(7), 3909–3917. <https://doi.org/10.1074/JBC.M509747200>
- Lee, J.-Y., Acosta, E. G., Stoeck, I. K., Long, G., Hiet, M.-S., Mueller, B., Fackler, O. T., Kallis, S., & Bartenschlager, R. (2014). Apolipoprotein E likely contributes to a maturation step of infectious hepatitis C virus particles and interacts with viral envelope glycoproteins. *Journal of Virology*, *88*(21), 12422–12437. <https://doi.org/10.1128/JVI.01660-14>
- Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., & Weber, P. C. (1999). Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nature Structural Biology*, *6*(10), 937–943. <https://doi.org/10.1038/13305>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. <https://doi.org/10.1093/BIOINFORMATICS/BTP324>

- Lin, C., Wu, J. W., Hsiao, K., & Su, M. S. (1997). The hepatitis C virus NS4A protein: interactions with the NS4B and NS5A proteins. *Journal of Virology*, *71*(9), 6465–6471. <https://doi.org/10.1128/JVI.71.9.6465-6471.1997>
- Lindenbach, B. D. (2013). Virion assembly and release. *Current Topics in Microbiology and Immunology*, *369*, 199–218. [https://doi.org/10.1007/978-3-642-27340-7\\_8](https://doi.org/10.1007/978-3-642-27340-7_8)
- Lindenbach, B. D., Prágai, B. M., Montserret, R., Beran, R. K. F., Pyle, A. M., Penin, F., & Rice, C. M. (2007). The C Terminus of Hepatitis C Virus NS4A Encodes an Electrostatic Switch That Regulates NS5A Hyperphosphorylation and Viral Replication. *Journal of Virology*, *81*(17), 8905–8918. <https://doi.org/10.1128/JVI.00937-07>
- Lindenbach, B. D., & Rice, C. M. (2005). *INSIGHT REVIEW Unravelling hepatitis C virus replication from genome to function Initial studies: HCV translation and polyprotein processing*. <https://doi.org/10.1038/nature04077>
- Lindenbach, B. D., & Rice, C. M. (2013). The ins and outs of hepatitis C virus entry and assembly. *Nature Reviews. Microbiology*, *11*(10), 688–700. <https://doi.org/10.1038/NRMICRO3098>
- Lohmann, V., Bartenschlager, R., & Lohmann, V. (2013). *Hepatitis C Virus RNA Replication*. [https://doi.org/10.1007/978-3-642-27340-7\\_7](https://doi.org/10.1007/978-3-642-27340-7_7)
- Lohmann, V., Roos, A., Körner, F., Koch, J. O., & Bartenschlager, R. (2000). Biochemical and structural analysis of the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. *Journal of Viral Hepatitis*, *7*(3), 167–174. <https://doi.org/10.1046/J.1365-2893.2000.00218.X>
- Lorenz, I. C. (2010). The Hepatitis C Virus Nonstructural Protein 2 (NS2): An Up-and-Coming Antiviral Drug Target. *Viruses*, *2*, 1635–1646. <https://doi.org/10.3390/v2081635>
- Lorenz, I. C., Marcotrigiano, J., Dentzer, T. G., & Rice, C. M. (2006). Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature*, *442*(7104), 831–835. <https://doi.org/10.1038/NATURE04975>
- Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T., & Hostomska, Z. (1996). The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell*, *87*(2), 331–342. [https://doi.org/10.1016/S0092-8674\(00\)81350-1](https://doi.org/10.1016/S0092-8674(00)81350-1)
- Love, R. A., Parge, H. E., Yu, X., Hickey, M. J., Diehl, W., Gao, J., Wriggers, H., Ekker, A., Wang, L., Thomson, J. A., Dragovich, P. S., & Fuhrman, S. A. (2003). Crystallographic Identification of a Noncompetitive Inhibitor Binding Site on the Hepatitis C Virus NS5B RNA Polymerase Enzyme. *Journal of Virology*, *77*(13), 7575–7581. <https://doi.org/10.1128/JVI.77.13.7575-7581.2003>
- Lozano Mérida, J. A. (2005). *Hepatitis C | Offarm*. <https://www.elsevier.es/es-revista-offarm-4-articulo-hepatitis-c-13073445>

- Lundin, M., Lindström, H., Grönwall, C., & Persson, M. A. A. (2006). Dual topology of the processed hepatitis C virus protein NS4B is influenced by the NS5A protein. *The Journal of General Virology*, 87(Pt 11), 3263–3272. <https://doi.org/10.1099/VIR.0.82211-0>
- Maertens, G., & Stuyver, L. (2003). HCV Genotyping by the Line Probe Assay INNO-LiPA HCV II. *Hepatitis C Protocols*, 183–198. <https://doi.org/10.1385/0-89603-521-2:183>
- Mazumdar, B., Banerjee, A., Meyer, K., & Ray, R. (2011). Hepatitis C virus E1 envelope glycoprotein interacts with apolipoproteins in facilitating entry into hepatocytes. *Hepatology (Baltimore, Md.)*, 54(4), 1149–1156. <https://doi.org/10.1002/HEP.24523>
- Minosse, C., Salichos, L., Taibi, C., Luzzitelli, I., Nardozi, D., Capobianchi, M. R., D’Offizi, G., McPhee, F., & Garbuglia, A. R. (2021). Phylogenetic and Phylodynamic Analyses of HCV Strains Circulating among Patients Using Injectable Drugs in Central Italy. *Microorganisms*, 9(7), 1432. <https://doi.org/10.3390/MICROORGANISMS9071432>
- Mirarab, S., Bayzid, M. S., & Warnow, T. (2016). Evaluating Summary Methods for Multilocus Species Tree Estimation in the Presence of Incomplete Lineage Sorting. *Systematic Biology*, 65(3), 366–380. <https://doi.org/10.1093/SYSBIO/SYU063>
- Morikawa, K., Lange, C. M., Gouttenoire, J., Meylan, E., Brass, V., Penin, F., & Moradpour, D. (2011a). Nonstructural protein 3-4A: the Swiss army knife of hepatitis C virus. *Journal of Viral Hepatitis*, 18(5), 305–315. <https://doi.org/10.1111/J.1365-2893.2011.01451.X>
- Morikawa, K., Lange, C. M., Gouttenoire, J., Meylan, E., Brass, V., Penin, F., & Moradpour, D. (2011b). Nonstructural protein 3-4A: the Swiss army knife of hepatitis C virus. *Journal of Viral Hepatitis*, 18(5), 305–315. <https://doi.org/10.1111/J.1365-2893.2011.01451.X>
- Moustafa, R. I., Haddad, J. G., Linna, L., Hanouille, X., Descamps, V., Mesalam, A. A., Baumert, T. F., Duverlie, G., Meuleman, P., Dubuisson, J., & Lavie, M. (2018). Functional Study of the C-Terminal Part of the Hepatitis C Virus E1 Ectodomain. *Journal of Virology*, 92(20). <https://doi.org/10.1128/JVI.00939-18>
- Murphy, D. G., Sablon, E., Chamberland, J., Fournier, E., Dandavino, R., & Tremblay, C. L. (2015). Hepatitis C virus genotype 7, a new genotype originating from Central Africa. *Journal of Clinical Microbiology*, 53(3), 967–972. <https://doi.org/10.1128/JCM.02831-14>
- Nátalia, S., JJ, B., KM, do Ó., CE, B.-M., & NM, A. (2019). The First Complete Genome Sequences of Hepatitis C Virus Subtype 2b from Latin America: Molecular Characterization and Phylogeographic Analysis. *Viruses*, 11(11). <https://doi.org/10.3390/V11111000>
- Niepmann, M. (2013). Hepatitis C virus RNA translation. *Current Topics in Microbiology and Immunology*, 369, 143–166. [https://doi.org/10.1007/978-3-642-27340-7\\_6](https://doi.org/10.1007/978-3-642-27340-7_6)
- Niepmann, M., Shalamova, L. A., Gerresheim, G. K., & Rossbach, O. (2018). Signals Involved in Regulation of Hepatitis C Virus RNA Genome Translation and Replication. *Frontiers in Microbiology*, 9(MAR). <https://doi.org/10.3389/FMICB.2018.00395>

