

# UACM

Universidad Autónoma  
de la Ciudad de México

---

*Nada humano me es ajeno*

COLEGIO DE CIENCIA Y TECNOLOGÍA

**MAESTRÍA EN CIENCIAS GENÓMICAS**

***“Diversidad genómica de virus de Dengue circulantes  
en mosquitos *Aedes aegypti* del estado de Guerrero”***

TESIS QUE PARA OBTENER EL GRADO DE  
MAESTRA EN CIENCIAS GENÓMICAS

PRESENTA:

**Biól. Cindy Fabiola Hernández Pérez**

Directora de tesis

**Dra. Claudia Selene Zárate Guerra**

Codirectora

**Dra. Rosa Martha Eugenia Yocupicio Monroy**

Ciudad de México, enero 2017.

## SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN



## UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA

### RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

### DERECHOS RESERVADOS ©

La presente obra y cada uno de sus elementos está protegido por la Ley Federal del Derecho de Autor; por la Ley de la Universidad Autónoma de la Ciudad de México, así como lo dispuesto por el Estatuto General Orgánico de la Universidad Autónoma de la Ciudad de México; del mismo modo por lo establecido en el Acuerdo por el cual se aprueba la Norma mediante la que se Modifican, Adicionan y Derogan Diversas Disposiciones del Estatuto Orgánico de la Universidad de la Ciudad de México, aprobado por el Consejo de Gobierno el 29 de enero de 2002, con el objeto de definir las atribuciones de las diferentes unidades que forman la estructura de la Universidad Autónoma de la Ciudad de México como organismo público autónomo y lo establecido en el Reglamento de Titulación de la Universidad Autónoma de la Ciudad de México.

Por lo que el uso de su contenido, así como cada una de las partes que lo integran y que están bajo la tutela de la Ley Federal de Derecho de Autor, obliga a quien haga uso de la presente obra a considerar que solo lo realizará si es para fines educativos, académicos, de investigación o informativos y se compromete a citar esta fuente, así como a su autor ó autores. Por lo tanto, queda prohibida su reproducción total o parcial y cualquier uso diferente a los ya mencionados, los cuales serán reclamados por el titular de los derechos y sancionados conforme a la legislación aplicable.



# Índice general

<b>0. Resumen</b>	<b>12</b>
<b>1. Introducción</b>	<b>14</b>
1.1 Importancia del estudio de la diversidad genómica de los arbovirus	14
1.2 Dengue: impacto de la enfermedad	16
1.3 Dengue: enfermedad diversa causada por virus diversos	18
1.4 DENV: partícula viral y genoma	19
1.5 Ciclo de transmisión del DENV	21
1.6 Ciclo de vida de <i>Aedes aegypti</i>	23
1.7 Replicación del DENV en <i>Aedes aegypti</i>	24
1.8 Diversidad del DENV en América	24
1.9 Diversidad del DENV en México	26
1.10 Diversidad del DENV en el estado de Guerrero	29
<b>2. Antecedentes</b>	<b>30</b>
2.1 Serotipos y genotipos del DENV presentes en México	30
2.2 Diversidad del DENV aislados de <i>Aedes aegypti</i> en México	31
2.3 Variación intrahospedero del DENV en mosquito y/o humano	32
<b>3. Justificación</b>	<b>35</b>
<b>4. Hipótesis</b>	<b>36</b>
<b>5. Objetivos</b>	<b>37</b>
5.1 Objetivo general	37
5.2 Objetivos particulares	37
<b>6. Materiales y método</b>	<b>38</b>
6.1 Muestras	38
6.1.1 <i>Pools</i> de mosquitos	42
6.1.2 Sueros de pacientes infectados con DENV	42
6.2 Procesos llevados a cabo en laboratorio	42
6.2.1 Procesamiento de muestras de <i>pools</i> de mosquitos	42
6.2.1.1 Extracción de RNA	43
6.2.1.2 RT-PCR One step singleplex	43
6.2.1.3 Depleción de RNAr	43
6.2.1.4 Procedimientos llevados a cabo en la Unidad de Secuenciación Masiva y Bioinformática (USMB) del Instituto de Biotecnología (IBT), UNAM, Cuernavaca, Morelos	44
6.2.1.4.1 Verificación de la integridad del RNA / RIN	44
6.2.1.4.2 Construcción de las bibliotecas genómicas	44
6.2.1.4.3 Secuenciación masiva	45
6.2.2 Procesamiento de muestras de sueros de pacientes infectados con DENV	45
6.3 Procesos <i>in silico</i>	46

6.3.1	“Prefiltrado” de secuencias	46
6.3.1.1	Análisis de calidad de las secuencias	46
6.3.1.2	Limpieza de adaptadores presentes en las secuencias	47
6.3.1.3	Eliminación de las secuencias redundantes	48
6.3.2	Obtención del genoma consenso mapeando contra genoma de referencia	48
6.3.2.1	<i>Binning</i> , reclutamiento de fragmentos	49
6.3.2.2	Remapeo contra genoma de referencia, determinación de variantes y obtención de genoma consenso	51
6.3.2.3	Análisis de calidad del mapeo contra genoma de referencia	52
6.3.3	Ensamble de genoma <i>de novo</i>	52
6.4	Genotipificación	53
6.5	Análisis de variabilidad de genomas	53
6.5.1	Análisis de variabilidad de genoma completo, con respecto a genoma del DENV proveniente de suero humano y de mosquito	53
6.5.2	Análisis del uso relativo de codones sinónimos	53
6.6	Filogeografía	54
6.7	Estructura poblacional	55
<b>7.</b>	<b>Resultados</b>	<b>56</b>
7.1	Extracción de RNA	56
7.2	RT-PCR <i>One step singleplex</i> , serotipificación	56
7.3	Depleción de RNAr	57
7.4	Verificación de la integridad del RNA / RIN	58
7.5	RNA-seq	59
7.6	Análisis de calidad de la secuenciación	59
7.7	Limpieza de adaptadores y eliminación de secuencias redundantes	59
7.8	<i>Binning</i> , reclutamiento de fragmentos	61
7.9	Obtención del genoma consenso	61
7.9.1	Mapeo contra genoma de referencia, variantes poblacionales	62
7.9.2	Ensamble <i>de novo</i>	
7.10	Genotipificación	64
7.11	Análisis de variabilidad de genoma completo, con respecto a genoma del DENV proveniente de suero humano y de mosquito	65
7.12	Uso relativo de codones sinónimos	69
7.13	Filogeografía	70
7.14	Estructura poblacional	72
<b>8.</b>	<b>Discusión</b>	<b>74</b>
8.1	Genomas obtenidos: profundidades, coberturas, longitud de contigs	74
8.2	Genomas "consenso" y variantes	77
8.3	Serotipo y genotipo reportado, consistente con la diversidad reportada para Guerrero	78
8.4	Un nucleótido con variación poblacional relevante	79
8.5	Uso relativo de codones	84
8.6	Estructura poblacional	84
8.7	Selección natural	85
<b>9.</b>	<b>Conclusiones</b>	<b>87</b>

<b>10. Perspectivas</b>	<b>89</b>
<b>11. Referencias</b>	<b>90</b>
<b>12. Anexos</b>	<b>99</b>
<b>13. Resumen del trabajo presentado en el IX Congreso Nacional de Virología, septiembre de 2015</b>	<b>104</b>

## Índice de figuras

1.1	Número de secuencias de DENV presentes en la base de datos ViPR, derivadas de muestras obtenidas de humanos o de mosquitos (DENV-2_mosquito).	16
1.2	Promedio anual de casos de fiebre de dengue y fiebre de dengue hemorrágico reportados a la Organización Mundial de la Salud (WHO) / Número de países con reporte de dengue hasta 2007.	17
1.3	Distribución geográfica del dengue.	17
1.4	Proporción de los distintos grados de severidad de las infecciones por DENV.	18
1.5	Estructura del virión del DENV.	19
1.6	Genoma del DENV.	19
1.7	Topología propuesta para las proteínas estructurales y no estructurales del DENV en el retículo endoplásmico.	21
1.8	Ciclos de transmisión del DENV.	23
1.9	Línea de tiempo que muestra la introducción de varios genotipos de los cuatro serotipos del DENV.	25
1.10	Total de casos de DF y DHF en la República Mexicana reportados por la Secretaría de Salud durante 1984-2011.	26
1.11	Distribución geográfica de los casos de fiebre de dengue y fiebre de dengue hemorrágico.	28
1.12	Circulación de serotipos del DENV en México, 2000-2011.	29
6.1	Plan general de trabajo seguido para alcanzar los objetivos del presente estudio en coordinación con distintas instituciones.	39
6.2	Sitios de colecta de mosquitos y toma de muestras analizados en el presente estudio.	41
6.3	Construcción de bibliotecas genómicas.	45
6.4	Región amplificada del gen E para las muestras de sueros humanos.	46
6.5	Características de las bibliotecas genómicas generadas.	47
6.6	Secuencias <i>paired-end</i> .	47
6.7	Estructura del formato fastq.	47
6.8	Reporte de los adaptadores removidos generado por PEAT.	48

6.9	Formato vcf.	52
6.10	Fragmento de matriz de RSCU para tres de las secuencias analizadas en este estudio.	54
6.11	Países representados en las secuencias con las que se realizó el análisis filogeográfico.	55
7.1	Cuantificación de la concentración de RNA total extraído de cada muestra mediante espectrofotometría.	56
7.2	Comparación de la concentración de RNA total cuantificada por espectrofotometría de manera previa y posterior al tratamiento con el kit de depleción de RNAr.	58
7.3	Electroferogramas generados a partir de los datos de <i>Agilent 2100 Bioanalyzer</i> para cada una de las muestras analizadas.	58
7.4	<b>(A)</b> Número de <i>reads</i> obtenidos de la RNA-seq por muestra. <b>(B)</b> Número de <i>reads</i> que quedan después de la limpieza de adaptadores y duplicados. <b>(C)</b> Número máximo de <i>reads</i> que hicieron <i>match</i> con la base de datos de <i>Flaviviridae</i> .	60
7.5	Distribución de variantes poblacionales de la muestra MYF5 a lo largo del genoma.	61
7.6	Número de <i>reads</i> por muestra resultantes del <i>binning</i> . Proporción de secuencias que se logró ensamblar.	62
7.7	Esquema que representa la distribución de los <i>contigs</i> obtenidos mediante ensamble <i>de novo</i> .	62
7.8	Árbol filogenético (NJ) construido con secuencias de genoma completo, que refleja el genotipo Asiático-Americano de los genomas de DENV-2 provenientes de <i>Aedes aegypti</i> .	64
7.9	Árbol filogenético (NJ) construido con secuencias de 861 nucleótidos del gen E, que refleja el genotipo Asiático-Americano de las secuencias de DENV-2 extraídas a partir de sueros humanos.	65
7.10	Tipo de variantes comunes a los genomas MYF1, MYF2 y MYF5 al compararlos con genomas del DENV-2 (Asiático-Americano) provenientes de suero de humano (gb:GU131955, Quintana Roo, México, 2004) y de <i>Aedes aegypti</i> (gb:KJ189308 y gb:KJ189309, Yucatán, México).	67
7.11	Dendograma de agrupamiento de las secuencias según su RSCU.	69
7.12	MCC ( <i>Maximum Clade Credibility Tree</i> ) construido a partir de 174 secuencias del DENV de genotipo Asiático-Americano y selvático con 861 nucleótidos del gen E (incluidas las generadas en el presente estudio).	70
7.13	Acercamiento del MCC construido a partir de 174 secuencias del DENV de genotipo Asiático-Americano y selvático con 861 nucleótidos del gen E (incluidas las generadas en el presente estudio).	71

7.14	MCC ( <i>Maximum Clade Credibility Tree</i> ) construido a partir de las secuencias de 861 nucleótidos del gen E obtenidas de muestras de sueros humanos y mosquitos en el presente estudio.	72
7.15	Resultados de la prueba Slatkin & Maddison considerando estructuración poblacional (Arriba) por geografía y (Abajo) por temporalidad. Gráficas construidas con R a partir de los datos generados por HyPhy.	73
8.1	Relación entre Ct y profundidad de secuenciación obtenida por mapeo contra genoma de referencia para las distintas muestras.	76
8.2	Incidencia de casos estimados por FD y FHD por 100,000 habitantes durante las 52 semanas epidemiológicas de 2012 en el estado de Guerrero (a partir de los datos del InDRE, 2012).	79
8.3	Sitios variantes que favorecen la replicación del genotipo Asiático-Americano en las glándulas salivales de los mosquitos, incluida la isoleucina 162, variante poblacional de las muestras analizadas en el presente estudio.	82
8.4	Distribución de variantes de la tercera posición del codón del aminoácido 162 de la proteína E en las muestras de DENV de Guerrero. Se observa que es un sitio “muy variable”.	83

## Índice de tablas

1.1	Comparación de porcentajes de identidad de las proteínas estructurales y no estructurales para los cuatro serotipos.	18
1.2	Funciones de las proteínas del DENV.	20
1.3	Datos de los brotes epidémicos de dengue ocurridos en México.	26
1.4	Frecuencia relativa de los serotipos del DENV aislados de sueros de pacientes del estado de Guerrero registrados como casos de DF y DHF.	29
6.1	Datos de colecta de las muestras de <i>pools</i> de mosquitos analizadas en el presente estudio.	40
6.2	Datos de procedencia de las muestras de DENV-2 obtenidas de sueros humanos del estado de Guerrero.	42
6.3	Oligonucleótidos (DENV, NS5) empleados para las reacciones de RT-PCR <i>One step</i> .	43
6.4	Oligonucleótidos (DENV, E) empleados para las reacciones de PCR.	46
6.5	Virus de la base de datos de <i>Flaviviridae</i> (obtenidos de <i>Nucleotide</i> de NCBI) con mayor número de secuencias.	49
7.1	Resultados de la amplificación por RT-PCR <i>One step singleplex</i> .	57
7.2	Valores de Ct ("Threshold cycle") para la RT-PCR <i>singleplex</i> .	57
7.3	Posiciones variables con respecto a la referencia, presentes en los archivos vcf correspondientes.	61
7.4	Detalles del mapeo contra genoma de referencia y ensamble <i>de novo</i> de las secuencias.	63
7.5	Número de mutaciones sinónimas y no sinónimas (conservativas y no conservativas) comunes a los genomas de MYF1, MYF2 y MYF5 con respecto a genomas de DENV provenientes de suero humano (GU131955) y <i>Aedes aegypti</i> (KJ189309, KJ189308).	66
7.6	Proporción de mutaciones por gen en relación con su longitud.	66
7.7	Posiciones del genoma comunes a MYF1, MYF2, MYF5, KJ189308 y KJ189309, pero variantes con respecto a GU131955.	68
8.1	Registros de DENV-2 presentes en la base de datos especializada ViPR para México.	80
8.2	Variantes poblacionales del fragmento del gen analizado.	84

*A Mónica Ángeles,  
por su inteligente y sensible acompañamiento,  
por enseñarme que podemos ser otros que también somos nosotros.*

## Agradecimientos

El presente trabajo se realizó gracias a la colaboración de diversas personas e instituciones a las que agradezco profundamente.

A la **Universidad Autónoma de la Ciudad de México**, por permitirme acceder a una formación de posgrado.

Al **Conacyt**, por haberme otorgado la beca No. 336972 para realizar los estudios de maestría durante el periodo de julio de 2014 a julio de 2016.

A la doctora **Selene Zárate**, por permitirme desarrollar la presente investigación en el laboratorio de Bioinformática de la UACM, por mostrarme su interesante campo de estudio y por la dirección de esta tesis.

A la doctora **Martha Yocupicio**, por hacerme partícipe del proyecto Conacyt No. 233977 “Perfil de RNA mensajeros y MicroRNAs de sueros de pacientes con Dengue del estado de Guerrero”, por la gestión de la obtención de las muestras con las que se realizó el presente estudio, por el financiamiento para el procesamiento y la secuenciación de éstas, y por la dirección de esta tesis.

Al **Laboratorio Estatal de Salud Pública del Estado de Guerrero**, por la amable donación de las muestras con las que se realizó la presente investigación.

A la **Unidad de Secuenciación Masiva y Bioinformática (USMB)** del Instituto de Biotecnología (IBT), UNAM, por su eficiente servicio de secuenciación, y especialmente al doctor **Ricardo Grande**, por su disposición para proporcionar toda la información adicional que solicité.

A la doctora **Blanca Taboada**, por sus sugerencias para el procesamiento *in silico* de las secuencias y por la lectura de esta tesis.

A las doctoras **Sofía Alcaraz** y **Lilia López**, por su asesoramiento y comentarios durante el desarrollo de este trabajo.

A la doctora **Elisa Azuara**, por sus acertados comentarios sobre el proyecto y el escrito.

A **tod@s l@s profesor@s del posgrado de Ciencias Genómicas**, por contribuir a la formación de las nuevas generaciones de científicos.

Al M. en C. G. **Hugo Castelán**, por su gran y siempre desinteresada ayuda para resolver mis dudas bioinformáticas. Por todo el tiempo que le dedicó a este proyecto.

A la M. en C. **Violeta Romero**, por facilitarme el acceso a muchas de las publicaciones científicas que se encuentran en la bibliografía de esta tesis.

A **Iliana Juárez**, por su valiosa colaboración en la corrección de estilo de esta tesis.

A la bióloga **Jimena Manzanares**, por su valiosa colaboración en el diseño y edición de las figuras.

A los compañeros del **Departamento de Salud Pública de los estados de Morelos y Oaxaca**, que me permitieron ver la dimensión humana de nuestro trabajo.

## **Agradecimientos personales**

A **Jime**, con quien comparto gustosa las metamorfosis.

A **mi madre**, quien me brindó el apoyo económico para poder concluir el presente trabajo.

A **mi familia**, por su cariño.

A mis amigas de siempre, **Violeta, Adriana, Ceci, Susanita**, por inspirarme con sus respectivos caminos, por estar siempre.

A mis amig@s uacemitas, **Jaquie, Raquel, Miguel, Caty, Vivi y Rosi**, por la compañía, las charlas y el buen humor.

## 0. Resumen

El virus de dengue (DENV) es en la actualidad el arbovirus con mayor impacto para los seres humanos. Debido a su RNA polimerasa sin capacidad de corrección y su elevada tasa de replicación entre dos hospederos (mosquitos del género *Aedes* y humanos), tiene una tasa evolutiva característicamente elevada. La transición entre hospederos hace que sobre las mutaciones generadas actúen fuerzas evolutivas con diferentes intensidades a lo largo de su ciclo de transmisión. Se sabe además que la diversidad de los DENV circulantes en una región determinada se relaciona con los posibles impactos que tenga la enfermedad en la salud.

A pesar de los avances que las tecnologías de secuenciación han alcanzado al día de hoy, y de la importancia que tiene para el mantenimiento de la diversidad el paso del DENV en las poblaciones de mosquitos, los análisis de la diversidad de DENV presentes en las poblaciones de los hospederos artrópodos, tanto durante brotes epidémicos como en épocas entre brotes, no ha sido analizada a profundidad.

En el presente estudio se analiza la secuencia genómica de la población de DENV presentes en *pools* de mosquitos colectados en las cercanías y el interior de los hogares con casos probables o confirmados de dengue. Observamos que la diversidad de DENV presente en mosquitos refleja la que también está presente en los pacientes infectados, lo que muestra el dinamismo interhospedero y la importancia de la vigilancia entomoviológica para el monitoreo de la enfermedad.

Se reporta también la presencia de variantes distribuidas de manera homogénea a lo largo de todo el genoma y no sólo concentradas en el gen de la Envoltura (E), como suele considerarse tradicionalmente. Se presentan de hecho variantes no sinónimas y no conservativas concentradas en el gen de la RNA polimerasa (NS5). Algunas de las posiciones variables reportadas podrían ser específicas de mosquitos; el papel que éstas puedan tener en el *fitness* del DENV queda por investigar.

Tanto en los genomas de DENV completos, aislados de mosquitos, como en los fragmentos del gen E secuenciados a partir de sueros de pacientes infectados con dengue, el número de variantes sinónimas identificadas supera por mucho al número de variantes no sinónimas. Esto se

interpreta como evidencia de selección negativa (purificadora); sin embargo, las mutaciones tradicionalmente silenciosas podrían enmascarar una naturaleza no neutral en contextos de reconocimiento de RNAi como parte de la respuesta inmune de los mosquitos hacia el DENV. Se reportan tres sitios variables del gen E para los que la población de DENV presenta polimorfismos destacados. Uno de ellos (en el codón 162 del gen E) ha sido reportado como importante para el aumento de competencia del DENV que tuvo el genotipo Asiático-Americano, en comparación con el genotipo nativo Americano. La relación que estas mutaciones silenciosas puedan tener con los RNAi del mosquito también quedan por investigar.

Finalmente, mediante análisis filogeográficos se corroboró que la población del DENV a escala continental se estructura o agrupa de acuerdo con su geografía y también con su temporalidad. Es posible que el número de eventos de invasión del DENV a las localidades sea menor a lo que se ha considerado hasta ahora, y que los procesos de generación y mantenimiento de la diversidad en los mosquitos tenga un papel más importante que el que hasta ahora ha sido reconocido para la circulación del DENV.

# 1. Introducción

## 1.1 Importancia del estudio de la diversidad genómica de los arbovirus

Los brotes recientes de virus de RNA transmitidos por artrópodos (arbovirus) pertenecientes a distintas familias, incluidas *Flaviviridae* (Zika, West Nile, Dengue) y *Togaviridae* (Chikungunya), evidencian su potencial para causar enfermedades humanas que repercuten de manera importante en la salud pública (De Coffey *et al.*, 2014). El efecto de los arbovirus puede ser potenciado por su tasa evolutiva inherentemente elevada (Holmes, 2008), así como por la cambiante ecología viral. Por ello, la vigilancia epidemiológica de los arbovirus tiene un papel crucial como sistema de alerta temprana con el fin de mitigar las consecuencias que puedan tener estas infecciones causantes de enfermedades en humanos y en otros animales (De Coffey *et al.*, 2014).

La rápida evolución de los arbovirus se debe tanto a la característica acción de la RNA polimerasa dependiente de RNA, sin capacidad de corrección de errores, como a sus rápidas tasas de replicación (Holmes, 2008). Las mutaciones ocurren durante la replicación viral y sobre ellas actúan distintas fuerzas evolutivas, como la deriva génica y la selección natural positiva o negativa (Thai *et al.*, 2012).

En el Dengue, la alta tasa de errores de replicación se refleja en el hecho de que las poblaciones del virus (DENV, *Dengue Virus*) son muy diversas (Lambrechts *et al.*, 2011), tanto a nivel del hospedero infectado (humano o mosquito), en donde el virus existe como poblaciones de variantes relacionadas (Sim & Hibberd, 2016; Sim *et al.*, 2015; Thai *et al.*, 2012), como a escala global. El DENV es un complejo de cuatro serotipos distintos, cada uno con diferentes genotipos, que varían en grados de virulencia y distribución geográfica (Twiddy *et al.*, 2002).

Se considera que para los arbovirus en general, y para el DENV en particular, la diversidad intrahospedero puede ofrecer una ventaja para superar las distintas presiones de selección a las que son sometidos durante su paso entre el hospedero humano y el artrópodo (mosquito) (Beltramello *et al.*, 2010; Sim *et al.*, 2014). Se ha propuesto que, en mosquitos, el RNA de

interferencia (RNAi, un mecanismo de defensa antiviral clave en insectos) es una fuerza directriz de la diversidad genética intrahospedero. La alternancia de hospedero somete a los arbovirus a frecuentes y significativas reducciones en el tamaño poblacional. Sólo un porcentaje pequeño del total de la población viral circulante en el humano es ingerido por el hospedero mosquito en su alimento de  $\sim 2 \mu\text{L}$ , e incluso un número menor de partículas virales logra replicarse en su intestino. Las reducciones subsecuentes en el tamaño poblacional también ocurren durante la dispersión, a través de los distintos tejidos y órganos del insecto, así como durante la inyección de los microlitros de saliva infectada en el hospedero humano. ¿Cómo es que estos procesos moldean la diversidad y el repertorio de la población viral? Aún no se ha explicado a profundidad (Sim & Hibberd, 2016).

A pesar de que el conocimiento de la diversidad en ambos hospederos es crucial para el entendimiento de la evolución de los arbovirus y, como consecuencia, para el desarrollo de tratamientos y estrategias de control (Sim & Hibberd, 2016), el número de estudios que se han enfocado en el análisis de la diversidad dentro del hospedero artrópodo es muy reducido, en comparación con aquellos realizados atendiendo al hospedero humano (destacan: Sim *et al.*, 2015; Carrington *et al.*, 2015; Lambrechts *et al.*, 2011; Coffey *et al.*, 2011a., Coffey *et al.*, 2011b; Lin *et al.*, 2004). El grado de variación de las secuencias del DENV en el mosquito y cómo la estructura poblacional cambia durante la transmisión entre humanos y mosquitos permanece casi desconocido (Lin *et al.*, 2004).

La amplificación y secuenciación del genoma del virus del Dengue directamente de muestras de pacientes se han vuelto rutinarias en muchos laboratorios, por lo que ha aumentado la disponibilidad de secuencias del DENV para llevar a cabo estudios de variación genética y evolución viral (Sim *et al.*, 2016).

Cuando analizamos las entradas de ViPR (*Virus Pathogen Database and Analysis Resource*<sup>1</sup>), una de las bases de datos especializadas en DENV más importantes, es posible observar que el número de secuencias de DENV-2<sup>2</sup> (serotipo 2) derivadas de muestras extraídas de humano es de alrededor de 5,600, en contraste con sólo 86 registros de secuencias provenientes de mosquitos (figura 1.1). De éstas, un número mucho más reducido corresponde a genoma completo: alrededor de 1,300 muestras provenientes de humanos y tan solo tres provenientes de mosquitos.

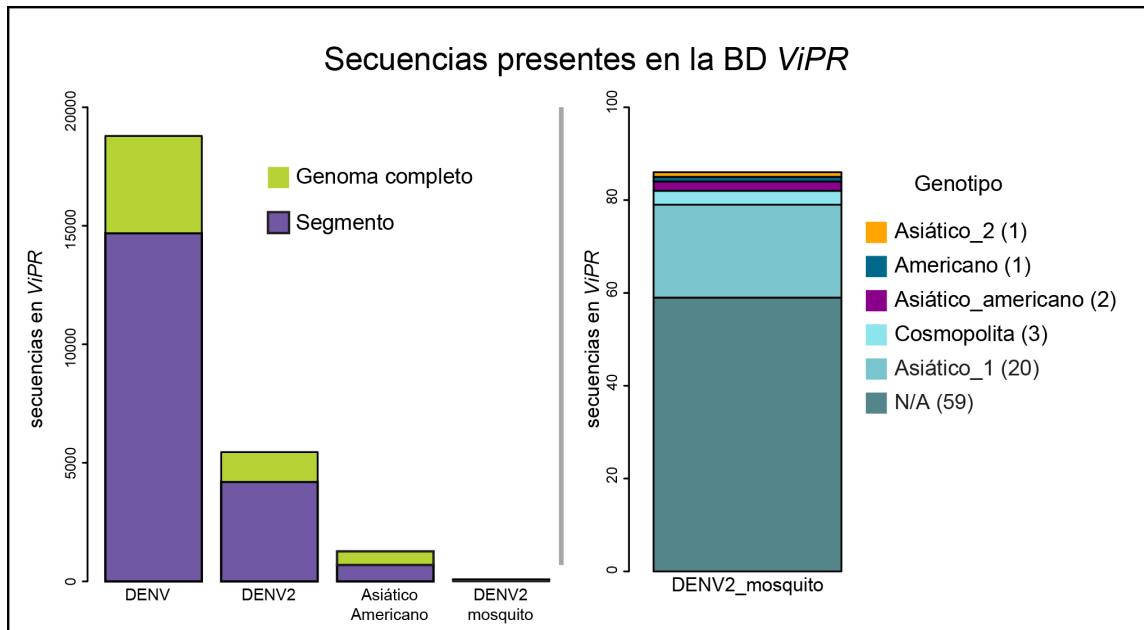
---

<sup>1</sup> Consultada en octubre de 2016. <http://www.viprbrc.org>

<sup>2</sup> Debido a su asociación con epidemias severas y frecuentes, el DENV-2 es el serotipo para el cual existe un mayor número de datos disponibles (Sim & Hibberd, 2016).

<sup>2</sup> Debido a su asociación con epidemias severas y frecuentes, el DENV-2 es el serotipo para el cual existe un mayor número de datos disponibles (Sim & Hibberd, 2016).

La secuenciación de genomas completos y el análisis filogenético de virus emergentes circulando globalmente, como el Dengue, son necesarios para entender la diversidad genética y la evolución de patógenos responsables de altas tasas de mortalidad y morbilidad humana a escala global cada año (Dugan *et al.* 2012).



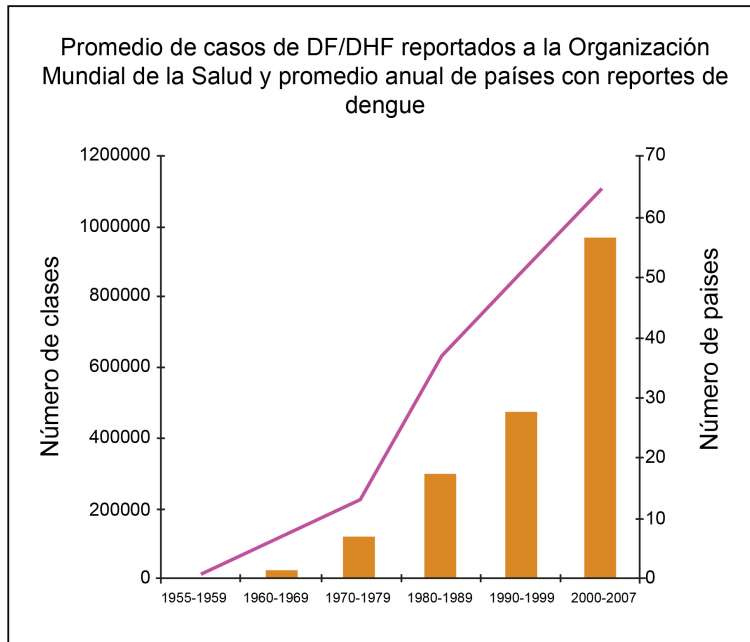
**Figura 1.1. (Izquierda)** Número de secuencias de DENV presentes en la base de datos ViPR, derivadas de muestras obtenidas de humanos o de mosquitos (DENV-2\_mosquito). Se muestra la proporción de genomas completos contra segmentos de genomas (predominantemente el gen E). **(Derecha)** Genotipos representados en las 86 secuencias de DENV-2 derivadas de muestras obtenidas de mosquito. Las tres secuencias de genoma completo corresponden al genotipo Asiático-Americano (2) y al genotipo no asignado (1). DENV-2: dengue serotipo 2.