- Oxford Technologies. (2021). *DNA sequencing | Oxford Nanopore Technologies*.  
<https://nanoporetech.com/applications/dna-nanopore-sequencing>
- Ouyang, B., Xie, S., Berardi, M. J., Zhao, X., Dev, J., Yu, W., Sun, B., & Chou, J. J. (2013). Unusual architecture of the p7 channel from hepatitis C virus. *Nature*, *498*(7455), 521–525.  
<https://doi.org/10.1038/NATURE12283>
- Pár, A., Pár, G., Tornai, I., Szalay, F., Várszegi, D., Fráter, E., Papp, M., Lengyel, G., Fehér, J., Varga, M., Gervain, J., Schuller, J., Nemes, Z., Péterfi, Z., Tusnádi, A., Hunyady, B., Haragh, A., Szinku, Z., Vincze, Á., ... Melegh, B. (2014). IL28B and IL10R –1087 polymorphisms are protective for chronic genotype 1 HCV infection and predictors of response to interferon-based therapy in an East-Central European cohort. *BMC Research Notes*, *7*(1), 12.  
<https://doi.org/10.1186/1756-0500-7-12>
- Paul, D., Hoppe, S., Saher, G., Krijnse-Locker, J., & Bartenschlager, R. (2013). Morphological and biochemical characterization of the membranous hepatitis C virus replication compartment. *Journal of Virology*, *87*(19), 10612–10627. <https://doi.org/10.1128/JVI.01370-13>
- Paul, D., Madan, V., & Bartenschlager, R. (2014). Hepatitis C virus RNA replication and assembly: living on the fat of the land. *Cell Host & Microbe*, *16*(5), 569–579.  
<https://doi.org/10.1016/J.CHOM.2014.10.008>
- Pawlotsky, J. M., Negro, F., Aghemo, A., Berenguer, M., Dalgard, O., Dusheiko, G., Marra, F., Puoti, M., & Wedemeyer, H. (2020). EASL recommendations on treatment of hepatitis C: Final update of the series☆. *Journal of Hepatology*, *73*(5), 1170–1218.  
<https://doi.org/10.1016/J.JHEP.2020.08.018/ATTACHMENT/AC59514F-6EFF-4B59-83AF-1F5F4FBCDE61/MMC1.PDF>
- Phan, T., Kohlway, A., Dimberu, P., Pyle, A. M., & Lindenbach, B. D. (2011). The acidic domain of hepatitis C virus NS4A contributes to RNA replication and virus particle assembly. *Journal of Virology*, *85*(3), 1193–1204. <https://doi.org/10.1128/JVI.01889-10>
- Poordad, F., & Dieterich, D. (2012). Treating hepatitis C: current standard of care and emerging direct-acting antiviral agents. *Journal of Viral Hepatitis*, *19*(7), 449–464.  
<https://doi.org/10.1111/J.1365-2893.2012.01617.X>
- Popescu, C.-I., Callens, N., Trinel, D., Roingeard, P., Moradpour, D., Ronique Descamps, V., Duverlie, G., Penin, F., Hé Liot, L., Rouillé, Y., & Dubuisson, J. (2011). *NS2 Protein of Hepatitis C Virus Interacts with Structural and Non-Structural Proteins towards Virus Assembly*.  
<https://doi.org/10.1371/journal.ppat.1001278>
- Prentoe, J., Velázquez-Moctezuma, R., Fong, S. K. H., Law, M., & Bukh, J. (2016). Hypervariable region 1 shielding of hepatitis C virus is a main contributor to genotypic differences in neutralization sensitivity. *Hepatology (Baltimore, Md.)*, *64*(6), 1881–1892.  
<https://doi.org/10.1002/HEP.28705>

- Raney, K. D., Sharma, S. D., Moustafa, I. M., & Cameron, C. E. (2010). Hepatitis C virus non-structural protein 3 (HCV NS3): A multifunctional antiviral target. *Journal of Biological Chemistry*, 285(30), 22725–22731. <https://doi.org/10.1074/JBC.R110.125294>
- Rebelo, A. R., Bortolaia, V., Kjeldgaard, J. S., Pedersen, S. K., Leekitcharoenphon, P., Hansen, I. M., Guerra, B., Malorny, B., Borowiak, M., Hammerl, J. A., Battisti, A., Franco, A., Alba, P., Perrin-Guyomard, A., Granier, S. A., de Frutos, C., Escobar, Malhotra-Kumar, S., Villa, L., ... Hendriksen, R. S. (2018). Multiplex PCR for detection of plasmid-mediated colistin resistance determinants, *mcr-1*, *mcr-2*, *mcr-3*, *mcr-4* and *mcr-5* for surveillance purposes. *Eurosurveillance*, 23(6), 17–00672. <https://doi.org/10.2807/1560-7917.ES.2018.23.6.17-00672/CITE/PLAINTEXT>
- Riaz, N., Leung, P., Barton, K., Smith, M. A., Carswell, S., Bull, R., Lloyd, A. R., & Rodrigo, C. (2021). Adaptation of Oxford Nanopore technology for hepatitis C whole genome sequencing and identification of within-host viral variants. *BMC Genomics*, 22(1). <https://doi.org/10.1186/s12864-021-07460-1>
- Rodrigo, C., Eltahla, A. A., Bull, R. A., Luciani, F., Grebely, J., Dore, G. J., Applegate, T., Page, K., Bruneau, J., Morris, M. D., Cox, A. L., Osburn, W., Kim, A. Y., Shoukry, N. H., Lauer, G. M., Maher, L., Schinkel, J., Prins, M., Hellard, M., & Lloyd, A. R. (2017). Phylogenetic analysis of full-length, early infection, hepatitis C virus genomes among people with intravenous drug use: the InC3 Study. *Journal of Viral Hepatitis*, 24(1), 43. <https://doi.org/10.1111/JVH.12616>
- Romero-Brey, I., & Bartenschlager, R. (2014). Membranous replication factories induced by plus-strand RNA viruses. *Viruses*, 6(7), 2826–2857. <https://doi.org/10.3390/V6072826>
- Romero-Brey, I., Merz, A., Chiramel, A., Lee, J. Y., Chlanda, P., Haselman, U., Santarella-Mellwig, R., Habermann, A., Hoppe, S., Kallis, S., Walther, P., Antony, C., Krijnse-Locker, J., & Bartenschlager, R. (2012). Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS Pathogens*, 8(12). <https://doi.org/10.1371/JOURNAL.PPAT.1003056>
- Sabahi, A., Uprichard, S. L., Wimley, W. C., Dash, S., & Garry, R. F. (2014). Unexpected structural features of the hepatitis C virus envelope protein 2 ectodomain. *Journal of Virology*, 88(18), 10280–10288. <https://doi.org/10.1128/JVI.00874-14>
- Salvatierra, K. (2017). Epidemiología molecular del virus de la hepatitis C. *Infectio*, 21(2), 117–125. <https://doi.org/10.22354/IN.V21I2.655>
- Sharma, N. R., Mateu, G., Dreux, M., Grakoui, A., Cosset, F. L., & Melikyan, G. B. (2011). Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *Journal of Biological Chemistry*, 286(35), 30361–30376. <https://doi.org/10.1074/JBC.M111.263350>
- Shi, G., & Suzuki, T. (2018). Molecular Basis of Encapsidation of Hepatitis C Virus Genome. *Frontiers in Microbiology*, 9(MAR). <https://doi.org/10.3389/FMICB.2018.00396>