## 1.2 Dengue: impacto de la enfermedad

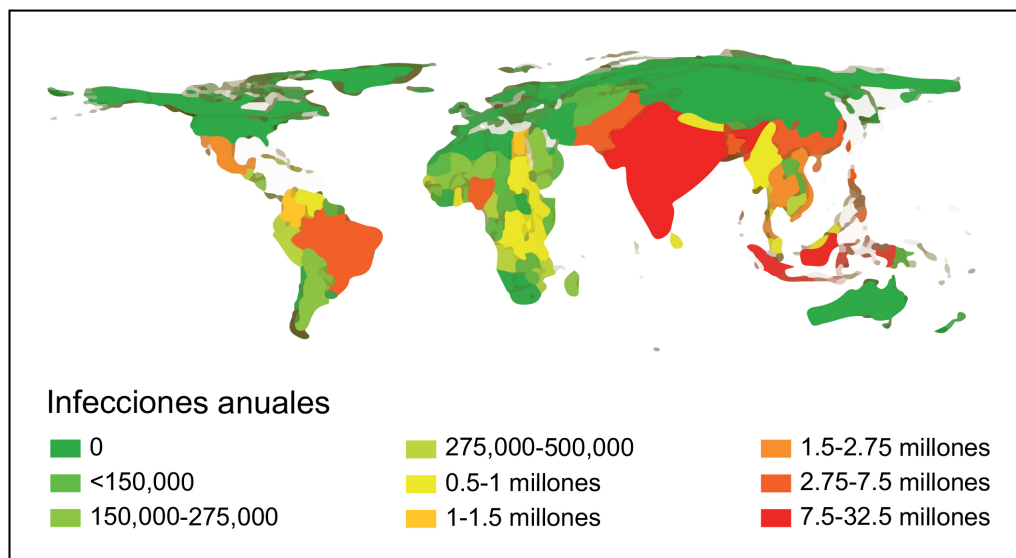
En la actualidad, el dengue es la infección viral transmitida por artrópodo con mayor impacto para los seres humanos (Kyle & Harris, 2008). Se estima que a escala mundial ocurren más de 390 millones de infecciones al año (I. C. 95% = 284-528) y que 96 millones de casos (I. C. 95% = 67-136) alcanzan gravedad clínica o subclínica (Bhatt *et al.*, 2013). Desde su emergencia, hace poco más de 50 años, la incidencia de las formas severas del dengue ha aumentado más de 500 veces (figura 1.2) y en el presente más de 100 países son afectados por este padecimiento (figura 1.3), particularmente en zonas tropicales y subtropicales (Kyle & Harris, 2008).

La extensión espacial del dengue está en función de la distribución de sus vectores (mosquitos del género *Aedes*), que a su vez se ha ampliado en respuesta al elevado crecimiento

demográfico, al aumento en la movilidad de los seres humanos y al grado de urbanización de la población, así como a cambios climáticos globales (Kyle & Harris, 2008). El virus es propagado por la picadura de la hembra de *Aedes aegypti* y de *Aedes albopictus*, y ésta puede transmitirlo vía transovárica a su descendencia (Vasilakis *et al.*, 2011).



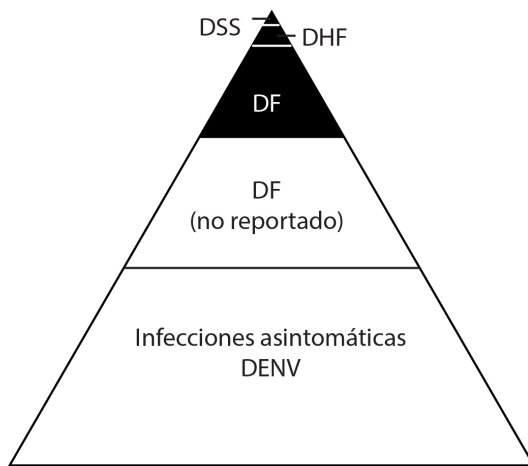
**Figura 1.2.** Promedio anual de casos de fiebre de dengue y fiebre de dengue hemorrágico reportados a la Organización Mundial de la Salud (WHO) / Número de países con reporte de dengue hasta 2007 (Awaidi *et al.*, 2014).



**Figura 1.3.** Distribución geográfica del dengue. Imagen tomada de Bhatt *et al.*, 2013.

### 1.3 Dengue: enfermedad diversa causada por virus diversos

El dengue es una enfermedad con grados de severidad variable que comprende desde infecciones asintomáticas (de 50% a 90% de los casos, figura 1.4) hasta casos letales. Se identifican tres grados de severidad: fiebre de dengue o dengue clásico (DF, del inglés *Dengue Fever*), fiebre de dengue hemorrágico (DHF, del inglés *Dengue Hemorrhagic Fever*) y Síndrome de choque por dengue (DSS, *Dengue Shock Syndrome*), que puede ser mortal (Kyle & Harris, 2008).



**Figura 1.4.** Proporción de los distintos grados de severidad de las infecciones por DENV representadas por una pirámide. Entre 10 y 50% de las infecciones son sintomáticas. Sólo una fracción de los casos de DF es reportado. Las manifestaciones más severas de la enfermedad representan la punta de la pirámide y se presentan con mayor frecuencia en infecciones secundarias que en infecciones primarias. (Modificada de Kyle & Harris, 2008).

El DENV es un complejo de cuatro serotipos distintos antigénicamente relacionados y que han sido designados como DENV-1, 2, 3, y 4. A nivel de proteína, los distintos serotipos tienen una identidad de entre 63% y 78% (Tabla 1.1).

**Tabla 1.1.** Comparación de porcentajes de identidad de las proteínas estructurales (por encima de la diagonal) y no estructurales (por debajo de la diagonal) para los cuatro serotipos (Tomada de Westaway & Blok 1997).

Abreviaciones: D1 = dengue-1 Singapore S275/90 (Fu *et al.*, 1992); D2 = dengue-2 Jamaica (Deubel *et al.*, 1988); D3 = dengue-3 H87 (Osatami and Sumiyoshi, 1990); y D4 = dengue 4-Dominica (Zhao *et al.*, 1986; Mackow *et al.*, 1987).

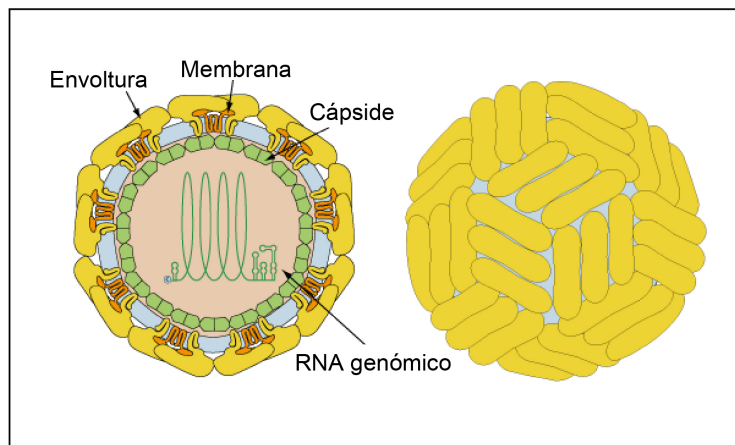
	D1	D2	D3	D4
D1	-	71.1%	78.7%	65.4%
D2	67.1%	-	67.1%	66.1%
D3	73.0%	73.3%	-	63.2%
D4	64.3%	69.1%	65.3%	-

Para cada serotipo se distinguen genotipos en función de su diversidad genética. Los genotipos varían en su virulencia y distribución geográfica. Para DENV-2 se reconocen seis genotipos: Americano, Asiático 1 y 2, Asiático-Americano, Cosmopolita y Selvático (Twiddy *et*

al., 2002).

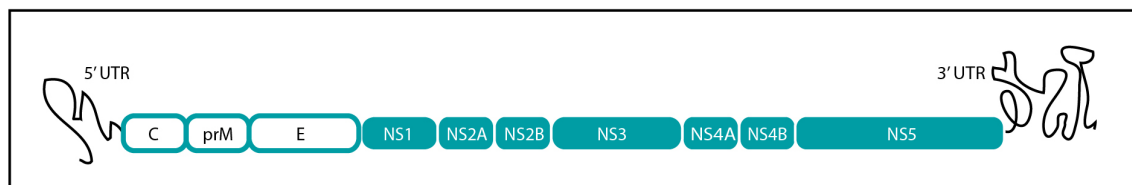
#### 1.4 DENV: partícula viral y genoma

El DENV es un virus de RNA de cadena sencilla y positiva (ssRNA+). El virión es una partícula icosaédrica, de 40-50 nm con membrana lipídica (obtenida de las células del hospedero), sobre la cual se insertan las proteínas de membrana (M) y envoltura (E) (figura 1.5). El interior del virus contiene el complejo riboproteico conformado por la proteína de la cápside y el genoma viral (Velandia & Castellanos, 2011).



**Figura 1.5.** Estructura del virión del DENV. Imagen tomada de [http://viralzone.expasy.org/all\\_by\\_species/24.html](http://viralzone.expasy.org/all_by_species/24.html).

El DENV tiene un genoma que va de 10.3 a 11 kb y codifica para tres proteínas estructurales y siete no estructurales (figuras 1.6 y 1.7). Las proteínas estructurales comprenden las proteínas de la cápside (C) membrana (M) y envoltura (E), mientras que las proteínas no estructurales incluyen NS1, NS2A, NS2B, NS3, NS4A, NS4B y NS5. (La función de cada una de las proteínas se encuentra descrita brevemente en la tabla 1.2). Las regiones codificantes del genoma del DENV están flanqueadas por 94 nucleótidos (nt) en la región 5'UTR y por 388-462 nt en la región 3'UTR (Azhar *et al.* 2015). Se ha observado que las regiones UTR se asocian con factores del hospedero y forman estructuras secundarias y terciarias requeridas para la viabilidad del virus (Fernández-García *et al.*, 2009).



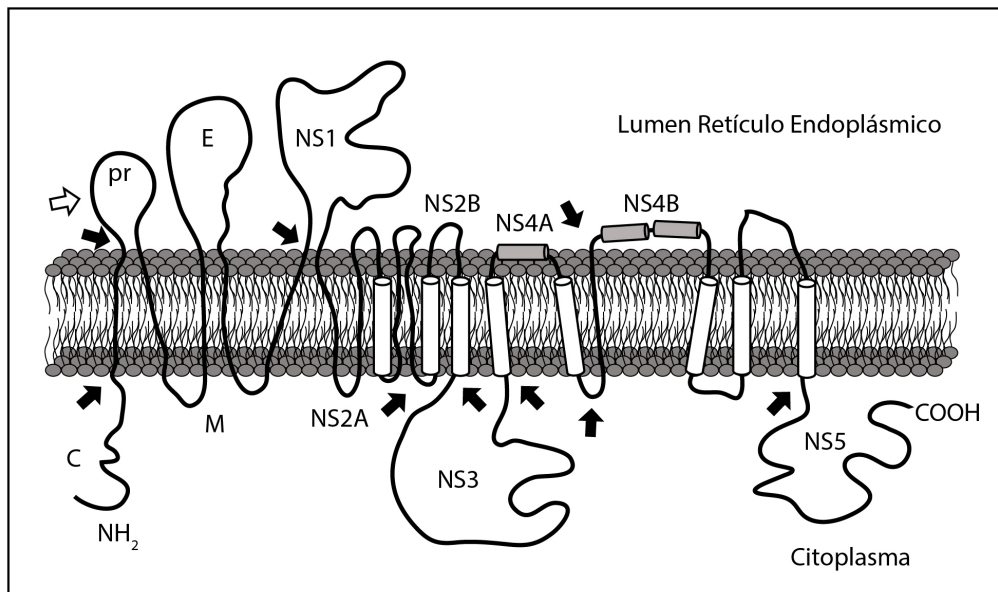
**Figura 1.6.** Genoma del DENV. Un único marco de lectura (ORF) codifica tres proteínas estructurales

(cápside = C, glicoproteínas de membrana = M y envoltura E) y siete proteínas no estructurales (NS1, NS2A, NS2B, NS3, NS4A, NS4B y NS5). (Imagen tomada de Guzmán *et al.*, 2010).

**Tabla 1.2.** Funciones de las proteínas del DENV (Velandia & Castellanos, 2011).

Proteína	Masa molar	Función
Proteína C	11kDa	Constituida por cuatro hélices alfa: 1: Ubicada en el extremo N-terminal y orientada hacia el citoplasma, posee aminoácidos de carácter básico que se asocian y unen fuertemente al RNA genómico formando el complejo riboprotéico o nucleocápside que protege al RNA viral de la degradación y promueve la organización del RNA en el interior de la partícula viral. 2: De naturaleza muy hidrofóbica. Actúa como bisagra que favorece el acercamiento del RNA viral al resto de la proteína C anclada en la membrana del retículo endoplásmico. Recluta pequeñas gotas lipídicas presentes en el citoplasma que promueven la formación de la proteína viral. 3 y 4: Hidrofóbicas, anclan la proteína a la membrana del retículo endoplásmico.
Proteína precursora de la membrana (prM) / proteína de la membrana (M)	26kDa / 8kDa	Presente en los viriones inmaduros, junto con la proteína M, participa fundamentalmente en el proceso de maduración de la partícula viral. La prM es procesada después de la traducción por la proteasa celular furina, que la divide en dos y genera el péptido pr y la proteína M. Se ha identificado la presencia de un dominio ApoptoM que induce muerte celular.
Proteína de envoltura E	50kDa	Glicoproteína que posee tres dominios I, II y III, se distribuye sobre la superficie del virus, formando complejos homodiméricos de tipo cabeza-cola. Los dominios II y III de cada una de las proteínas del homodímero son determinantes para las interacciones entre el virus y los receptores de las células vulnerables. Interactúa con las moléculas receptoras de la membrana plasmática de las células vulnerables que favorecen la endocitosis del virus. Estimula la respuesta inmune del individuo e induce la producción de anticuerpos neutralizadores.
NS1	46kDa	Forma dímeros o hexámeros asociados a balsas lipídicas de la membrana plasmática. También se puede hallar soluble en el citoplasma y el espacio extracelular por lo que puede estimular al sistema inmunitario.
NS2A	22kDa	Promueve el ensamblaje y la replicación viral. Coordina de un modo en el que aún no se describe por completo si el RNA genómico producido en cada ciclo de replicación se utiliza como nueva plantilla para generar las formas replicativas o si se asocia dentro de la nucleocápside durante el ensamblaje viral.
NS2B	14kDa	Posee una región hidrofóbica que ancla a la membrana del retículo endoplásmico el complejo NS2B/NS3. Por un procesamiento proteolítico, un pequeño dominio hidrofílico de NS2B recién liberado interactúa con el dominio proteasa de la proteína NS3 para actuar como cofactor de ésta.
NS3	70kDa	Proteína bipartita que posee en el extremo N-terminal un dominio proteasa similar a la tripsina (NS3pro) y en el extremo C-terminal posee un dominio con diferentes actividades enzimáticas, que actúa como trifosfatasa de nucleótidos estimulada por RNA (NTPasa) y como helicasa del RNA (NS3Hel); ambas funciones son indispensables en la replicación viral. El dominio NS3Pro actúa hidrolizando los complejos NS2A/NS2B, NS2B/NS3, NS3/NS4A y NS4B/NS5 del polipéptido. La función del dominio NS3Pro depende de su asociación con NS2B, que le confiere estabilidad durante su actividad proteolítica, mientras que la función helicasa permanece inhibida. NS3Hel actúa como helicasa al desenrollar las estructuras secundarias que se forman en el extremo 3' del RNA viral, para favorecer la unión de la polimerasa

		NS5 sobre el RNA y dar inicio a la replicación. NS3 también genera el ambiente lipídico apropiado alrededor del retículo endoplásmico, al reclutar enzimas celulares de la vía de síntesis de lípidos, lo que garantiza el inicio del ensamblaje.
NS5		Proteína multifuncional. En el extremo N-terminal posee actividad enzimática de metiltransferasa y guanidiltransferasa, responsables del <i>capping</i> y la metilación del extremo 5' del RNA genómico. En el extremo C-terminal, se ubica el dominio de la RNA polimerasa dependiente de RNA. Aunque estos procesos suceden exclusivamente en el citoplasma de la célula infectada, se ha identificado una señal de localización nuclear en la proteína NS5 que facilita su importación al núcleo. La posible función de NS5 en el núcleo es desconocida.