- Simmonds, P. (1995). Variability of hepatitis C virus. *Hepatology*, 21(2), 570–583. <https://doi.org/10.1002/HEP.1840210243>
- Simmonds, P., Bukh, J., Combet, C., Deléage, G., Enomoto, N., Feinstone, S., Halfon, P., Inchauspé, G., Kuiken, C., Maertens, G., Mizokami, M., Murphy, D. G., Okamoto, H., Pawlotsky, J. M., Penin, F., Sablon, E., Shin-I, T., Stuyver, L. J., Thiel, H. J., ... Widell, A. (2005). Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology*, 42(4), 962–973. <https://doi.org/10.1002/HEP.20819>
- Singer, J. B., Thomson, E. C., McLauchlan, J., Hughes, J., & Gifford, R. J. (2018). GLUE: a flexible software system for virus sequence data. *BMC Bioinformatics*, 19(1). <https://doi.org/10.1186/S12859-018-2459-9>
- Sitnikova, T., Rzhetsky, A., & Nei, M. (1995). Interior-branch and bootstrap tests of phylogenetic trees. *Molecular Biology and Evolution*, 12(2), 319–333. <https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A040205>
- Smith, D. B., Bukh, J., Kuiken, C., Muerhoff, A. S., Rice, C. M., Stapleton, J. T., & Simmonds, P. (2014). Expanded Classification of Hepatitis C Virus Into 7 Genotypes and 67 Subtypes: Updated Criteria and Genotype Assignment Web Resource. *Hepatology (Baltimore, Md.)*, 59(1), 318. <https://doi.org/10.1002/HEP.26744>
- Soriano, V., Vispo, E., de Mendoza, C., Labarga, P., Fernandez-Montero, J. V., Poveda, E., Treviño, A., & Barreiro, P. (2013). Hepatitis C therapy with HCV NS5B polymerase inhibitors. *Expert Opinion on Pharmacotherapy*, 14(9), 1161–1170. <https://doi.org/10.1517/14656566.2013.795543>
- Stewart, H., Bingham, R. J., White, S. J., Dykeman, E. C., Zothner, C., Tuplin, A. K., Stockley, P. G., Twarock, R., & Harris, M. (2016). Identification of novel RNA secondary structures within the hepatitis C virus genome reveals a cooperative involvement in genome packaging. *Scientific Reports*, 6. <https://doi.org/10.1038/SREP22952>
- Strosberg, A. D., Kota, S., Takahashi, V., Snyder, J. K., & Mousseau, G. (2010). Core as a novel viral target for hepatitis C drugs. *Viruses*, 2(8), 1734–1751. <https://doi.org/10.3390/V2081734>
- Sulkowski, M. S., Asselah, T., Lalezari, J., Ferenci, P., Fainboim, H., Leggett, B., Bessone, F., Mauss, S., Heo, J., Datsenko, Y., Stern, J. O., Kukulj, G., Scherer, J., Nehmiz, G., Steinmann, G. G., & Böcher, W. O. (2013). Faldaprevir combined with pegylated interferon alfa-2a and ribavirin in treatment-naïve patients with chronic genotype 1 HCV: SILEN-C1 trial. *Hepatology (Baltimore, Md.)*, 57(6), 2143–2154. <https://doi.org/10.1002/HEP.26276>
- Syed, G. H., Khan, M., Yang, S., & Siddiqui, A. (2017). Hepatitis C Virus Lipoviroparticles Assemble in the Endoplasmic Reticulum (ER) and Bud off from the ER to the Golgi Compartment in COPII Vesicles. *Journal of Virology*, 91(15). <https://doi.org/10.1128/JVI.00499-17>
- Thomas, D. L., Thio, C. L., Martin, M. P., Qi, Y., Ge, D., Ohuigin, C., Kidd, J., Kidd, K., Khakoo, S. I., Alexander, G., Goedert, J. J., Kirk, G. D., Donfield, S. M., Rosen, H. R., Tobler, L. H., Busch, M.

- P., McHutchison, J. G., Goldstein, D. B., & Carrington, M. (2009). Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature*, *461*(7265), 798–801. <https://doi.org/10.1038/NATURE08463>
- Uchida, Y., Kouyama, J. ichi, Naiki, K., Uemura, H., Tsuji, S., Sugawara, K., Nakao, M., Motoya, D., Nakayama, N., Imai, Y., Tomiya, T., & Mochida, S. (2019). A case of genotype-3b hepatitis C virus in which the whole genome was successfully analyzed using third-generation nanopore sequencing. *Hepatology Research*, *49*(9), 1083–1087. <https://doi.org/10.1111/HEPR.13339>
- V.F., M., & A.López, S. R. (2015, October). *Pruebas de función hepática: B, AST, ALT, FA y GGT*. [https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S1130-01082015001000017](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1130-01082015001000017)
- Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology* *2021 39:11*, *39*(11), 1348–1365. <https://doi.org/10.1038/s41587-021-01108-x>
- Yi, M., Ma, Y., Yates, J., & Lemon, S. M. (2009). trans-Complementation of an NS2 Defect in a Late Step in Hepatitis C Virus (HCV) Particle Assembly and Maturation. *PLoS Pathog*, *5*(5), 1000403. <https://doi.org/10.1371/journal.ppat.1000403>
- Yun, Z., Lara, C., Johansson, B., Lorenzana De Rivera, I., & Sönnnerborg, A. (1996). Discrepancy of hepatitis C virus genotypes as determined by phylogenetic analysis of partial NS5 and core sequences. *Journal of Medical Virology*, *49*(3), 155–160. [https://doi.org/10.1002/\(sici\)1096-9071\(199607\)49:3<155::aid-jmv1>3.0.co;2-3](https://doi.org/10.1002/(sici)1096-9071(199607)49:3<155::aid-jmv1>3.0.co;2-3)
- Zayas, M., Long, G., Madan, V., & Bartenschlager, R. (2016). Coordination of Hepatitis C Virus Assembly by Distinct Regulatory Regions in Nonstructural Protein 5A. *PLoS Pathogens*, *12*(1). <https://doi.org/10.1371/JOURNAL.PPAT.1005376>
- Zayas, M., Madan, V., & Bartenschlager, R. (2016). *Coordination of Hepatitis C Virus Assembly by Distinct Regulatory Regions in Nonstructural Protein 5A*. <https://doi.org/10.1371/journal.ppat.1005376>
- Zhang, Y., Gao, Z., Wang, S., Liu, J., Paul, N., He, T., Liu, C., Zhang, H., Lv, Y., Cao, R., Mao, W., Wan, J., Ma, H., Huang, M., Liu, Y., Wang, J., Liao, P., Zeng, P., He, M., & Shan, H. (2020). Hepatitis C virus genotype/subtype distribution and evolution among Chinese blood donors: Revealing recent viral expansion. *PLOS ONE*, *15*(7), e0235612. <https://doi.org/10.1371/JOURNAL.PONE.0235612>
- Zulay Sulbarán, M., Lérida Montañó, Yoneira Sulbarán, Carmen Luisa Loureiro, Carmen Rosa Flores, Yurviris Farías, Antonio Maldonado, Genny Guillen, Héctor Rangel, & Flor Pujol. (2016, April). *Caracterización molecular del virus de la Hepatitis C en pacientes remitidos al Laboratorio Regional de Salud Pública del Hospital Universitario Antonio Patricio de Alcalá Cumaná Venezuela*. [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S0535-51332016000100003](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0535-51332016000100003)

## 12. Anexos

### 12.1. Anexo I Reactivos para la extracción de ARN

Tabla 1. Materiales requeridos para aislamiento de ARN, DNA y proteínas

Ítem	Fuente
Equipo	
Centrífuga y rotor capaz de alcanzar las 12, 000 x g y 4°C	MLS
Tubos	
Tubos de polipropileno para microcentrífuga	MLS
Reactivos	
Cloroformo	MLS

Tabla 2. Materiales requeridos para la extracción de ARN

Ítem	Fuente
Equipo	
Baño de agua o termo bloc a 55-60°C	MLS
Reactivos	
Isopropanol	MLS
Etanol al 75%	MLS
Agua libre de ARNsas	MLS



## 12.2. Anexo II Síntesis de ANDc

Tabla 1. Materiales requeridos para síntesis de ANDc con M-MLV RT

Reacción de 20uL

Reactivo	μL
5X-First Strand Buffer	4
dNTP Mix 10mM	1
0.1M DTT Mix	1
M-MLV RT (200 U/ μL)	1
Agua libre de nucleasas	a 12
Plantilla de RNA	5-10 μL

Memoria de cálculo para preparar random oligonucleótidos 50-250 ng/μL

$$C1V1 = C2V2$$

$$V1 = \frac{C2V2}{C1}$$

$$V1 = \frac{(100ng/\mu L)(50\mu L)}{3000}$$

$$V1 = 5\mu L \text{ de Random Primers}$$

$$Vol. H2O = 5\mu L \text{ de Random Primers} - 50\mu L = 95\mu L$$

### 12.3. Anexo III Reactivos y condiciones para los ciclos de PCR punto final

Tabla 1. Reactivos requeridos para PCR

Reactivo	$\mu\text{L}$
10x Taq Buffer	2.5
dNTP Mix 10mM	0.50
25mM $\text{MgCl}_2$	0.75
Oligonucleótido F 10 $\mu\text{M}$	1.25
Oligonucleótido R 10 $\mu\text{M}$	1.25
Taq DNA Polimerasa (200U/ $\mu\text{L}$ )	0.4
Agua libre de nucleasas	16.53
Plantilla de ADNc	2

Tabla 2. Condiciones para PCR punto final

Paso	Temperatura $^{\circ}\text{C}$	Tiempo	# de ciclos
<b>Desnaturalización inicial</b>	95	3 min	1
<b>Desnaturalización</b>	95	30s	30
<b>Alineamiento</b>	TM	30s	
<b>Extensión</b>	72	1 min	
<b>Extensión final</b>	72	10 min	1

#### 12.4. Anexo IV Preparación de Buffer TBE 5x

Reactivo	Stock/L
<b>Tris base</b>	54g
<b>Ácido bórico</b>	27.5g
<b>EDTA 0.5M pH 8</b>	20ml

\*Sambrooke

## 12.5. Anexo V Condiciones de temperatura para PCR Multiplex

Se realizó a 35 ciclos

Paso	Temperatura (°C)	Tiempo
Pre-Desnaturalización	98	30s
Desnaturalización	98	15s
Alineamiento	65	5 min
Almacenamiento	4	Infinito

## 12.6. Anexo VI Primer diseño de oligonucleótidos GT 1a Primal

### Scheme

Nombre	Secuencia	Tamaño	%gc	TM (use 65)
Scheme_1a_1_LEFT	AGCCATAGTGGTCTGCGGAA	20	55	61.28
Scheme_1a_1_RIGHT	ATAGGCTGACGTCTACCTCGAG	22	54.55	60.99
Scheme_1a_2_LEFT	GTTTACTTGTTGCCGCGCAG	20	55	61.01
Scheme_1a_2_RIGHT	TTCCCTGTTGCATAGTTCACGC	22	50	61.63
Scheme_1a_3_LEFT	ATGGGGTACATTCCGCTCGT	20	55	61.35
Scheme_1a_3_RIGHT	CTCCCGACAAGCAGATCGAT	20	55	59.63
Scheme_1a_4_LEFT	CCCTAACTCGAGTATTGTGTACGA	24	45.83	59.98
Scheme_1a_4_RIGHT	ATGTCCAAGATGGCTTGTGGGA	22	50	62.15
Scheme_1a_5_LEFT	ACTGTTACACCTTCTCTCCAGG	22	54.55	61.54
Scheme_1a_5_RIGHT	CAGTTCAAGGCCGTGCTATTGA	22	50	61.37
Scheme_1a_6_LEFT	GTGCTGTTGCTATTTGCCGG	20	55	60.47
Scheme_1a_6_RIGHT	AGTGAAGCAATACACCGGGC	20	55	60.69
Scheme_1a_7_LEFT	TGGCCCGGTGTATTGCTTCA	20	55	62.15
Scheme_1a_7_RIGHT	CCTCCACGTACATCCTGACTT	22	54.55	61.66
Scheme_1a_8_LEFT	CTGATTGTTTCCGCAAGCATCC	22	50	60.91
Scheme_1a_8_RIGHT	TCCCACTTAATGGCCAGGA	20	55	60.57
Scheme_1a_9_LEFT	TGCTGCTGTCTACCACACAGT	21	52.38	61.75
Scheme_1a_9_RIGHT	AGGAAAGGCCACATCCCGTA	20	55	60.93
Scheme_1a_10_LEFT	TTCCTTCTGCTTGCAGACGC	20	55	61.25
Scheme_1a_10_RIGHT	GCTTCTGCTCTGGTCAGGAA	20	55	59.69
Scheme_1a_11_LEFT	CTTGTCGGGTTAATGGCGCT	20	55	61.04
Scheme_1a_11_RIGHT	AGAGGAGTGAGATGGTTATAAACATA GG	28	39.29	60.56
Scheme_1a_12_LEFT	GGAGGCCATTACGTGCAAATG	21	52.38	60.3
Scheme_1a_12_RIGHT	TTGACACAATCTGGACCTCACC	22	50	60.41
Scheme_1a_13_LEFT	CCGGGACAAAAACCAAGTGA	21	52.38	61.13
Scheme_1a_13_RIGHT	CACACCGCAGCCCTGAATAT	20	55	60.48
Scheme_1a_14_LEFT	CTTACTTGAAAGGCTCCTCGGG	22	54.55	60.86
Scheme_1a_14_RIGHT	GGCTGCCAGTGGTAATTGTTCT	22	50	61.32
Scheme_1a_15_LEFT	GCTACAAGGTGCTAGTGCTCAA	22	50	60.79
Scheme_1a_15_RIGHT	GCCTTGCCGTAAAAAGGGATCT	22	50	61.39

<b>Scheme_1a_16_LEFT</b>	CTGTCCTTGACCAAGCAGAGAC	22	54.55	61.05
<b>Scheme_1a_16_RIGHT</b>	GGTCAAGGCTGAAATCGACTGT	22	50	61.05
<b>Scheme_1a_17_LEFT</b>	ATGCTCTCATGACCGGCTTTAC	22	50	60.92
<b>Scheme_1a_17_RIGHT</b>	TGGGATAGGAAGTGGGCATCTA	22	50	60.28
<b>Scheme_1a_18_LEFT</b>	GGCTGTGCTTGGTATGAGCT	20	55	60.41
<b>Scheme_1a_18_RIGHT</b>	AGGTCAGCCGACATGCATGT	20	55	62.21
<b>Scheme_1a_19_LEFT</b>	TGCTATACAGACTGGGCGCT	20	55	61.06
<b>Scheme_1a_19_RIGHT</b>	CGAGTTTTTGCCAGTTGGTCTG	22	50	60.72
<b>Scheme_1a_20_LEFT</b>	CAGTTCAAGCAGAAGGCCCT	20	55	60.27
<b>Scheme_1a_20_RIGHT</b>	CCATACCCTGCAAGAATGTCCA	22	50	60.54
<b>Scheme_1a_21_LEFT</b>	GCCAAACCCTCCTCTTCAACAT	22	50	61
<b>Scheme_1a_21_RIGHT</b>	CACATAGTGC GTGGGGAAA	20	55	60.34
<b>Scheme_1a_22_LEFT</b>	GCAGCAATACTACGCCGACA	20	55	60.82
<b>Scheme_1a_22_RIGHT</b>	GTGGCAGCGAGTGTGCATAAT	21	52.38	61.71
<b>Scheme_1a_23_LEFT</b>	AAAGCCAAGCTCATGCCACAA	21	47.62	61.48
<b>Scheme_1a_23_RIGHT</b>	CACCCCGTCCAATTCAGTGA	20	55	59.98
<b>Scheme_1a_24_LEFT</b>	TTCCACTACGTGACGGGTATGA	22	50	61.06
<b>Scheme_1a_24_RIGHT</b>	AGGAGGTTAGACTCTATGAGCTCG	24	50	61.07
<b>Scheme_1a_25_LEFT</b>	ATGCTCACTGATCCCTCCCA	20	55	60.34
<b>Scheme_1a_25_RIGHT</b>	GGTGGTTCGTAGTCAGGCTTTT	22	50	60.99
<b>Scheme_1a_26_LEFT</b>	CGTGAAAAAGCCTGACTACGA	22	50	61.04
<b>Scheme_1a_26_RIGHT</b>	GTGATAAGTGCGCCTGTCCA	20	55	60.41
<b>Scheme_1a_27_LEFT</b>	TGAGTCCTGTTCTTCCATGCC	21	52.38	60.1
<b>Scheme_1a_27_RIGHT</b>	TTGGATTTGGCTGAGTGTGGG	21	52.38	60.93
<b>Scheme_1a_28_LEFT</b>	AGCGGCGTCAAAAGTGAAGG	20	55	61.52
<b>Scheme_1a_28_RIGHT</b>	AGGAATTC AACCCGCTGTCCT	21	52.38	61.84
<b>Scheme_1a_29_LEFT</b>	TGCGAGAAGATGGCCCTGTA	20	55	60.99
<b>Scheme_1a_29_RIGHT</b>	GGGCCTTGATGTAGCAAGTGAG	22	54.55	61.44
<b>Scheme_1a_30_LEFT</b>	GCCCTCTTACCAATTCAAGGGG	22	54.55	61.13
<b>Scheme_1a_30_RIGHT</b>	GAGTGTGTCTTGCTGTCTCCCA	22	54.55	62.36
<b>Scheme_1a_31_LEFT</b>	ATCATGCTCCTCCAACGTGTC	21	52.38	60.51
<b>Scheme_1a_31_RIGHT</b>	CGGTGTCTCCAAGTTCGCAA	20	55	60.89
<b>Scheme_1a_32_LEFT</b>	AGAACCACTGGATCTACCTCCA	22	50	60.14
<b>Scheme_1a_32_RIGHT</b>	TTGGGGAGGAGGTAGATGCCTA	22	54.55	61.98