**Figura 1.7.** Topología propuesta para las proteínas estructurales y no estructurales del DENV en el retículo endoplásmico. Los dominios transmembranales se muestran como cilindros blancos y los dominios asociados con la membrana, como cilindros grises. Los sitios reconocidos y cortados por la proteasa celular furina se señalan con una flecha blanca y los sitios de clivaje hechos por la NS3 se muestran con flechas negras. (Tomada de Velandía & Castellanos, 2011).

## 1.5 Ciclo de transmisión del DENV

El virus del DENV es mantenido en dos ciclos de transmisión que en apariencia son ecológica y evolutivamente distintos (figura 1.8), uno humano y otro selvático (Vasilakis *et al.*, 2011):

- Ciclo humano.** Involucra a la especie de mosquito doméstico *Aedes aegypti* y a la especie peridoméstica *Aedes albopictus* (cada vez con un área de distribución mayor. Zhang *et al.*, 2010). Aunque hasta hace poco se consideraba que en este ciclo, el ser humano era el único hospedero reservorio y de amplificación conocido (Vasilakis *et al.*, 2011), ahora se sabe que

el mosquito también funge como reservorio del DENV.

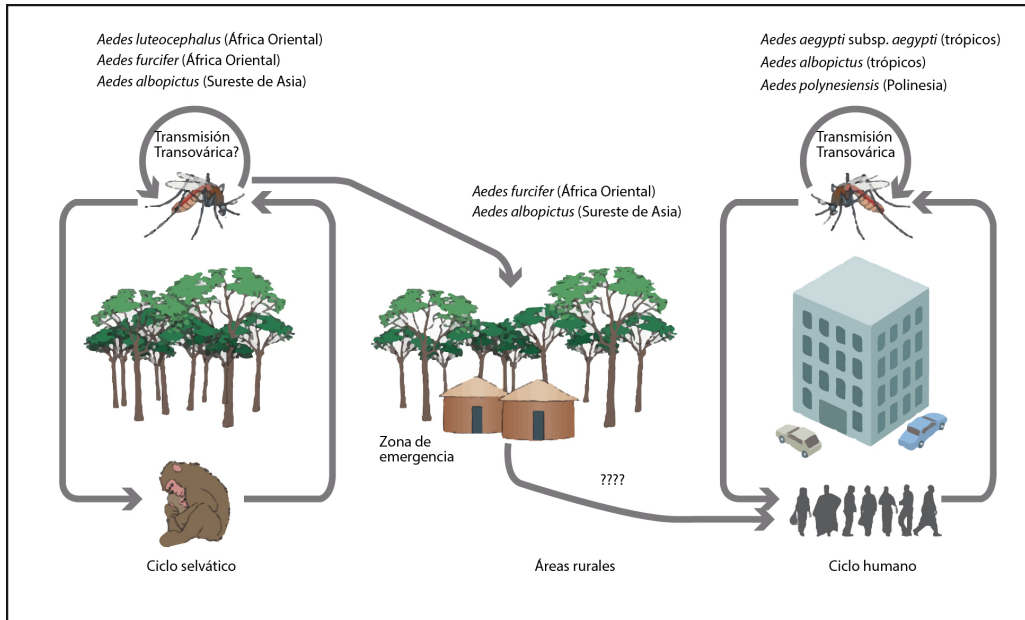
En este ciclo se presentan dos mecanismos de transmisión:

- *Horizontal (humano-mosquito-humano)*. La forma de transmisión mejor conocida y documentada.
- *Vertical o transovárica (mosquito hembra-progenie)*. Mecanismo reportado en la naturaleza desde 1984 por Hull y cols., aunque se corroboró de manera contundente mediante el empleo de métodos moleculares en fechas recientes (Le Goff *et al.*, 2011; Martinis *et al.*, 2012; Espinosa *et al.*, 2014; Cruz *et al.*, 2015). Éste puede constituir uno de los mecanismos por los que el DENV persiste en la naturaleza en ausencia de brotes epidémicos (Le Goff *et al.*, 2011).

- a) **Ciclo selvático**. En él, mosquitos arbóreos transmiten el DENV a primates no humanos que experimentan infecciones asintomáticas (Holmes & Twiddy, 2003), viremias transitorias que no generan el exantema ni los síntomas hemorrágicos característicos del DF o el DHF (Rico-Hesse, 2003).

Aunque existen claras evidencias de que ocurren interacciones entre ciclos (por ejemplo, en Senegal, Diallo y cols., 2003, se demostró que una variante selvática del DENV entra en contacto regular con los humanos, siendo *Aedes furcifer* el vector puente entre hábitats forestales y peridomésticos); no se ha determinado su frecuencia o amplitud ni, como consecuencia, el impacto que éstas puedan tener en la salud pública.

En América no se han reportado aislados del DENV que muestren evidencia filogenética de un linaje selvático. Sin embargo, a partir de suero de distintas especies de mamíferos sudamericanos (murciélagos, roedores y marsupiales) se han aislado y secuenciado variantes de los cuatro serotipos que atacan a humanos (De Thoisy *et al.*, 2009). El papel que los mamíferos no primates puedan tener en el ciclo de transmisión humano y/o selvático es desconocido.



**Figura 1.8.** Ciclos de transmisión del DENV. Los orígenes selváticos del DENV y la “zona de emergencia” donde los ciclos selváticos contactan poblaciones humanas en áreas rurales de África Oriental y el sureste asiático. Además, el DENV puede persistir en poblaciones de mosquito por transmisión transovárica (TOT), en la que mosquitos infectados con el virus transfieren el virus a sus huevos (esto se ha visto que ocurre en algunas especies, pero no en todas). (Tomada de Vasilakis *et al.*, 2011.)

## 1.6 Ciclo de vida del *Aedes aegypti*

Las hembras de *Ae. aegypti* necesitan alimentarse de sangre para lograr la maduración de sus huevecillos. Una vez que esto ocurre, la hembra produce entre 50 y 100 huevecillos en cada ovipostura, que deposita en cuerpos de agua ubicados en distintos lugares (Sinave, DGE & Salud, 2012). Los huevecillos pueden ser resistentes a la desecación hasta por varios meses (diapausa), lo que facilita la supervivencia en épocas adversas, de sequía. En los cuerpos de agua ocurren los eventos de eclosión de los huevecillos y la transformación y maduración de las formas de larva y pupa, que en conjunto tienen una duración de siete a 10 días. La pupa es la última fase acuática; una vez que emerge continúa la fase aérea adulta. La alimentación comienza entre las 20 y 72 horas posteriores a este evento. Las hembras se alimentan aproximadamente cada tres días. El intervalo entre la ingesta de sangre y la ovipostura puede ser también de sólo tres días. A diferencia de otras especies de mosquitos, *Ae. aegypti* pica o se alimenta varias veces de uno o varios humanos hasta satisfacer sus necesidades alimenticias (2 a 3 mg de sangre). En condiciones naturales, *Ae. aegypti* sobrevive en promedio de cuatro a ocho semanas (Sinave, DGE & Salud, 2012).

## 1.7 Replicación del DENV en *Aedes aegypti*

La replicación del DENV comienza después de que una hembra de mosquito ingiere sangre de un paciente infectado durante el periodo de viremia (que dura generalmente cinco días) (Guzmán, 2010). Esta replicación inicia en el epitelio de la región posterior del intestino medio (*midgut*) y subsecuentemente se disemina a otros órganos vía la hemolinfa, para finalmente infectar las glándulas salivales. El DENV aparece en estas glándulas luego de un periodo de incubación extrínseca de ocho a 12 días, cuando el mosquito se vuelve capaz de transmitirlo a otro huésped humano (Lin *et al.*, 2004). El virus es secretado a la saliva del mosquito e inyectado al hospedero humano en eventos de alimentación subsecuentes. Los mosquitos permanecen en vida libre entre dos y tres semanas, e infectivos de por vida (Sim & Hibberd, 2016; Salazar *et al.*, 2007), aunque el DENV parece no ejercer un costo en el *fitness* del mosquito durante las infecciones naturales (Carrington *et al.*, 2015); la infección que les produce no es citolítica (Lindenbach & Rice, 2001).

La competencia del mosquito, es decir, su capacidad intrínseca para ser infectado por un patógeno, soportar su replicación y transmitirlo, varía ampliamente entre y dentro de poblaciones de mosquitos, está genéticamente determinada y también es influida por factores ambientales (Bennett *et al.*, 2002, Lozano-Fuentes *et al.*, 2009; Sylla *et al.*, 2009; Sim *et al.*, 2013).

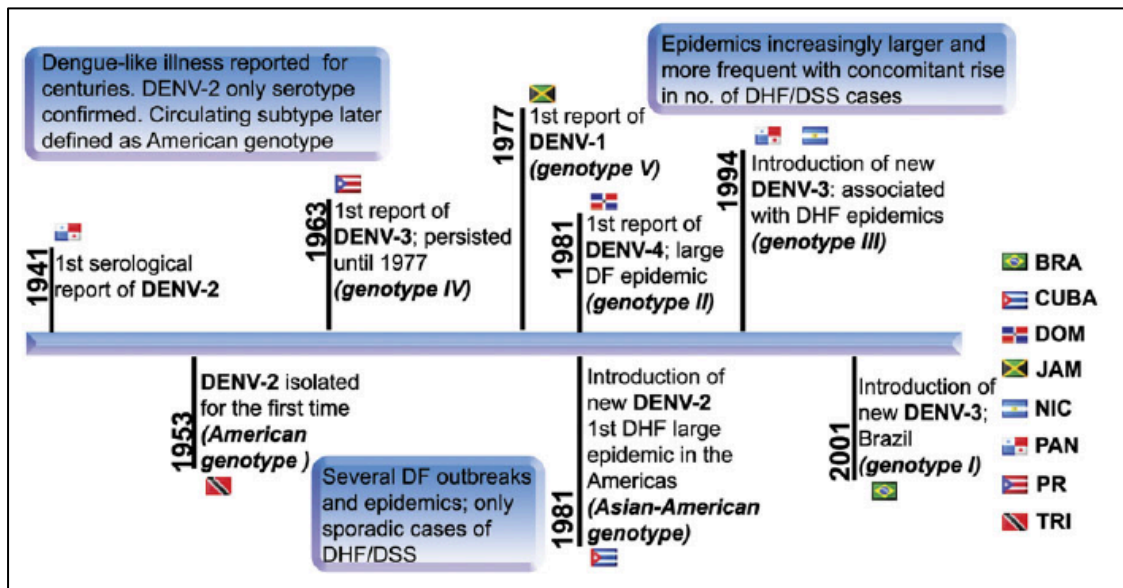
## 1.8 Diversidad del DENV en América

En América, el registro más antiguo de una enfermedad con síntomas que sugieren dengue data de 1780, en Filadelfia, aunque los brotes de ésta eran más bien raros. Antes de 1963 sólo se tiene reporte del genotipo Americano de DENV-2, pero desde entonces la región ha estado sujeta a introducciones repetidas de otros genotipos (figura 1.9) (Allicock *et al.*, 2012).

Durante las décadas de 1950 y 1960, la Organización Panamericana de la Salud promovió una campaña para erradicar el *Ae. aegypti*, que resultó en la desaparición de DF en la mayor parte de América Central y Sudamérica. Sin embargo, DENV-2 y DENV-3 permanecieron en algunas islas del Caribe durante esos años. En la década de 1970, la campaña para erradicar al vector se redujo gradualmente y *Ae. aegypti* reapareció en diversos países. El dengue apareció poco después y, en 1977 y 1981, respectivamente, fueron introducidos DENV-1 y DENV-4, causando brotes epidémicos en la región. El DF fue la presentación clínica más frecuente, aunque también se reportaron algunos casos severos. DFH/DS emergieron abruptamente en 1981 en Cuba y causaron cerca de 150,000 hospitalizaciones y 150 muertes en menos de cuatro meses. Un

programa intensivo de erradicación detuvo el brote y el DHF no se dispersó fuera de ese país. DHF reapareció en Venezuela en 1989 y poco después fue detectado en otros países de Sudamérica. DHF no fue diagnosticado en América Central hasta 1994, concurrentemente con la reintroducción de DENV-3, que no había sido detectado en América desde 1977 (Díaz *et al.*, 2006).

Después de la introducción del genotipo Asiático-Americano de DENV-2 en 1981 hubo un incremento en el número de casos reportados de DHF (4.5 veces) y DSS. Actualmente, los cuatro serotipos están presentes en América y, desde 1963, la región ha experimentado brotes causados por al menos un genotipo de DENV-1, dos genotipos de DENV-2, tres genotipos de DENV-3 y uno de DENV-4 (Allicock *et al.*, 2012).



**Figura 1.9.** Línea de tiempo que muestra la introducción de varios genotipos de los cuatro serotipos del DENV. Las banderas indican los países asociados a los primeros reportes. (Tomada de Allicock y cols. 2012).

Allicock y cols. (2012) delinean en términos generales el comportamiento histórico del DENV en América. Señalan que los cuatro serotipos del DENV han experimentado una fase de invasión inicial caracterizada por un incremento rápido en la diversidad genética, coincidente con los primeros casos confirmados de cada genotipo en la región. Posteriormente ha ocurrido una rápida dispersión geográfica, después de la cual los linajes individuales se mantienen localmente, estableciendo un flujo génico con localidades más distantes.

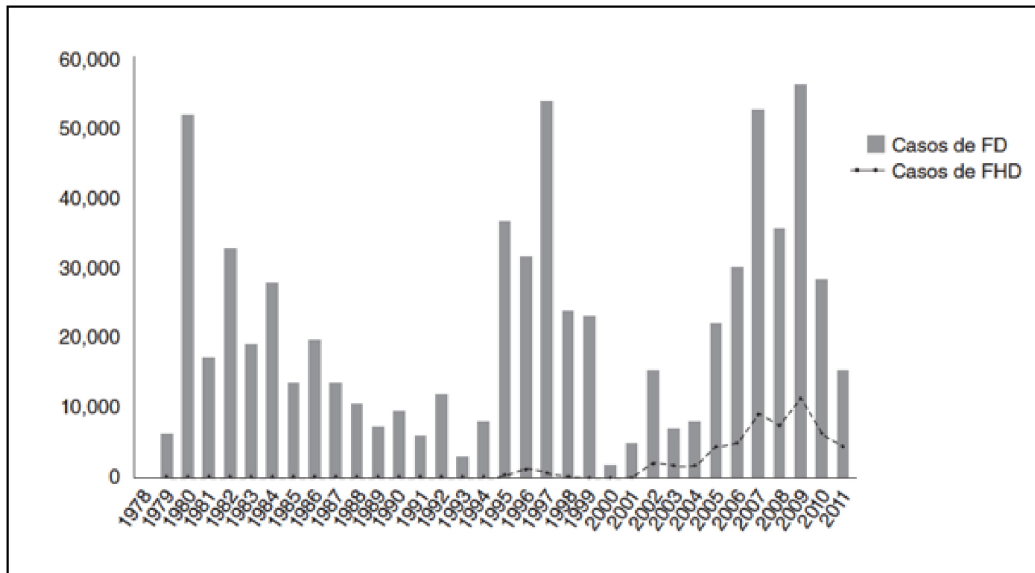
## 1.9 Diversidad del DENV en México

De 1947 a la década de 1960, las autoridades de salud (en sincronía con la estrategia continental de erradicación del vector) condujeron una campaña intensiva para controlar el *Ae. aegypti*, y el mosquito fue declarado erradicado del país en 1963 (Torres-Galicia *et al.*, 2014). Aunque el mosquito reapareció sólo dos años después, el país estuvo libre de dengue desde el comienzo de la década de 1960 hasta 1978 (Gómez-Dantés *et al.*, 2014).

Desde entonces se han reportado cuatro brotes epidémicos (tabla 1.3) (Torres-Galicia *et al.*, 2014) con un notorio aumento en la proporción de casos de DHF y DSS del total de casos de dengue reportados en la última década (figura 1.10) (Díaz *et al.*, 2006).