## 12.7. Anexo VII Primer diseño de oligonucleótidos GT 1b Primal

### Scheme

Nombre	Secuencia	Tamaño	%gc	TM (use 65)
Scheme_1b_1_LEFT	TAGTATGAGTGTCGTGCAGCCT	22	50	61.39
Scheme_1b_1_RIGHT	ACCGCTCGGAAGTCTTCCTA	20	55	60.34
Scheme_1b_2_LEFT	TCGTTGGTGGAGTTTACCTGTT	22	45.45	60.01
Scheme_1b_2_RIGHT	TAATTCACGCCGTCTCCAGAA	22	50	61.91
Scheme_1b_3_LEFT	GGTCATCGATACCCTCACATGC	22	54.55	61.05
Scheme_1b_3_RIGHT	CGTCGTATTGTCGTAGTGGGGA	22	54.55	62
Scheme_1b_4_LEFT	TCACGAACGACTGCTCCAAC	20	55	60.61
Scheme_1b_4_RIGHT	GATCCGGAGTAACTGCGACA	20	55	59.56
Scheme_1b_5_LEFT	TATTCCATGGTGGGGAAGTGGG	22	54.55	61.96
Scheme_1b_5_RIGHT	AGGTGCGTAGTGCCAGCAATA	21	52.38	62.22
Scheme_1b_6_LEFT	CACATCAACAGAAGTGCCTGA	22	50	60.99
Scheme_1b_6_RIGHT	TTCATCCATGTACAGCCGAACC	22	50	60.86
Scheme_1b_7_LEFT	GACGTGCTGCTCCTCAACAA	20	55	60.9
Scheme_1b_7_RIGHT	GGCTGAGCTCTGATCTATCCCT	22	54.55	61.01
Scheme_1b_8_LEFT	ACTACCCCTGCACTGTCAACTT	22	50	61.74
Scheme_1b_8_RIGHT	ACCACCAAGTTCTCTAAGGCGG	22	54.55	62.44
Scheme_1b_9_LEFT	GCAATACCTGTACGGCATAGGG	22	54.55	61.05
Scheme_1b_9_RIGHT	TAGTGCGGTGACAAGGTCAG	20	55	59.7
Scheme_1b_10_LEFT	TATGCCATGGACCGGGAGAT	20	55	60.49
Scheme_1b_10_RIGHT	AGCTTCATGAGAGCCATTTGGA	22	45.45	60.21
Scheme_1b_11_LEFT	CTGGGGGCCATTATGTCCAAAT	22	50	60.88
Scheme_1b_11_RIGHT	GAAACCACTTGAACCTCCCCCT	22	54.55	62.35
Scheme_1b_12_LEFT	AGTAGAGCCCGTCGTCTTCT	20	55	60.05
Scheme_1b_12_RIGHT	AGGTCCTGGTCTACATTGGTGT	22	50	60.95
Scheme_1b_13_LEFT	TTAACGGCGTGTGTTGGACTG	21	52.38	61.5
Scheme_1b_13_RIGHT	TGGTTTCCATAGACTCAACGGG	22	50	60.21
Scheme_1b_14_LEFT	TGTGGCAGCTCGGACCTTTA	20	55	61.49
Scheme_1b_14_RIGHT	TCAGGACGAGTACCTTGACCC	22	54.55	61.06
Scheme_1b_15_LEFT	TTCCAAGTGGCCCATCTACAC	21	52.38	60.1

<b>Scheme_1b_15_RIGHT</b>	ATGTTGGGATGTGGCACGGT	20	55	62.44
<b>Scheme_1b_16_LEFT</b>	ATCATAATATGTGATGAGTGCCACTC	26	38.46	59.56
<b>Scheme_1b_16_RIGHT</b>	CCGGTAAAGCCCGTCATTAGAG	22	54.55	60.98
<b>Scheme_1b_17_LEFT</b>	TAGCGTATTACCGGGTCTTGA	22	50	60.87
<b>Scheme_1b_17_RIGHT</b>	GCAACCCTGGTGTGTTTAGGTA	22	50	60.67
<b>Scheme_1b_18_LEFT</b>	TGTGTGAGTGCTATGACGCG	20	55	60.75
<b>Scheme_1b_18_RIGHT</b>	CAGCCGACATACATGCCATGAT	22	50	61.31
<b>Scheme_1b_19_LEFT</b>	ACACCCCTGCTGTATAGGCTAG	22	54.55	61.21
<b>Scheme_1b_19_RIGHT</b>	AAGGGTTCGCCACTTGGACT	20	55	61.72
<b>Scheme_1b_20_LEFT</b>	TACATCGAACAGGGAATGCAGC	22	50	61.18
<b>Scheme_1b_20_RIGHT</b>	ATGTCCACAAGCACCTTCCCA	21	52.38	62.06
<b>Scheme_1b_21_LEFT</b>	ACCCAAAACACCCTCCTGTTTA	22	45.45	59.94
<b>Scheme_1b_21_RIGHT</b>	CGTCGCTTTCAGGCACATAGT	21	52.38	61.1
<b>Scheme_1b_22_LEFT</b>	CAGTGGATGAACCGGCTGATAG	22	54.55	60.98
<b>Scheme_1b_22_RIGHT</b>	GGAACCGTTTTTGACATGTCCG	22	50	60.78
<b>Scheme_1b_23_LEFT</b>	TCATGTCAACGTGGGTACAAGG	22	50	60.73
<b>Scheme_1b_23_RIGHT</b>	TCCCGTAGAAGAGGCTTGCA	20	55	60.92
<b>Scheme_1b_24_LEFT</b>	TAAAATGCCCATGCCAGGTTCC	22	50	61.66
<b>Scheme_1b_24_RIGHT</b>	TCTCTGACTCCACTCGGGTGAT	22	54.55	62.47
<b>Scheme_1b_25_LEFT</b>	TCTTCAGCCAGCCAGTTGTCT	21	52.38	61.76
<b>Scheme_1b_25_RIGHT</b>	TCTTTCTCCGTGGAGGTGGT	20	55	60.49
<b>Scheme_1b_26_LEFT</b>	CAGAGATCCTGCGGAAATCCAG	22	54.55	60.99
<b>Scheme_1b_26_RIGHT</b>	GACCAAGACCCGTCGCTAAGAT	22	54.55	62.28
<b>Scheme_1b_27_LEFT</b>	CGATCTTAGCGACGGGTCTT	20	55	59.64
<b>Scheme_1b_27_RIGHT</b>	GTCCTTTGCCCATAGCCAAA	21	52.38	61
<b>Scheme_1b_28_LEFT</b>	CCACAGTTAAGGCCAAGCTTCT	22	50	61
<b>Scheme_1b_28_RIGHT</b>	CTTTGATTTCCAGGCATTCACCA	23	43.48	60
<b>Scheme_1b_29_LEFT</b>	AAGCCAGCTCGCCTTATCGT	20	55	61.98
<b>Scheme_1b_29_RIGHT</b>	GGTATTACCGCAGCTAGTCGTC	22	54.55	60.78
<b>Scheme_1b_30_LEFT</b>	AGACAGGCCATAAGGTCGCT	20	55	61
<b>Scheme_1b_30_RIGHT</b>	AGTACACCCTTTTGCCAGATGC	22	50	61.32
<b>Scheme_1b_31_LEFT</b>	CCGGAATACGACTTGGAGTTGA	22	50	60.27
<b>Scheme_1b_31_RIGHT</b>	AGCCACCCTATTGATCTCACCT	22	50	60.82
<b>Scheme_1b_32_LEFT</b>	CCATTGAGCCACTTGACCTACC	22	54.55	61.12
<b>Scheme_1b_32_RIGHT</b>	AGATGCCTACCCCTACAGAAAGT	23	47.83	60.83



## 12.8. Anexo VIII Oligonucleótidos 1a sintetizados por Oligo T4

<b>Primer diseño</b>		<b>Segundo Diseño</b>	
1a_1_LEFT	AGCCATAGTGGTCTGCGGAA	1a_1_LEFT	CTCAATGCCTGGAGATTTGGGC
1a_1_RIGHT	ATAGGCTGACGTCTACCTCGAG	1a_1_RIGHT	CCTCATTACCATAGAGGGGCCA
1a_2_LEFT	GTTTACTTGTTGCCGCGCAG	1a_2_LEFT	TAGACGRCAGCCTATCCCCAAG
1a_2_RIGHT	TTCCCTGTTGCATAGTTCACGC	1a_2_RIGHT	TTGCGCACTTGGTAGGCTGA
1a_3_LEFT	ATGGGGTACATTCCGCTCGT	1a_3_LEFT	AGTTCTGGAAGACGGCGTGA
1a_3_RIGHT	CTCCCGACAAGCAGATCGAT	1a_3_RIGHT	AAGAAAGACAGACCCGCACA
1a_4_LEFT	CCCTAACTCGAGTATTGTGTACGA	1a_4_LEFT	TGTGTCCCTTGCGTTCGTGA
1a_4_RIGHT	ATGTCCAAGATGGCTTGTGGGA	1a_4_RIGHT	ATGGAGAAATACGCTATGCCCG
1a_5_LEFT	ACTGTTACCTTCTCTCCCAGG	1a_5_LEFT	CGCATGGCATGGGATATGATGA
1a_5_RIGHT	CAGTTCAAGGCCGTGCTATT	1a_5_Rmodi	GGACAGCCTGAAGAGTTGAARTTG
1a_6_LEFT	GTGCTGTTGCTATTTGCCGG	1a_6_LEFT	ACATCAATAGCACGGCCTTGAA
1a_6_RIGHT	AGTGAAGCAATACACCGGGC	1a_6_RIGHT	TTCATCCAGGTACAACCGAACC
1a_7_LEFT	TGGCCCGGTGTATTGCTTCA	1a_7_LEFT	GTGTATTGCTTCACTCCCAGCC
1a_7_RIGHT	CCTCCCACGTACATCCTGACTT	1a_7_RIGHT	GCTCGACCCCTCCTATGTACAT
1a_8_LEFT	CTGATTGTTTCCGCAAGCATCC	1a_8_LEFT	GAAGCCACATACTCTCGGTGTG
1a_8_RIGHT	TCCCCTTAATGGCCCAGGA	1a_8_RIGHT	GAGAACGACGTACTIONCCACTT
1a_9_LEFT	TGCTGCTGTCTACCACACAGT	1a_9_LEFTM	ATCCACCTCCACCAGAACATTG
1a_9_RIGHT	AGGAAAGGCCACATCCCGTA	1a_9_RIGHT	AGCGCCATTAACCCGACAAG
1a_10_LEFT	TTCCTTCTGCTTGCAGACGC	1a_10_LEFT	TCTACGCCTTCTACGGGATGTG
1a_10_RIGHT	GCTTCTGCTCTGGTCAGGAA	1a_10_RIGHT	CCTTGAACGCGCACGAAGTA
1a_11_LEFT	CTTGTGCGGTTAATGGCGCT	1a_11_LEFT	TTGACATCACCAAAGTCTGCT
1a_11_RIGHT	CCGGAGAGGAGTGAGATGGT	1a_11_RIGHT	TTGGAGACCATTCCATCGGC
1a_12_LEFT	GGAGGCCATTACGTGCAAATG	1a_12_LEFT	GTGACATCATCAACGGCTTGC
1a_12_RIGHT	TTGACACAATCTGGACCTCACC	1a_12_RIGHT	CAGGTGCAGGGTGTCAATGA
1a_13_LEFT	CCGGGACAAAAACCAAGTGGA	1a_13_LEFTB	GCGTACGCCAGCAGACAAGG
1a_13_RIGHT	CACACCGCAGCCCTGAATAT	1a_13_RIGHT	AACACCGGGGATCTCATGGTT
1a_14_LEFT	CTTACTTGAAAGGCTCCTCGGG	1a_14_LEFT	ATATTCAGGGCCGCAAGTGTG
1a_14_RIGHT	GGCTGCCAGTGGTAATTGTTCT	1a_14_RIGHT	TTATGTCATAAGCACCCCCTGAG
1a_15_LEFT	GCTACAAGGTGCTAGTGCTCAA	1a_15_LEFT	ATCCTAACATCAGGACTGGGGT
1a_15_RIGHT	GCCTTGCCGTAAAAAGGGATCT	1a_15_RIGHT	GAGCTCGTCGCACTTCTTCTTT
1a_16_LEFT	CTGTCTTGACCAAGCAGAGAC	1a_16_LEFTB	GAGGTTGCTCTGTCCACCACCG
1a_16_RIGHT	GGGTCAAGTCGAAATCGACTGT	1a_16_RIGHTB	GCTCCCCCGGTGCCACAAATC

1a\_17\_LEFT ATGCTCTCATGACCGGCTTTAC  
 1a\_17\_RIGHT TGGGATAGGAAGTGGGCATCTA  
 1a\_18\_LEFT GGCTGTGCTTGGTATGAGCT  
 1a\_18\_RIGHT AGGTCAGCCGACATGCATGT  
 1a\_19\_LEFT GCTATACAGACTGGGCGC  
 1a\_19\_RIGHT CGAGTTTTTGCCAGTTGGTCTG  
 1a\_20\_LEFT CAGTTCAAGCAGAAGGCCCT  
 1a\_20\_RIGHT CCATACCCTGCAAGAATGTCCA  
 1a\_21\_LEFT GCCAAACCCTCCTCTTCAACAT  
 1a\_21\_RIGHT CACATAGTGCCTGGGGGAAA  
 1a\_22\_LEFT GCAGCAATACTGCGCCGACA  
 1a\_22\_RIGHT GGCAACGAATGTGCATGACGC  
 1a\_23\_LEFT AAAGCCAAGCTCATGCCACAA  
 1a\_23\_RIGHT CACCCCGTCCAATTCAGTGA  
 1a\_24\_LEFT TTCCACTACGTGACGGGTATGA  
 1a\_24\_RIGHT AGGAGGTTAGACTCTATGAGCTCG  
 1a\_25\_LEFT ATGCTCACTGATCCCTCCCA  
 1a\_25\_RIGHT GGTGGTTCGTAGTCAGGCTTTT  
 1a\_26\_LEFT CGTGAAAAAGCCTGACTACGA  
 1a\_26\_RIGHT GTGATAAGTGCGCCTGTCCA  
 1a\_27\_LEFT TGAGTCCTGTTCTTCCATGCC  
 1a\_27\_RIGHT TTGGATTTGGCTGAGTGTGGG  
 1a\_28\_LEFT AGCGGCGTCAAAAGTGAAGG  
 1a\_28\_RIGHT AGGAATTCAACCCGCTGTCCT  
 1a\_29\_LEFT TGCAGAAAGATGGCCCTGTA  
 1a\_29\_RIGHT GGGCCTTGATGTAGCAAGTGAG  
 1a\_30\_LEFT GCCCTTACCAATTCAAGGGG  
 1a\_30\_RIGHT GAGTGTGTCTTGCTGTCTCCCA  
 1a\_31\_LEFT ATCATGCTCCTCCAACGTGTC  
 1a\_31\_RIGHT CGGTGTCTCCAAGTTCGCAA  
 1a\_32\_LEFT AGAACCACTGGATCTACCTCCA  
 1a\_32\_RIGHT TTGGGGAGGAGGTAGATGCCTA