**Tabla 1.3.** Datos de los brotes epidémicos de dengue ocurridos en México. Construida a partir de los datos de Torres-Galicia *et al.*, 2014.

Año	Casos totales/Casos DHF	Serotipo predominante
1980	51,406	DENV-1
1997	53,541 / 9,433	DENV-3
2007	52,369 / 9,433	DENV-1
2009	55,961 / 11,396	DENV-1 y DENV-2



**Figura 1.10.** Total de casos de DF y DHF en la República Mexicana reportados por la Secretaría de Salud durante 1984-2011 (adaptado de Cenavece/Boletines epidemiológicos 1984-2011). Tomada de Torres-Galicia y cols. (2014).

El área geográfica de transmisión del DENV equivale al área de distribución del insecto vector (Ibáñez & Gómez, 1995). Gómez-Dantés y cols. (2014) afirman que México es un país particularmente vulnerable al dengue debido a que presenta muchas de las características que favorecen la dispersión de los mosquitos transmisores, entre las que destacan: ubicación geográfica en región tropical y subtropical, elevada densidad poblacional, crecimiento demográfico acelerado (de 97 millones en el año 2000 a más de 112 millones en 2010), rápida urbanización durante los últimos 20 a 30 años, elevados niveles de turismo nacional y extranjero, así como migración intensiva desde países centroamericanos. Ibáñez y Gómez (1995) mencionan que otro factor determinante en la dispersión del vector del DENV en nuestro país radica en la dificultad que representa elevar la calidad de vida de un gran número de personas en términos de vivienda, cobertura de servicios (drenaje, agua potable con disponibilidad continua en cada domicilio, disposición adecuada de desechos y basura) y educación higiénica.

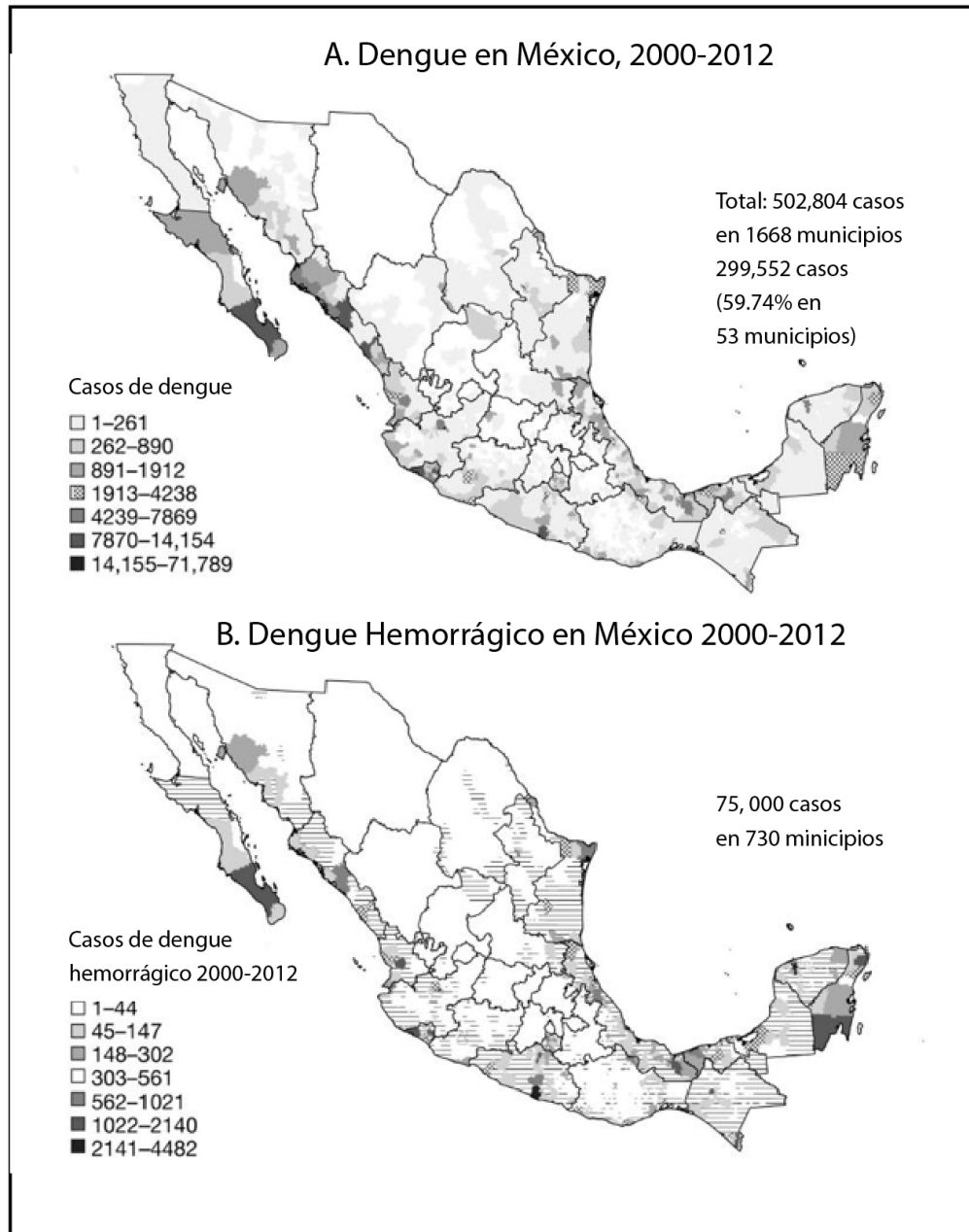
El factor ambiental que limita la distribución de *Ae. aegypti* es la temperatura, y en un país con intrincada orografía como el nuestro, ésta se asocia directamente con la altitud. Para 1987, Ibáñez-Bernal reporta por primera vez la presencia de *Ae. aegypti* a una altitud de 1,630 msnm, y en 1988 se registra un brote de dengue en Taxco, Guerrero, a 1,735 msnm. En 2009, Lozano-Fuentes y cols. reportan la presencia moderadamente abundante de mosquitos *Ae. aegypti* a 1,700 msnm y la presencia (aunque rara) a altitudes de hasta 2,300 msnm en el estado de Veracruz.

La dispersión del dengue y el DHF a través del territorio muestra un patrón claramente regional de incidencia, con elevadas concentraciones de casos en las áreas costeras y tropicales (figura 1.11), pues 50% de los casos de dengue se concentran sólo en 65 municipios (Gómez-Dantés *et al.*, 2014).

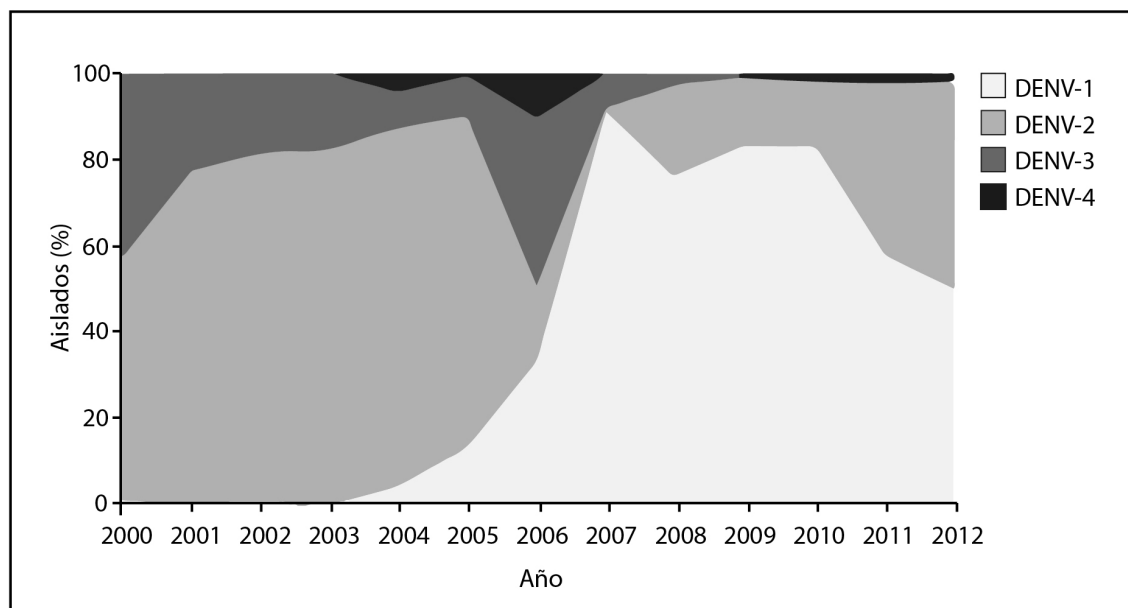
Se han observado dos patrones epidemiológicos distintos de transmisión: uno endémico, en los estados del Golfo de México, y otro estacional, en la costa del Pacífico y la Península de Yucatán, con la mayoría de los casos con fecha entre septiembre y diciembre. De manera histórica se ha asumido que la llegada de la estación lluviosa (mayo a octubre), desencadena el comienzo del periodo de reproducción de *Ae. aegypti*, pero los datos de años recientes sugieren que éste ha comenzado a reproducirse durante todo el año (Gómez-Dantés *et al.*, 2014).

En México, aunque en distintos momentos han circulado los cuatro serotipos del DENV, creando escenarios epidémicos, endémicos e hiperendémicos (figura 1.12), a pesar de la importancia que tiene el dengue en la salud pública de México, ha habido relativamente pocos estudios de su epidemiología molecular (Carrillo-Valenzo *et al.*, 2010). Sin embargo, los que se han realizado permiten conocer en términos generales el comportamiento del DENV en territorio nacional. De manera análoga a lo que ocurre en el continente americano, se ha observado que la

diversidad del DENV se caracteriza por eventos sucesivos de introducción de linajes virales y reemplazo de linajes preexistentes (Díaz *et al.*, 2006; Carrillo-Valenzo *et al.*, 2010), de tal manera que existe poca cocirculación y predominancia de un genotipo dentro de cada serotipo específico (Gardella-García *et al.*, 2008; Rivera-Osorio *et al.*, 2011).



**Figura 1.11.** Distribución geográfica de los casos de fiebre de dengue y fiebre de dengue hemorrágico, 2000-2012. Las poblaciones regionales en México están expuestas a diferentes magnitudes de riesgo de dengue debido a que hay áreas epidémicas, endémicas e hiperendémicas. **(B)** Casos de dengue. **(C)** Casos de dengue hemorrágico con números elevados en centros turísticos y comerciales importantes en áreas costeras y tropicales. (Imagen tomada de Gómez-Dantés *et al.*, 2014.)



**Figura 1.12.** Circulación de serotipos del DENV en México, 2000-2011. (Imagen tomada de Gómez-Dantés y cols., 2014.)

### 1.10 Diversidad del DENV en el estado de Guerrero

Para el periodo de 1995 a 2012, Guerrero se encontró en la categoría de los cinco estados con mayor número de casos de DF y de DHF (Sinave, DGE & Salud, 2012). Aunque para el periodo de 2013 a 2014, la incidencia del dengue se redujo en ese estado, continuó dentro de las 10 entidades con mayor número de casos de DF y DHF. A la semana epidemiológica No. 32 de 2015, Guerrero ocupa el segundo lugar en el número de casos de DF y el primer lugar en DHF (Sinave / DGE / Salud/ Sistema Especial de Vigilancia Epidemiológica de Dengue).

Los serotipos que se han registrado en los sueros de pacientes de Guerrero durante el periodo de 2010 a 2014 se muestran en la tabla 1.4.

**Tabla 1.4.** Frecuencia relativa de los serotipos del DENV aislados de sueros de pacientes del estado de Guerrero registrados como casos de DF y DHF. Fuente: Sinave / DGE / Salud / Sistema Especial de Vigilancia Epidemiológica de Dengue.

Año	2010	2011	2012	2013	2014
<b>DENV-1</b>	99.48%	90.32%	38.88%	43.62%	66.92%
<b>DENV-2</b>	-	9.67%	58.58%	54.90%	27.69%
<b>DENV-3</b>	0.52%	-	-	-	-
<b>DENV-4</b>	-	-	2.52%	1.47%	5.38%

## **2. Antecedentes**

### **2.1 Serotipos y genotipos del DENV presentes en México**

Como ocurre a escala global, el conocimiento que se tiene sobre la diversidad del DENV en México se deriva fundamentalmente del análisis de muestras provenientes de hospedero humano, en particular de casos sintomáticos con relevancia clínica.

Díaz y cols. (2006) analizaron la secuencia del gen E de aislados de pacientes obtenidos entre 1980 y 2002 en México. Señalan que DENV-1 está presente en el país desde finales de la década de 1970, mientras que DENV-2 y -4, desde comienzos de la década de 1980, aunque su presencia no fue detectada cada año. Esos mismos autores encuentran que la circulación del DENV en nuestro país se caracteriza por episodios frecuentes de introducción, expansión, extinción aparente y reintroducción del mismo o de distintos genotipos virales. También afirman que la selección positiva contribuye en una mínima proporción, al tiempo que la deriva génica y el flujo genético (debido a las múltiples introducciones de cepas extranjeras) son las fuerzas predominantes en la evolución del DENV en México.

Gardella-García y cols. (2008) encuentran que el serotipo predominante en los brotes de DF y DHF de 2005 y 2006 en Oaxaca y de 2006 en Veracruz fue DENV-2, con 100% de las secuencias pertenecientes al genotipo Asiático-Americano.

En secuencias de DENV-2 provenientes de pacientes del estado de Oaxaca, Pérez-Ramírez y cols. (2009) detectaron recombinación entre genotipos Asiático-Americano y Cosmopolita, y entre Asiático-Americano y Americano. Este fue el primer reporte de recombinación de DENV-2 en México. Sin embargo, no se ha establecido la importancia de la recombinación en la evolución de este virus.

Carrillo-Valenzo y cols. (2010) generalizan las características de la diversidad del DENV en México. Señalan que para cada uno de los cuatro serotipos han existido múltiples introducciones de linajes virales con DENV similares a aquellos observados a lo largo del continente americano, aunque en presencia de poca cocirculación. Más bien, el DENV en México

se caracteriza por un frecuente reemplazo de linajes, tal que sólo un linaje viral domina en un serotipo específico en un momento dado.

Rivera-Osorio y cols. (2011) reportan que de 100 secuencias de DENV-1 (el serotipo circulante más común para el periodo de estudio) provenientes de suero de pacientes con manifestaciones clínicas de dengue en el periodo de enero de 2006 a diciembre de 2007 de 22 estados distintos (elegidas aleatoriamente), 100% correspondió al genotipo III.

## **2.2 Diversidad de DENV aislados de *Aedes aegypti* en México**

García-Rejón y cols. (2011) determinan la tasa de hembras *Ae. aegypti* infectadas con DENV a partir de mosquitos colectados en escuelas de Mérida, Yucatán. Identifican también los serotipos de DENV presentes en los *pools* de mosquitos. Señalan que el porcentaje de *pools* de mosquitos positivos para DENV no se mantuvo estable, sino que osciló en un rango de 3.7% en diciembre de 2008 a 20.9% en noviembre-diciembre de 2009. También indican que los serotipos presentes en *pools* positivos mostraron un cambio de DENV-1 en 2008 a DENV-2 y DENV-3 en 2009, y que datos de pacientes en Yucatán mostraron similitudes en el incremento de infecciones causadas por DENV-2, en relación con las causadas por DENV-1 de 2008 a 2009. Sin embargo, los autores no muestran las cifras ni las fuentes que respaldan esa aseveración.

Espinosa-Gómez y cols. (2011) realizaron una prospección entomológica buscando Flavivirus (entre los que se incluye el DENV) en mosquitos para evaluar su potencial de transmisión en Colima (una zona hiperendémica del DENV). Sin embargo, de los 60 *Ae. aegypti* capturados y analizados, no se obtuvo ninguna secuencia del DENV, aunque sí de otros Flavivirus que no tienen rol patogénico en humanos.