1a\_17\_LNuevo ACCATTGAGACAACCACGCT  
 1a\_17\_R CTTCCACATCTGGTCCCACGAT  
 1a\_18\_LEFT TTTCTTACCTGGTAGCGTACC  
 1a\_18\_RIGHT TGCTGAGAGCACTCTTCCATCT  
 1a\_19\_LEFT GGTCATAGTGGGCAGGATTGTC  
 1a\_19\_RIGHT CCCCCTAAGATGTTGAAGAGGAG  
 1a\_20\_LEFT CCGCCATTGCTTCATTGATGG  
 1a\_20\_RIGHT CGGTTTCATCCATTGCACTGC  
 1a\_21\_LEFT AATCTACTGCCCGCCATCCT  
 1a\_21\_RNUEVO CAGTGGCAGCGAGTGTGCAT  
 1a\_22\_LEFT GAGGTGCTGAGCGATTTCAAGA  
 1a\_22\_RIGHT CTGTGAAAAATTTCGGGCGATGG  
 1a\_23\_LEFT TGGGGGACTTCCACTACGTA  
 1a\_23\_RIGHT AGGTTAGCCTCTATGAGCTCGG  
 1a\_24\_LEFT GTGTTGACGTCCATGCTCACT  
 1a\_24\_RIGHT GGTGGTTCGTAGTCAGGCTTTT  
 1a\_25\_LEFT2 AGTGGTGATTCTGGACTCCT  
 1a\_25\_RIGHT CACTACTGACCGTCGACCATGA  
 1a\_26\_LEFT TCTTGTTCTTCCATGCCCCC  
 1a\_26\_RIGHT CGTCAGGCTACAAGCTTCCCT  
 1a\_27\_LEFT CCAAAGGCAGAAGAAAGTCACAT  
 1a\_27\_RIGHT GAGCTTGCTAACCACGTCGTAC  
 1a\_28\_LEFT TTTTCTGCGTTCAGCCTGAGAA  
 1a\_28\_RIGHT CGCAGTTCTCCCCTCTTGAATT  
 1a\_29\_LEFT TACCAATGTTGTGACCTGGACC  
 1a\_29\_RIGHT GCGACTGACACGTTGGAAGA  
 1a\_30\_LEFT ACGGAGGCTATGACCAGGTA  
 1a\_30\_RIGHT AACTGTGGAGTGAAAATGCGCT  
 1a\_31\_LEFT GCCTGCTACTCCATAGAACCAC  
 1a\_31\_RIGHT AGCAGGAGTAGGCAAAACCAG

## 12.9. Anexo IX Oligonucleótidos 1b sintetizados por Oligo T4

	Primer Diseño		Segundo Diseño
		1b_1LEFTN	GTGAGGAACTACTGTCTTCA
1b_1_LEFT	TAGTATGAGTGTCTGCAGCCT	1b_1RIGHTN	GTCCTGGCAATTCCGGTGTACT
1b_1_RIGHT	ACCGCTCGGAAGTCTTCCTA	1b_2L	TCGTTGGTGGAGTTTACCTGTT
1b_2_LEFT	TCGTTGGTGGAGTTTACCTGTT	1b_2R	CACGTTGCGCACTTCATAGGC
1b_2_RIGHT	TAATTCACGCCGTCCTCCAGAA	1b_3LN	ATGAGGCAGCGGACATGAT
1b_3_LEFT	GGTCATCGATACCCTCACATGC	1b_3.5RN	GGCAGTCCTGTTGATGTGCC
1b_3_RIGHT	CGTCGTATTGTCGTAGTGGGGA	1b_7_LN	CTGTACATGGATGAATAGCA
1b_4_LEFT	TCACGAACGACTGCTCCAAC	1b_7_RIGHTN	CGAGCCTATGCTCCACGCC
1b_4_RIGHT	GATCCGGAGTAACTGCGACA	1b_9_LEFTN	GCTRAGCCCGTTGCTGCTGT
1b_5_LEFT	TATTCATGGTGGGGAAGTGGG	1b_9_RIGHTN	GGGGACCAGCCTGCCCTTGA
1b_5_RIGHT	AGGTGCGTAGTGCCAGCAATA	1b_9.5_LEFTN	CATCAAGGGCAGGCTGGTCCC
1b_6_LEFT	CACATCAACAGAACTGCCCTGA	1b_9.5_RIGHTN	TGAGCGGACCGAGTATGGCGA
1b_6_RIGHT	TTCATCCATGTACAGCCGAACC	1b_10_LEFTN	TACGCCATGGACCGGGAGAT
1b_7_LEFT	GACGTGCTGCTCCTCAACAA	1b_10_RIGHTN	AACATGCATGCACGAATGAGCC
1b_7_RIGHT	GGCTGAGCTCTGATCTATCCCT	1b_11_LEFTN	CTGGGGGRCATTATGTCCAAAT
1b_8_LEFT	ACTACCCCTGCACTGTCAACTT	1b_11_RIGHTN	TCTGTTGGGAGTAGGCCGTGAT
1b_8_RIGHT	ACCACCAAGTTCTCTAAGGCGG	1b_15_LEFTN	TTCCAAGTGGCCARCTACAC
1b_9_LEFT	GCAATACCTGTACGGCATAGGG	1b_15_RIGHTN	ATYTTDGYYTGTGGCACGGT
1b_9_RIGHT	AGTATGGTGCCAAGGTCAAGAAT	1b_16_LEFTN	GCCTCACAGGCCGGGACAAG
1b_10_LEFT	TTATGCCATGGACCGGGAGAT	1b_16_RIGHTN	AAGATGCCACACYGCGTGCCC
1b_10_RIGHT	AGCCTCATGAGAGCCATTTGGAC	1b_17.5_LEFTN	GCAGGACTGGTAGGGGCAGG
1b_11_LEFT	CTGGGGGCCATTATGTCCAAAT	1b_17.5_RIGHTN	CTGGCAYACGGGCAACCCTG
1b_11_RIGHT	GAAACCACTTGAACCTCCCCCT	1b_21_LEFTN	ACCCAABRACCCTCCTGTTTA
1b_12_LEFT	AGTAGAGCCCGTCTCTTCT	1b_21_RIGHTN	CGTCGCTCTCAGGCACATAGT
1b_12_RIGHT	AGGTCCTGGTCTACATTGGTGT	1b_22_LEFTN	GCTCTGGCCGCGTARTGCCTG
1b_13_LEFT	TTAACGGCGTGTGTTGGACTG	1b_22_RIGHTN	CGGCGAGCTGCATTCCCTGRTC
1b_13_RIGHT	TGGTTTCCATAGACTCAACGGG	1b_24_LEFTN	CGGTTCCATGAGGATCGTTGGGC
1b_14_LEFT	TGTGGCAGCTCGGACCTTTA	1b_24_RIGHTN	GTG CAY CCG CAC CCC ATC CA
1b_14_RIGHT	TCAGGACGAGTACCTTGACCC	1b_26_LEFTN	TCCATGCTCACCGACCCCTCC
1b_15_LEFT	TTCCAAGTGGCCCATCTACAC	1b_26_RIGHTN	TGCCGCCACAGGAGGTTGG
1b_15_RIGHT	ATGTTGGGATGTGGCACGGT	1b_28_LEFTN	GTAGGCTGGCCAGGGGTCT
1b_16_LEFT	ATCATAATATGTGATGAGTGCCAC TC	1b_28_RIGHTN	CGGAGGGACGTAGTCCGGGT