Martínez y cols. (2014) llevaron a cabo estudios pioneros en nuestro país relacionados con la búsqueda del DENV en poblaciones naturales de mosquitos *Aedes aegypti* en los distintos estados de su ciclo vital. Adultos, larvas y pupas fueron colectados en distintas colonias del municipio de Acapulco, Guerrero, durante la estación lluviosa de 2011, en el interior o la cercanía de hogares con casos recientes de dengue. Los autores reportan la presencia de DENV-1 en larvas, constituyendo éste el primer registro de transmisión transovárica de DENV-1 en poblaciones naturales de *Aedes aegypti* del país.

Dzul-Manzanilla y cols. (2015) buscaron DENV y virus de chikungunya (CHIKV) también en poblaciones naturales de *Ae. Aegypti* en el estado de Guerrero durante el periodo de 2012 a 2014. Ellos reportan la presencia de DENV-1, DENV-2 y DENV-4 para 2012 y únicamente

DENV-1 durante 2013 y 2014; y también la presencia de CHIKV en mosquitos colectados en Acapulco y Juchitán en 2014.

### **2.3 Variación intrahospedero del DENV en mosquito y/o humano**

La intención primaria de los estudios sobre diversidad genética intrahospedero fue estimar su magnitud en ambos hospederos y encontrar una posible correlación con aspectos de relevancia clínica, epidemiológica o filogenética. Sólo durante los últimos dos años, los estudios han comenzado a monitorear la dinámica evolutiva de las poblaciones de DENV a nivel intrahospedero (Lambrechts & Lequime, 2016).

La diversidad intrahospedero del DENV fue inicialmente demostrada por transcripción reversa de RNA viral purificado de muestras individuales de plasma humano, seguidos de clonación y secuenciación con ampliaciones de PCR para los genes E (Wang *et al.*, 2002a), C y NS2B (Wang *et al.*, 2002b) de DENV-3.

La mayoría de los estudios subsecuentes se enfocó en el hospedero humano (Chao *et al.*, 2005, DENV-3; Descloux *et al.*, 2009 y Thai *et al.*, 2012, DENV-1; Parameswaran *et al.*, 2012, DENV-2), aunque también se ha estudiado la diversidad genética intrahospedero en mosquitos.

Lin y cols. (2004) analizaron la variabilidad del gen E (dominio III) y concluyeron que la diversidad es mayor en el DENV aislado de humanos que en el de mosquitos (aunque estudios posteriores no sustentan este hecho, p. ej. Aaskov *et al.*, 2006 y Sim & Hibberd., 2016). Los autores también describen la dramática reducción de la variabilidad (>90% de las variantes) que ocurre durante la transmisión del virus desde el humano al abdomen del mosquito y de éste hacia las glándulas salivales.

Aaskov y cols. (2006) reportaron que linajes de DENV-1 que adquirieron una mutación de codón de paro en el gen E se detectaron tanto en mosquitos como en humanos, en un periodo de 18 meses en Myanmar. Proponen que los genomas defectuosos, en los que largas secciones han sido deletadas, tienen una ventaja de replicación sobre los genomas parentales, pero sólo se replican por complementación en la presencia de virus coinfectantes funcionales.

Lambrechts y cols. (2011) analizaron el papel que tienen las interacciones DENV-mosquito en la evolución del DENV, probando la capacidad de 23 aislados de DENV entre 1985 y 2009, para ser transmitidos en la población actual de *Aedes aegypti* de Tailandia. Los autores encontraron que un evento principal de reemplazo de clado a mediados de la década de 1990 se asocia con aislados del DENV que exhiben mayores títulos virales en el hemocele del mosquito,

lo que podría traducirse en una mayor probabilidad de transmisión. Este descubrimiento es consistente con la hipótesis de que la selección de una transmisión mejorada es uno de los posibles mecanismos que subyacen en los principales eventos de reemplazo de clados (tan característico del DENV).

Sim y cols. (2015) rastrearon cambios en la diversidad genética intrahospedero de DENV-1 durante transmisiones humano-mosquito. Los pacientes virémicos de Vietnam fueron expuestos experimentalmente a mosquitos *Ae. aegypti* vírgenes. Estos autores compararon la diversidad genética intrahospedero en muestras de DENV aislados de humano y mosquito, lo que reveló cambios sustantivos en el repertorio de variantes entre sangre venosa y los órganos de los mosquitos que fueron infectados al alimentarse de la persona correspondiente. La mayoría de las variantes genéticas se perdió durante la transición de sangre venosa al abdomen del mosquito, así como del abdomen del mosquito a las glándulas salivales. Sin embargo, varias medidas de diversidad genética por muestra fueron similares entre plasma humano, abdomen de mosquito y glándulas salivales del mosquito, indicando que el nivel general de diversidad genética intrahospedero se mantuvo.

Sessions y cols. (2015) examinaron la diversidad genética en 12 muestras de sueros humanos pareadas colectadas en momentos tempranos y tardíos (con cuatro días de diferencia) durante el curso de la infección, y tanto en *Ae aegypti* como en *Ae albopictus* después de inoculación intratorácica con las mismas muestras de suero. Los autores detectaron *hotspots* y *coldspots* de mutaciones en el genoma viral, algunos de los cuales fueron independientes del tipo de hospedero. Adicionalmente hubo evidencia de reversiones frecuentes de variantes de baja frecuencia durante la transmisión humano-mosquito, y algún grado de evolución mosquito-especie-específico.

Lequime y cols. (2016) investigaron las fuerzas evolutivas que subyacen en la evolución intrahospedero en diferentes genotipos del vector *Aedes aegypti*. Estos autores usaron secuenciación masiva para monitorear la diversidad genética de genomas completos durante una infección de mosquito por aislados *wild type* de DENV-1. Como resultado, detectaron una reducción estocástica en el tamaño de la población en la infección inicial, siendo de cinco a 42 los genomas fundadores que lograron ser infectivos en el intestino medio (dos órdenes de magnitud menor que la dosis del virus ingerido por mosquitos individuales en su experimento). El nivel de diversidad genética observado en este estudio incrementó sustancialmente durante la expansión de la población en el *midgut* sin evidencia de selección positiva. Se observó que al igual que la diversidad genética viral, la cantidad de diversidad genética viral difirió dependiendo de los *backgrounds* genéticos de los mosquitos.

Hasta donde esta investigación fue posible, no se ha realizado ningún análisis de la diversidad de DENV provenientes de mosquitos circulantes en poblaciones naturales (capturados en campo y no alimentados en el laboratorio, ya sea de manera natural o intratorácica) en ningún país. Tampoco se ha llevado a cabo una comparación entre la variabilidad presente en los DENV circulantes en poblaciones naturales de mosquitos y en humanos con la misma distribución geográfica y temporalidad.

### 3. Justificación

Dado que:

- La diversidad de los DENV circulantes en mosquitos puede repercutir directamente en los patrones de transmisión hacia el ser humano (*e.g.* sustitución de clados).
- Se desconocen los patrones de diversidad de los DENV en las poblaciones de mosquitos que coexisten con humanos que presentan manifestaciones clínicas de la enfermedad.
- La secuenciación del genoma completo del DENV es ahora accesible y permite el análisis diferencial de las fuerzas evolutivas que pueden actuar en las distintas regiones de su genoma.

Se propone realizar el análisis de la diversidad de genomas completos de DENV extraídos de mosquitos *Aedes aegypti* de poblaciones naturales (colectados en campo), así como su comparación con DENV extraídos de humanos infectados y con manifestaciones clínicas coexistentes en la misma región geográfica.

## **4. Hipótesis**

La diversidad de los DENV circulantes en mosquitos reflejará la diversidad de los DENV circulantes en pacientes con manifestaciones clínicas de la misma región geográfica y temporalidad.

## **5. Objetivos**

### **5.1 Objetivo general**

Determinar la variabilidad de genomas completos de DENV circulantes en mosquitos *Aedes aegypti* en el estado de Guerrero.

### **5.2 Objetivos particulares**

1. Obtener los serotipos y genotipos de los DENV secuenciados, circulantes en mosquitos *Aedes aegypti* en el estado de Guerrero.
2. Obtener las secuencia del genoma completo de DENV circulantes en mosquitos *Aedes aegypti* en el estado de Guerrero.
3. Obtener la filogeografía y la estructura poblacional de DENV secuenciados, circulantes en mosquitos *Aedes aegypti* en el estado de Guerrero.
4. Determinar el papel de la selección natural en la diversidad de los genomas de DENV secuenciados, circulantes en mosquitos *Aedes aegypti* en el estado de Guerrero.

## 6. Materiales y métodos

Para alcanzar los objetivos del presente estudio se requirió de la participación de distintas instituciones. La figura 6.1 muestra el plan general de trabajo seguido.

### 6.1 Muestras

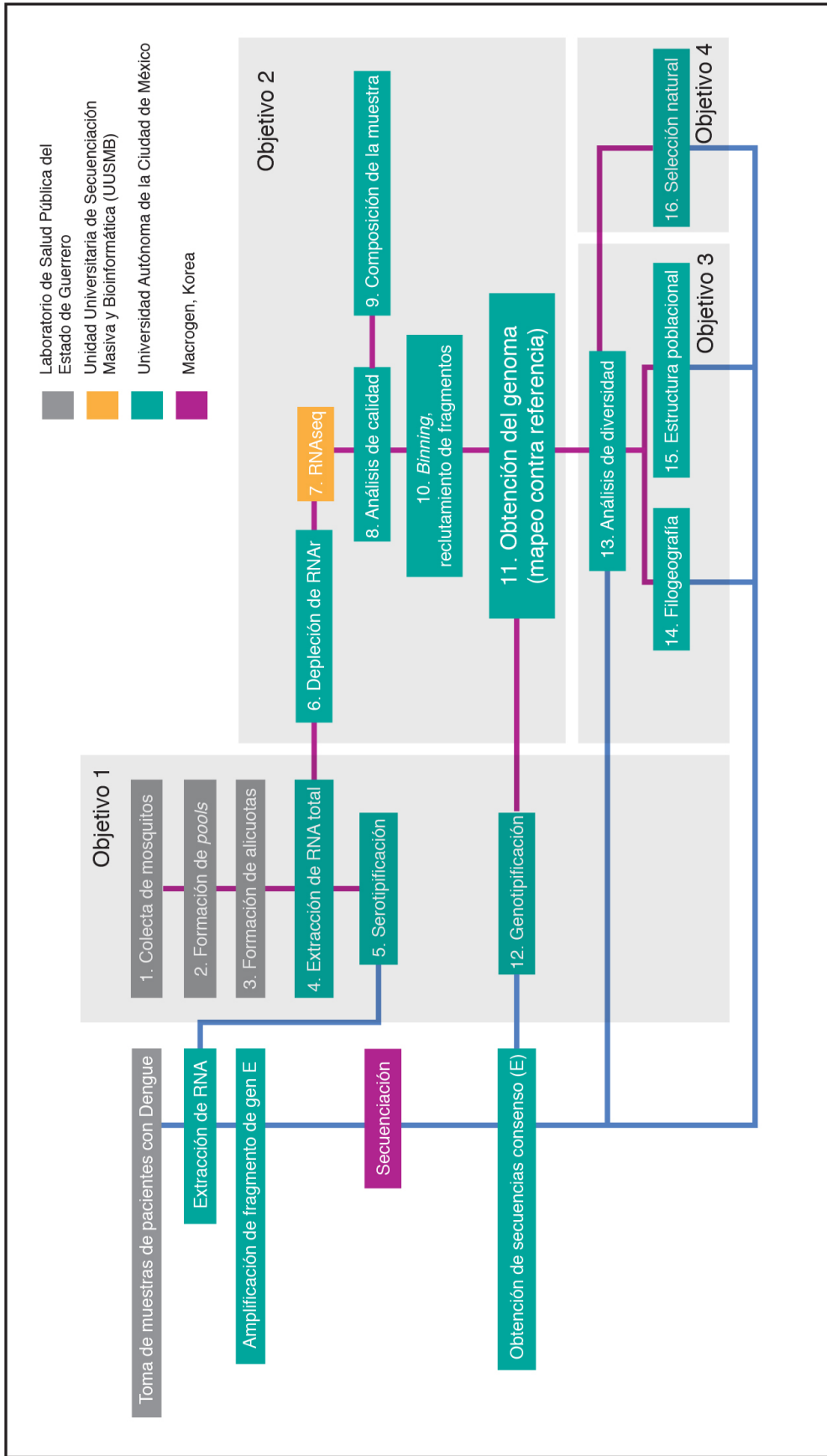
#### 6.1.1 Pools de mosquitos

Las muestras de mosquitos con las que se efectuó el presente estudio fueron donadas por el Laboratorio Estatal de Salud Pública del Estado de Guerrero a la doctora Martha Yocupicio (UACM) y consisten en siete alícuotas de *pools* de mosquitos adultos hembra de *Aedes aegypti* colectados en el municipio de Acapulco de Juárez (figura 6.2) durante la época lluviosa de 2012, preservadas en medio esencial mínimo suplementado (MEM) a  $-70^{\circ}\text{C}$ .

Las colectas se llevaron a cabo en el interior de hogares con reportes de casos confirmados y/o probables de dengue y sus cercanías, mediante el uso de aspiradores portátiles de mochila CDC.

Los *pools* se formaron con 15 a 35 especímenes de mosquitos provenientes de las cercanías del mismo punto de colecta. Cada grupo fue homogeneizado en 2,000  $\mu\text{L}$  de MEM con un disruptor de tejidos (Qiagen). De cada *pool* se realizaron alícuotas de aproximadamente 500  $\mu\text{L}$ . Una de las alícuotas de cada *pool* fue analizada por el Laboratorio Estatal de Salud Pública del estado de Guerrero, mientras que el resto fue almacenado para su posterior procesamiento.

La tabla 6.1 contiene los datos de colecta de las muestras de mosquitos analizadas en el presente estudio.



**Figura 6.1.** Plan general de trabajo seguido para alcanzar los objetivos del presente estudio en coordinación con distintas instituciones.

**Tabla 6.1.** Datos de colecta de las muestras de *pools* de mosquitos analizadas en el presente estudio.

<b>No. de muestra</b>	<b>Fecha de colecta</b>	<b>Colonia del municipio de Acapulco donde se hizo la colecta</b>	<b>Tipo de caso de dengue</b>	<b>Número de hembras por <i>pool</i></b>
M1	18/07/2012	Morelos	caso confirmado	30
M2	26/07/2012	Coloso	caso confirmado	35
M3	12/10/2012	La Garita	caso probable	33
M4	23/10/2012	Jardín Mangos	caso probable	22
M5	23/10/2012	Jardín Mangos	caso probable	22
M6	07/11/2012	Sta. Rosita	caso probable	23
M7	07/11/2012	Col. Centro	caso probable	15

De acuerdo con los datos proporcionados por el laboratorio donante, a partir del análisis realizado a la alícuota que se procesó, el serotipo de los DENV presentes en los *pools* de procedencia corresponde a DENV-2. Sin embargo, dada la naturaleza heterogénea de la muestra y por sugerencia del propio laboratorio, fue necesario corroborar la positividad.



**Figura 6.2.** Sitios de colecta de mosquitos y toma de muestras analizados en el presente estudio.

### 6.1.2 Sueros de pacientes infectados con DENV

El Laboratorio Estatal de Salud Pública del Estado de Guerrero realizó una donación de sueros positivos a la infección con DENV a la doctora Martha Yocupicio (UACM). Las muestras corresponden a pacientes de distintos municipios de la entidad (figura 6.2) que presentaron cuadros clínicos de dengue clásico o hemorrágico. De las muestras donadas del año 2012 se seleccionaron 16 positivas para DENV-2 que fueran coincidentes con la temporalidad de colecta de las muestras de mosquitos. Los datos de colecta de las muestras seleccionadas se presentan en la tabla 6.2.

**Tabla 6.2.** Datos de procedencia de las muestras de DENV-2 obtenidas de sueros humanos del estado de Guerrero.

Muestra	Municipio	Col. / Localidad	F. inicio	F. toma	Clasificación
MYC6	Acapulco	El quemado	19/08/2012	22/08/2012	Hemorrágico
MYC7	Chilpancingo	Acahuizotla	26/09/2012	26/09/2012	Clásico
MYE3	Iguala	Villa de Guadalupe	28/04/2012	30/04/2012	Hemorrágico
MYE4	Iguala	Centro	16/06/2012	19/06/2012	Hemorrágico
MYE5	Acapulco	Luis Donaldo Colosio	17/07/2012	19/07/2012	Clásico
MYE6	Acapulco	La poza	21/08/2012	22/08/2012	Clásico
MYE7	Chilpancingo	Centro	26/09/2013	27/09/2013	Clásico
MYG3	Iguala	Burocracia	10/05/2012	15/05/2012	Hemorrágico
MYG4	Cocula	Epifanio Rodríguez	29/06/2012	03/09/1920	Hemorrágico
MYG5	Iguala	24 de febrero	01/08/2012	05/08/2012	Clásico
MYG6	Coyuca de Benítez	Platanillo	02/09/2012	03/09/2012	Clásico
MYA3	N/A	N/A	31/03/2012	03/04/2012	Hemorrágico
MYA5	Iguala	Iguala	09/07/2012	11/07/2012	Clásico
MYA7	Iguala	Ejidal	06/09/2012	07/09/2012	Clásico
MYC4	Cocula	Ampliación panteón	05/06/2012	07/06/2012	Hemorrágico
MYC5	Iguala	Iguala	09/07/2012	11/07/2012	Clásico

## 6.2 Procesos llevados a cabo en laboratorio

### 6.2.1 Procesamiento de muestras de *pools* de mosquitos

#### 6.2.1.1 Extracción de RNA

Las alícuotas donadas por el Laboratorio de Salud Pública del estado de Guerrero (volumen inicial entre 500  $\mu$ L y 750  $\mu$ L) fueron homogeneizadas con micropipeta; se les agregó TRIzol

(Invitrogen, USA) en una relación de volumen 3:1, mezclando vigorosamente, y se incubaron durante 10 min a temperatura ambiente. Los clarificados se obtuvieron centrifugando a 4,500 rpm durante 5 min y recuperando todo el sobrenadante. A los clarificados se les agregó etanol en una relación de volumen 1:1, mezclando, para posteriormente filtrarlos en columnas Zymo-Spin™ a 8000 rpm durante 30 s; se lavaron con 600 µL de *buffer* de lavado (trisHCl 10mM, EDTA 1mM, etanol al 100%) y se centrifugaron a 8,000 rpm durante 1 min (dos veces). Se eliminó el exceso de *buffer* de lavado centrifugando a 10,000 rpm durante 2 min; se eluyó con 30 µL de H<sub>2</sub>O libre de nucleasas a 10,000 rpm durante 2 min. La concentración de cada extracto de RNA total se cuantificó por espectrofotometría utilizando el equipo NanoDrop 2000 (*Thermo Scientific*).

#### 6.2.1.2 RT-PCR One step singleplex

Por cada muestra se realizaron dos reacciones *singleplex* de RT-PCR *One step* empleando oligonucleótidos de DENV (NS5) (tabla 6.3), sondas TaqMan® para DENV-1 y DENV-2 y el kit *One Step RT-PCR SuperScript III (Invitrogen™)* en el equipo 7300 *Real Time* (Applied Biosystem). Para cada una de las reacciones se empleó 1µL de extracto de RNA al que se agregó 5 µL de Master Mix, 0.48 µL de sonda TaqMan®, 1.18 µL de *primer* Sentido (Fw), 1.16 µL de *primer* Antisentido (Rv), 0.4 µL de enzima SuperScript III y 0.78 µL de H<sub>2</sub>O DEPC. Las muestras se centrifugaron (Allegra X-22R) y se trataron a 50 °C por 15 min, 95 °C por 2 min, 45 ciclos a 95 °C por 15 s, 50 °C por 30 s y 60 °C por 30 s.

**Tabla. 6.3.** Oligonucleótidos (DENV, NS5) empleados para las reacciones de RT-PCR *One step*.

Oligonucleótido NS5 (DENV)	Secuencia	Amplicón
Fw	5' -TAC AAC ATG ATG GGA AAG CGA GAG AAA AA -3'	255pb
Rv	5' -GTG TCC CAG CCG GCG GTG TCA TCA GC- 3'	

#### 6.2.1.3 Depleción de RNAr

De cada extracto de RNA total (30 µL) se tomó el volumen requerido para obtener 2000 ng de RNA (3.81 µL, 0.87 µL, 4.06 µL, 4.05 µL y 3.69 µL para las muestras respectivas MYF1, MYF2, MYF3, MYF4 y MYF5<sup>3</sup>), considerando la concentración medida por espectrofotometría, y se llevó a un volumen total de 12 µL con H<sub>2</sub>O libre de nucleasas. Se empleó el kit NEBNext® rRNA Depletion Kit (Human/Mouse/Rat) (*New England, Biolabs, MA, USA*) con algunas modificaciones a las instrucciones de uso del proveedor.

<sup>3</sup> El número de muestras se redujo, ya que no todas amplificaron durante la RT-PCR (ver resultados).

A cada una de las alícuotas de 12  $\mu\text{L}$  se le agregó 3  $\mu\text{L}$  de mezcla RNA/Sonda (1 $\mu\text{L}$  de *NEBNext rRNA Depletion Solution* y 2 $\mu\text{L}$  de *Probe Hybridization Buffer*) y se mezcló. La mezcla se trató en el termociclador (*BIO-RAD T100*) a 95 °C por 2 min; posteriormente se hizo descender la temperatura a 22 °C con un “*ramp-rate*” de 0.1 °C/s (durante aproximadamente 12 min) y finalmente permaneció a 22 °C por 5 min. A cada muestra se le agregó 5  $\mu\text{L}$  de mezcla de digestión por RNasa H (2 $\mu\text{L}$  *NEBNext NRase H*, 2 $\mu\text{L}$  *RNase H Reaction Buffer* y 1  $\mu\text{L}$  de agua libre de nucleasas) y se mezcló. Se incubó en termociclador (*Gene Amp PCR System 9700*) a 37 °C por 30 min y se agitó. A cada muestra se le agregó 30  $\mu\text{L}$  de mezcla para digestión por DNasa I (5  $\mu\text{L}$  de *DNase I reaction buffer*, 2.5  $\mu\text{L}$  *DNase I RNase-free*, 22.5  $\mu\text{L}$  de agua libre de nucleasas) y se mezcló por pipeteo. A cada muestra se le agregó 50  $\mu\text{L}$  de H<sub>2</sub>O DEPC mezclando y filtrando en columnas *zymo* para centrifugar a 14,000 rpm durante 30 s. Se lavó con 700  $\mu\text{L}$  de *buffer* de lavado centrifugando a 14,000 rpm durante 30 s. Se agregó 500  $\mu\text{L}$  del mismo *buffer* para repetir el lavado centrifugando a 14,000 rpm durante 30 s. Se eluyó con 30  $\mu\text{L}$  de *buffer* de elución caliente, centrifugando a 14,000 rpm durante 3s. Se cuantificó la concentración de 1  $\mu\text{L}$  del RNA remanente por medio de espectrofotometría utilizando el equipo NanoDrop 2000 (Thermo Scientific). Las muestras limpias de RNAr se almacenaron a -70° C.

#### *6.2.1.4 Procedimientos llevados a cabo en la Unidad de Secuenciación Masiva y Bioinformática (USMB) del Instituto de Biotecnología (IBT), UNAM*

##### 6.2.1.4.1 Verificación de la integridad del RNA / RIN

De cada muestra con el RNAr depletado se enviaron 15  $\mu\text{L}$  a la USMB para su secuenciación masiva. Como parte del protocolo previo a la secuenciación, la USMB hace la separación de las moléculas de RNA por tamaño mediante electroforesis microcapilar (*Agilent 2100 Bioanalyzer*). El RNA es teñido con un agente intercalante, la separación es inducida por voltaje y se registra la detección de la fluorescencia inducida por laser. Los datos se archivan y se representan como electroferogramas, imágenes “*gel like*” o de forma tabular (*Agilent, 2005*). Con estos datos se determina el RIN (*RNA Integrity Number*), que refleja el grado de degradación de las muestras de RNA en función de la presencia de RNAr 28S y 18S.

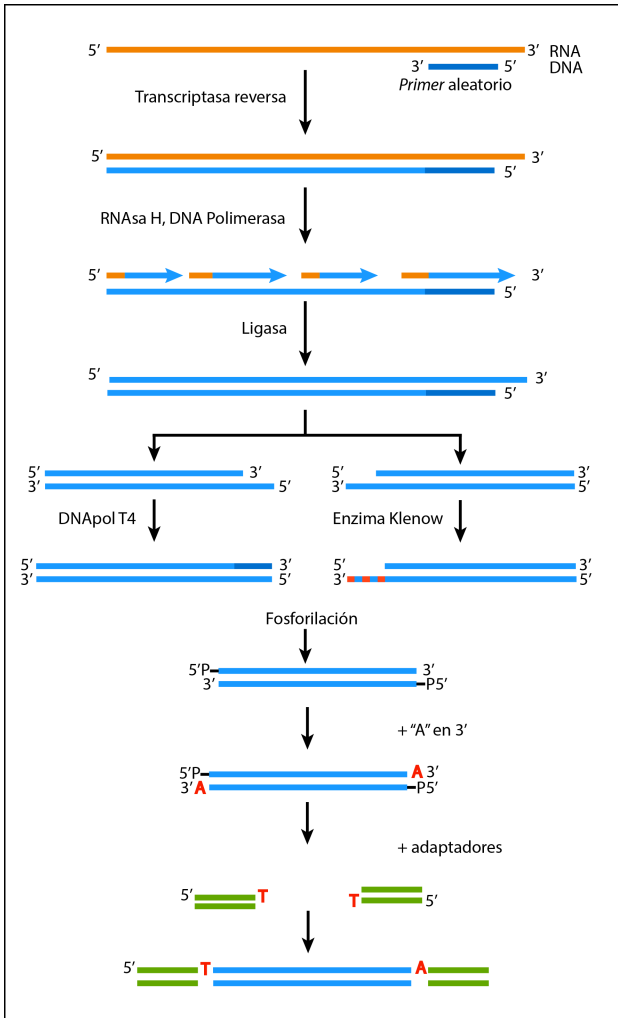
##### 6.2.1.4.2 Construcción de las bibliotecas genómicas

Se construyeron las bibliotecas genómicas siguiendo el protocolo del fabricante (figura 6.3).

Las bibliotecas genómicas resultantes tuvieron insertos de 85 a 315 pb, y adaptadores que, en conjunto, tuvieron un tamaño aproximado de 135 pb, lo que generó tamaños totales de 220 a 450 pb (figura 6.5).

### 6.2.1.4.3 Secuenciación masiva

Con las bibliotecas se llevó a cabo la secuenciación masiva RNaseq (*Shotgun*) empleando un sistema MiSeq® (*Illumina*) *paired end*. Esta es una tecnología de secuenciación por síntesis que integra los pasos de generación de *clusters*, amplificación, secuenciación y análisis de datos en un solo instrumento. La elección de secuencias *paired-end* tuvo la intención de facilitar el posterior ensamblado de los genomas. Las *reads* de secuenciación tuvieron una longitud de  $2 \times 75$  pb (figura 6.5).



**Figura 6.3.** Construcción de bibliotecas genómicas. El RNA fue fragmentado por la acción de una solución de catión divalente ( $\text{Cu}^{2+}$  o  $\text{Zn}^{2+}$ ) calentada a  $94^\circ\text{C}$  durante 8 min. Posteriormente se realizó la síntesis del cDNA; para la primera cadena (A) se emplearon *primers* aleatorios y una retrotranscriptasa, y para la segunda cadena (B) se empleó una DNA polimerasa y RNAsaH. Los extremos de los fragmentos de cDNA fueron reparados con enzima T4 DNAPol y enzima Klenow (C) y posteriormente fosforilados empleando una polinucleótido cinasa y ATP. Se adicionó una adenina a los extremos 3' (D) y se ligaron los adaptadores de doble cadena (E). Se realizó la selección de los fragmentos según el tamaño deseado, verificando por electroforesis capilar. Información obtenida por comunicación con el doctor Ricardo Grande, USMB, IBT, UNAM.

### 6.2.2 Procesamiento de muestras de sueros de pacientes infectados con DENV

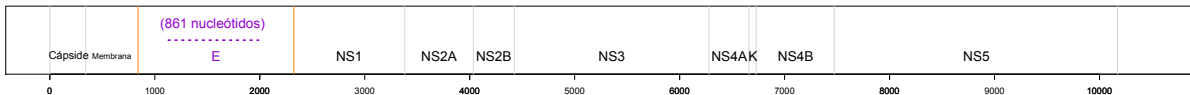
Las muestras fueron procesadas por personal del laboratorio de Virología de la UACM, quienes realizaron la extracción de RNA con el kit QIAamp Viral RNA Mini Kit (Qiagen); la serotipificación se llevó a cabo mediante reacciones singleplex de RT-PCR One Step (*One Step RT-PCR SuperScript III* de *Invitrogen*<sup>TM</sup>) usando oligonucleótidos

que amplifican la región correspondiente a la proteína NS5 de DENV y la sonda TaqMan® para DENV-2 (tabla 6.3) en el equipo 7300 *Real Time* de *Applied Biosystems*. Posteriormente se realizó la amplificación de un fragmento correspondiente a la región codificante de la proteína E empleando la DNA polimerasa de alta fidelidad Phusion™ (Thermo Scientific™) y los iniciadores que se muestran en la tabla 6.4. Los productos de la PCR se separaron en un gel de agarosa a 1% y el producto del tamaño esperado se cortó y se purificó utilizando un kit de extracción de DNA a partir de gel (Qiagen), el DNA eluido se cuantificó en nanodrop y finalmente se secuenció utilizando tanto el iniciador Fw como el Rv mediante el método de Sanger (Macrogen, Korea).

**Tabla 6.4.** Oligonucleótidos (DENV, E) empleados para las reacciones de PCR.

Oligonucleótido E (DENV)	Secuencia	Amplicón
<i>Fw</i>	5'-ATG AAG AGC AGG ACA AAA GGT TC-3'	937pb
<i>Rv</i>	5' TCA AAC ATT TGG CCG ATR GAA-3'	

En el laboratorio de bioinformática de la UACM se obtuvo el consenso de las 16 secuencias de 861 nt de longitud (figura 6.4) correspondientes a la región de la proteína E, empleado el software *Seq Trace* (Stucky, 2012).



**Figura 6.4.** Región amplificada del gen E para las muestras de sueros humanos.

## 6.3 Procesos *in silico*

### 6.3.1 “Prefiltrado” de secuencias

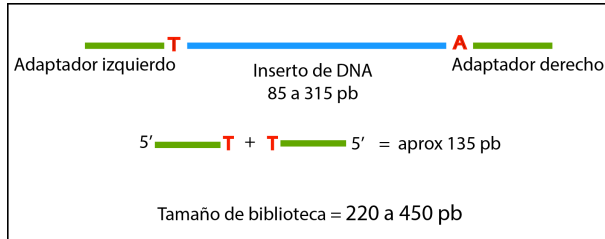
#### 6.3.1.1 Análisis de calidad de las secuencias

La USMB reportó las lecturas derivadas de la secuenciación en formato fastq (figura 6.7) que muestra tanto la secuencia de nucleótidos como su puntuación de calidad correspondiente en formato PHRED. El score PHRED Q, se define como:

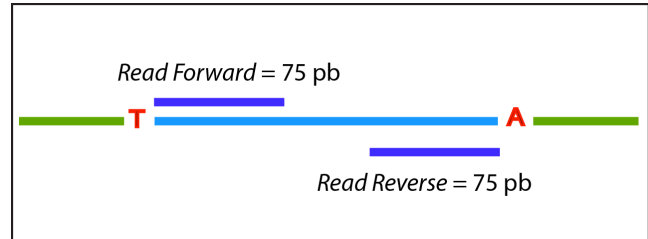
$$Q = -10\log_{10}P$$

donde P es la probabilidad de que la base en la secuencia sea incorrecta.