1b_16_RIGHT	CCGGTAAAGCCCGTCATTAGAG	1b_30_LEFTN	CCACCTACCAAGGCCCTCC
1b_17_LEFT	TAGCGTATTACCGGGTCTTGA	1b_30_RIGHTN	AGAAGCTTGGCCTTAACTGTGG
1b_17_RIGHT	GCAACCCTGGTGTGTTTAGGTA	1b_31_LEFTN	CCGGAATACGACTTGGAGTTGA
1b_18_LEFT	TGTGTGAGTGCTATGACGCG	1b_31_RIGHTN	AGCCACCCTATTGATCTCACCT
1b_18_RIGHT	CAGCCGACATACATGCCATGAT	1b_32_LEFTN	CCATTGAGCCACTTGACCTACC
1b_19_LEFT	ACACCCCTGCTGTATAGGCTAG	1b_32_RIGHT	AGATGCCTACCCCTACAGAAAGT
1b_19_RIGHT	AAGGGTTCGCCACTTGGACT	1b_32_RIGHTN	GGAGTGTTTACCTCCCCGTTTCAT
1b_20_LEFT	TACATCGAACAGGGAATGCAGC		
1b_20_RIGHT	ATGTCCACAAGCACCTTCCCA		
1b_21_LEFT	ACCCAAAACACCCTCCTGTTTA		
1b_21_RIGHT	CGTCGCTTTCAGGCACATAGT		
1b_22_LEFT	CAGTGGATGAACCGGCTGATAG		
1b_22_RIGHT	GGAACCGTTTTTGACATGTCCG		
1b_23_LEFT	TCATGTCAACGTGGGTACAAGG		
1b_23_RIGHT	TCCCGTAGAAGAGGCTTGCA		
1b_24_LEFT	TAAAATGCCCATGCCAGGTTCC		
1b_24_RIGHT	TCTCTGACTCCACTCGGGTGAT		
1b_25_LEFT	TCTTCAGCCAGCCAGTTGTCT		
1b_25_RIGHT	TCTTTCTCCGTGGAGGTGGT		
1b_26_LEFT	CAGAGATCCTGCGGAAATCCAG		
1b_26_RIGHT	GACCAAGACCCGTCGCTAAGAT		
1b_27_LEFT	CGATCTTAGCGACGGGTCTT		
1b_27_RIGHT	GTCCTTTGCCCATAGCCAAA		
1b_28_LEFT	CCACAGTTAAGGCCAAGCTTCT		
1b_28_RIGHT	CTTTGATTTCCAGGCATTACCA		
1b_29_LEFT	AAGCCAGCTCGCCTTATCGT		
1b_29_RIGHT	GGTATTACCGCAGCTAGTCGTC		
1b_30_LEFT	AGACAGGCCATAAGGTCGCT		
1b_30_RIGHT	AGTACACCCTTTTGCCAGATGC		
1b_31_LEFT	CCGGAATACGACTTGGAGTTGA		
1b_31_RIGHT	AGCCACCCTATTGATCTCACCT		
1b_32_LEFT	CCATTGAGCCACTTGACCTACC		
1b_32_RIGHT	AGATGCCTACCCCTACAGAAAGT		

## 12.10. Ejemplo de Resultados de evaluación en Amplify 4

Primer	EU155304.2	AF139594.1	EU256062.1	EU256091.1	EU255962.1	EU155224.2	AF054250.1	AF054249.1	AF054248.1	EU781832.1	EU781830.1	EU781829.1	EU781828.1	EU781826.1	EU781825.1
1b_1LEFTN	no	good	good	no	no	good	good	good	good	no	good	good	good	good	good
1b_1RIGHTN	no	good	good	no	no	good	good	good	good	no	good	good	good	good	good
1b_2L															
1b_2R	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	moderate
1b_3LN															
1b_3SRN	okay	moderate	moderate	moderate	moderate	weak	moderate	moderate	moderate	okay	moderate	moderate	moderate	okay	moderate
1b_7_LN															
1b_7_RBLAST	okay	okay	okay	okay	moderate	weak	okay	okay	okay	moderate	moderate	okay	no	okay	okay
1b_9_FBLAST															
1b_9_RBLAST	okay	moderate	weak	weak	okay	moderate	weak	weak	weak	weak	okay	moderate	moderate	okay	moderate
1b_9_FBLAST															
1b_9_RBLAST	okay	moderate	moderate	moderate	okay	okay	moderate	moderate	moderate	moderate	okay	okay	okay	okay	okay
1b_10_LEFT															
1b_10_RIGHT	okay	okay	moderate	okay	okay	moderate	okay	okay	okay	okay	moderate	okay	okay	okay	okay
1b_11_LEFTM															
1b_11_RIGHT	okay	okay	okay	okay	weak	okay	okay	okay	okay	moderate	moderate	okay	okay	okay	okay
1b_15_LEFT															
1b_15_RIGHT	weak	moderate	moderate	moderate	moderate	moderate	moderate	weak	weak	weak	moderate	moderate	moderate	moderate	moderate
1b_16_FBLAS															
1b_16_RBLAS	okay	okay	moderate	okay	okay	okay	no	moderate	moderate	no	okay	moderate	okay	okay	okay
1b_17.5_FBLA															
1b_17.5_RBLA	good	okay	good	good	good	good	okay	moderate	moderate	good	good	good	good	good	good
1b_21_LEFT															
1b_21_RIGHT	moderate	okay	okay	okay	weak	okay	moderate	okay	okay	moderate	moderate	okay	okay	okay	moderate
1b_22_NBLAS															
1b_22_NBLAS	okay	okay	no	good	okay	okay	no	okay	okay	good	good	okay	good	good	okay
1b_24_NBLAS															
1b_24_NBLAS	good	okay	okay	okay	no	okay	okay	okay	okay	okay	okay	moderate	okay	okay	okay
1b_26_NBLAS															
1b_26_NBLAS	good	good	good	good	good	good	okay	good	good	good	good	good	good	okay	good
1b_28_NBLAS															
1b_28_NBLAS	varios	okay	okay	varios	okay	varios	moderate	moderate	var	okay	okay	okay	okay	okay	okay
1b_30_NBLAS															
1b_30_NBLAS	moderate	weak	moderate	weak	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate
1b_31_LEFT															
1b_31_RIGHT	moderate	moderate	moderate	no	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate	moderate
1b_32_LEFT															
1b_32_RIGHT	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay	okay
1b_32_RIGHTN		okay					okay	okay	okay	okay	okay		okay	okay	

## 12.11. Anexo XI Script para ensamblaje y limpieza de secuencias con Porechop y Flye

```
#Para empezar a usar Flye assembly se debe utilizar también anaconda
#Instalar directo de la página u obteniendo el link para el sistema operativo
correspondiente y sustituirlo en la línea de comandos de abajo. TODO EN EL
MISMO LUGAR, por ejemplo en MAC, en disco Untitled en administrador
wget https://repo.anaconda.com/archive/Anaconda3-2021.11-MacOSX-x86_64.sh
#si wget no funciona entonces utilizar
sudo apt install wget
#Se hace un directorio con las secuencias a ensamblar, aquí se descargan las
secuencias que se quieren ensamblar. Si se descarga de internet se utiliza el
archivo .fastq
mkdir Genome_assembly
#Se activa anaconda para instalar algunas herramientas
conda activate
conda config --add channels bioconda
conda config --add channelsconda-forge
#para verificar que se hayan instalado se usa el siguiente comando y deben
aparecer ambos paquetes
conda config --show channels
#ahora se activan los paquetes de flye
conda create -n genome_assembly porechop flye
#Ahora se deben activar los ambientes, primero se visualizan y luego se
selecciona el de genome que es en donde vamos a trabajar
conda env list
conda activate genome_assembly
#Si se activó correctamente entonces ahora se verá del lado izquierdo esto
(genome_assembly)
#Primero se usa Porechop para remover los adaptadores, se puede utilizar el
archivo .fastq.gz
porechop -i archivo.fastq.gz -o nuevonombre.trimmed.fa --format fasta -t 10
#Ahora se visualiza el inicio del archivo fasta. Aquí se ve la secuencia en fasta
head nuevonombre.trimmed.fa
#ENSAMBLAJE
#puede ensamblar con el archivo trimmed o directo de los archivos .gz
flye --nano-raw nuevonombre.trimmed.fa -o carpetadondesealmacena -g 9k -t 10 -i
2
#contar contigs en ensamblaje
grep ">" carpetadondesealmacena/assembly.fasta |wc -l
```