Para verificar la calidad de las secuencias obtenidas se empleó el software FastQC (Andrews, 2010) que genera un reporte gráfico para las secuencias de cada *read* que refleja, entre otros aspectos, las estadísticas de calidad por posición en el *read*, el % GC esperado y obtenido, las secuencias sobrerrepresentadas y el contenido de adaptadores.



**Figura 6.5.** Características de las bibliotecas genómicas generadas.



**Figura 6.6.** Secuencias *paired-end*.

```
@NS500502_0019_FC_HYJWKBGXX:1:11102:26615:15478#ATTACTCG+GCCTCTAT
GGCAACAGTCACAGGAGGAATCTTCTTATTCTTAATGAGCGGAAAAGGTGTAGGGAAGATGACCCTGGGAATGCG
+
EAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA/EEEEEEEEAAAAA
@NS500502_0019_FC_HYJWKBGXX:1:11102:26615:15478#ATTACTCG+GCCTCTAT
GAACTGCCGGAGACCCTGGAAACACTGCTTCTACTGACACTCTGGCAACAGTCACAGGAGGAATCTTCTTATT
+
AAAAAEEEEEEEEEEAEEA/E/EEEEEEEEAAE/EEEEEEEAEEE/EE/EEEE/EEA/E<E/EEAE<EEE/E
```

**Figura 6.7** Estructura del formato fastq.

### 6.3.1.2 Limpieza de adaptadores presentes en las secuencias

Para eliminar los adaptadores que se encontraron en las secuencias, se empleó el software PEAT (Li *et al.*, 2015), que emplea un algoritmo de limpieza diseñado específicamente para secuencias pareadas. Al no requerir *a priori* secuencias de adaptadores, se convierte en una herramienta conveniente para estudios de meta-análisis. Para ejecutar el programa se empleó la siguiente línea de comandos:

```
PEAT_mac paired -1 PATH/MYF1_R1.fastq -2 PATH/MYF1_R2.fastq -o
PATH/MYF1_R1_sin_adaptadores.fastq
PATH/MYF1_R1_sin_adaptadores.fastq --adapter_contexts report.txt4
```

<sup>4</sup> En las líneas de comando descritas en este texto, PATH indica la ruta de ubicación del archivo o directorio en cuestión.

La ejecución del programa da como resultado los archivos fastq ya limpios de adaptadores y un reporte donde se muestran las secuencias de los adaptadores removidos, su longitud y frecuencia (figura 6.8).

```
@PEAT report
Mode: paired-end
@Total number of reads: 27307982
@Average length of reads in raw data: 75
@Average length of reads after adapter trimming: 74.8827
-----
@The top 50 frequent adapter contexts:
@minimal length (user defined): 10
@Context Length Number
AGATCGGAAGAG 12 3308
AGATCGGAAGA 11 3277
AGATCGGAAGAGC 13 2635
AGATCGGAAGAGCA 14 2585
AGATCGGAAGAGCAC 15 2297
AGATCGGAAGAGCACACA 16 1925
AGATCGGAAGAGCACACGTCTG 22 1808
AGATCGGAAGAGCACAC 17 1728
AGATCGGAAGAGCACACGTC 20 1614
AGATCGGAAGAGCACACGTCTGAA 24 1511
AGATCGGAAGAGCACACG 18 1494
AGATCGGAAGAGCACACGTCT 21 1421
AGATCGGAAGAGCACACGT 19 1353
CACAAATTGTCCTATTAAGCAGAAGTCAACCACGATGGGTTGA 43 1324
AGATCGGAAGAGCACACGTCTGA 23 1296
AGATCGGAAGAGCACACGTCTGAAC 25 1261
CACAAATTGTCCTATTAAGCAGAAGTCAACCACGATGGGTTGAC 44 1247
CACAAATTGTCCTATTAAGCAGAAGTCAACCACGATGGGTTGACG 45 1104
AGATCGGAAGAGCACACGTCTGAACT 26 1031
```

Figura. 6.8. Reporte de los adaptadores removidos generado por PEAT.

### 6.3.1.3 Eliminación de las secuencias redundantes

Para remover los duplicados de las *reads* de secuenciación se empleó la herramienta de CD-HIT (Li & Godzik, 2006). Para ejecutar el programa se empleó la siguiente línea de comandos:

```
PATH/cd-hit-dup -i PATH/MYF1_R1_sin_adaptadores.fastq -i2 PATH/MYF1_R2_sin_adaptadores.fastq -o MY1_sindups1.fastq -o2 MY1_sindups2.fastq
```

La ejecución del programa da como resultado los archivos fastq ya sin *reads* duplicados.

### 6.3.2 Obtención del genoma consenso mapeando contra genoma de referencia

Para obtener el genoma consenso se siguieron dos estrategias: mapeo contra genoma de referencia, empleando nuevamente SMALT, Bowtie2 2.2.9 (Langmead & Salzberg, 2012) y las

herramientas SAMtools y BCFtools 1.3.1 (Danecek *et al.*, 2011), y ensamblaje de novo, empleando Velvet 1.2.10 (Zerbino & Birney, 2008).

### 6.3.2.1 Binning, reclutamiento de fragmentos

Dada la naturaleza heterogénea de las muestras secuenciadas (*pools* de mosquitos macerados), se siguió una estrategia de metagenómica enfocada a DENV.

*Binning* es un proceso comúnmente empleado en estudios metagenómicos que consiste en la asociación de secuencias con las especies (o niveles taxonómicos superiores) a las que corresponden (Kunin *et al.*, 2008).

El *binning* que se realizó consistió en el reclutamiento de fragmentos descrito por Sharpton (2014). Con el uso del software SMALT (*Wellcome Trust Sanger Institute*, 2012) se seleccionaron todas las secuencias resultantes del proceso de “prefiltrado” que tuvieran identidad de al menos 60% con las secuencias presentes en una base de datos de referencia.

La base de datos de referencia se construyó con las secuencias de genoma completo presentes en la base de datos *Nucleotide* de NCBI<sup>5</sup> que pertenecieran a la familia *Flaviviridae*. Dicha base de datos constó de 7,020 secuencias y, aunque en ella están representados 94 Flavivirus (Anexo 1), cerca de 95% de los registros corresponden a sólo seis de ellos (tabla 6.4).

**Tabla 6.4.** Virus de la base de datos de *Flaviviridae* (obtenidos de *Nucleotide* de NCBI<sup>5</sup>) con mayor número de secuencias.

Virus	Número de secuencias en la base de datos
Dengue	3,884
Hepatitis	1,199
West Nile	995
Japanese encephalitis	174
Yellow Fever	57
Zika	43
Otros	668

Para realizar el *binning* fue necesario:

1. Crear el índice para la base de datos:

```
smalt index -k 13 -s 4 flavivirus_index  
PATH/sequence_flaviviridae.fasta
```

---

<sup>5</sup> Consultada en abril de 2016.

## 2. Mapear los *reads* de cada muestra contra el índice de la base de datos:

```
smalt map -f samsoft -o MYF1.sam -y 0.6 -i 320 -j 80 -l pe  
flavivirus_index PATH/MY1_sindups1.fastq PATH/MY1_sindups2.fastq
```

El resultado de mapear contra la base de datos es un archivo de salida de tipo SAM (Sequence Alignment/Map), del que es posible extraer las secuencias que hicieron “*match*” con la base de datos. Para ello se emplearon diversos comandos de SAMtools 1.3.1 (Li *et al.*, 2009) y Linux:

- Convertir a un formato binario (más compacto), BAM:

```
samtools view -b -S -o MYF1.bam MYF1.sam
```

- Seleccionar y escribir en otro archivo sam sólo aquellas *reads* que hicieron *match* con la base de datos:

```
samtools view -f 2 MYF1.bam | tee MYF1_hits.sam
```

- Seleccionar y escribir en otro archivo sam sólo aquellas *reads* que no hicieron *match* con la base de datos:

```
samtools view -F 2 MYF1.bam | tee MYF1_no_hits.sam
```

- Contar el número de *reads* que hicieron *match* con la base de datos:

```
samtools view -c -f 2 MYF1.bam
```

- Convertir a formato fasta las *reads* que hicieron *match*:

Secuencias *Forward*:

```
cut -f 1,2,10,11 MYF1.sam | awk '$2==83||$2==99' | cut -f 1,3,4  
|awk '{print $1"\n"$2"\n"+" "\n"$3}' | sed -e "s/NS500/@NS500/g"  
> MYF1_match_flavi_F.fastq
```

Secuencias *Reverse*:

```
cut -f 1,2,10,11 MYF1.sam | awk '$2==147||$2==163' | cut -f 1,3,4  
|awk '{print $1"\n"$2"\n"+" "\n"$3}' | sed -e "s/NS500/@NS500/g"  
> MYF1_match_flavi_R.fastq
```

### 6.3.2.2 Remapeo contra genoma de referencia, determinación de variantes y obtención de genoma consenso

Las secuencias que por *binning* hicieron *match* con la base de datos de *Flaviviridae* fueron remapeadas contra un genoma de referencia, también con SMALT. Para ello se eligió el genoma de DENV-2 gb: KJ189309, extraído de mosquito *Aedes aegypti* de México (Yucatán) en 2010 (longitud: 10,300 nucleótidos).

El *pipeline* y las líneas de comando empleados fueron:

1. Construir el índice del genoma de referencia:

```
PATH_bowtie2-2.2-2.9/bowtie2-build PATH_secuencia_referencia.fna  
PATH_indice
```

2. Mapear las secuencias contra la referencia:

```
PATH_bowtie2-2.2-2.9/bowtie2 -x PATH_indice -1  
path_secuencias_forward.fastq -2 PATH_secuencias_reverse.fastq -S  
archivo.sam
```

3. Generar archivo bam:

```
samtools view -b -S -o PATH_archivo.bam PATH_archivo.sam
```

4. Ordenar archivo bam:

```
samtools sort PATH_archivo_bam.bam -o PATH_archivo_sorted.bam
```

5. Generar archivo bcf:

```
samtools mpileup -g -f PATH_secuencia_referencia.fna  
PATH_archivo_sorted.bam > PATH_variantes.bcf
```

6. Realizar la determinación (llamado) de variantes:

```
bcftools call -c -v PATH_variantes.bcf > PATH_variantes.vcf
```

7. Indexar para visualizar:

```
samtools index PATH_sorted.bam
```

## 8. Visualizar:

```
samtools tview PATH_sorted.bam PATH_secuencia_referencia.fna
```

## 9. Obtención de genoma consenso en formato fastq:

```
samtools mpileup -uf PATH_secuencia_referencia.fna  
PATH_sorted.bam | bcftools call -c | vcfutils.pl vcf2fq >  
PATH_genoma_consenso.fq
```

Como resultado del punto 6 se obtiene un archivo en formato vcf (“variant call format”), que proporciona información sobre las posiciones en el genoma para las que los *reads* constituyen una variante, la base presente en el consenso, en las *reads* y su calidad PHRED, entre otros datos (figura 6.9).

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	MYF2_sorted.bam
gb:KJ189309	184.016		Dengue	T		C			
DP=9;VDB=0.163983;SGB=-0.662043;MQSB=0.974597;MQ0F=0;AF1=1;AC1=2;DP4=0,0,5,4;MQ=42;FQ=-53.9861								GT:PL	1/1:217,27,0
gb:KJ189309	222.001		Dengue	C		T			
DP=12;VDB=0.952754;SGB=-0.680642;MQSB=0.95494;MQ0F=0;AF1=1;AC1=2;DP4=0,0,5,7;MQ=42;FQ=-62.9862								GT:PL	1/1:255,36,0

Figura 6.9. Formato vcf.

### 6.3.2.3 Análisis de calidad del mapeo contra genoma de referencia

Mediante el software QualiMap2.2 (Okonechnikov *et al.*, 2015), se evaluó la calidad del mapeo contra el genoma de referencia.

### 6.3.3 Ensamble de genoma *de novo*

Como alternativa al mapeo contra el genoma de referencia, se realizó el ensamble *de novo* empleando Velvet 1.2.10. Las líneas de comando empleadas fueron:

```
velveth ensamble17/ 17 -fastq -separate -shortPaired  
MYF_match_flavi_F.fastq MYF_match_flavi_R.fastq
```

```
velvetg ensamble17/ -cov_cutoff_ auto -exp_cov auto
```

Dado que diferentes tamaños de k-meros generan ensambles de distinta calidad, se buscó aquel que permitiera ensamblar los contigs (nodos) más largos. Una vez encontrado el k-mero óptimo, el ensamble generado se empleó como referencia para realizar nuevamente el *binning* (indexado y mapeado) con SMALT, seguido del ensamble *de novo* con Velvet. La identidad de los productos de los ensambles se corroboró mediante BLASTx.

## 6.4 Genotipificación

La genotipificación de secuencias del DENV (y de otros flavivirus) se lleva a cabo mediante la construcción de árboles filogenéticos con genomas de referencia por su agrupación con el más cercano (Vecna Technologies, 2016). Para genotipificar las secuencias del presente estudio se emplearon genomas contenidos en la base de datos Virus Pathogen Database and Analysis Resource (ViPR) (<https://www.viprbrc.org>). Los 31 genomas descargados (Anexo 3) representan a los seis genotipos reconocidos para DENV-2 (Asiático I, Asiático II, Americano, Asiático-Americano, Cosmopolita y Selvático). Con éstos y las secuencias problema se construyó un árbol Neighbour-joining con 1,000 réplicas de *bootstrap* empleando el software SeaView 4.6 (Gouy *et al.*, 2010).

## 6.5 Análisis de variabilidad de genomas

### 6.5.1 Análisis de variabilidad de genoma completo con respecto al genoma del DENV proveniente de suero humano y de mosquito

A partir del análisis de los archivos vcf (generados con BCFtools) se determinaron las posiciones en común de los genomas de DENV-2 aislados de *Aedes aegypti* que constituyen variantes, comparándolos contra un genoma de referencia proveniente de suero humano (gb:GU131955, Quintana Roo, México, 2004) y otro proveniente de mosquito *Aedes aegypti* (gb:KJ189309, Yucatán, México, 2010). Se identificó además la naturaleza de las mutaciones (sinónimas, no sinónimas: conservativas o no conservativas). Para llevar a cabo dicha tarea se creó un script *ex profeso* en Python, en el que el listado de variantes para cada genoma se consideró como un conjunto, encontrándose las intersecciones entre ellos tanto a nivel de nucleótido como de aminoácido.

### 6.5.2 Análisis del uso relativo de codones sinónimos

Se ha observado que el patrón de uso de codones en distintos arbovirus es consistente con el uso de codones de sus respectivos hospederos (Velázquez-Salinas *et al.*, 2016). Para averiguar si el uso relativo de codones de las secuencias provenientes de mosquito difiere del que tienen las secuencias provenientes de sueros humanos, se realizó un análisis de agrupamiento jerárquico (empleando el paquete R). Para ello se descargaron 176 secuencias de ViPR correspondientes al genotipo Asiático-Americano, se alinearon con las obtenidas en el presente estudio y se

recortaron a la longitud de 861 nucleótidos de la región de E con SeaView. Se calculó el uso de codones sinónimos relativos (*RSCU*, *Relative Synonymous Codon Usage*) empleando el servidor CAIcal (<http://genomes.urv.cat/CAIcal/E-CAI/>), generando una matriz como la que se muestra en la figura 6.10. Con los valores de la matriz de RSCU se realizó el análisis de agrupamiento.

<b>CODONS</b>	<b>TTT</b>	<b>TTC</b>	<b>TTA</b>	<b>TTG</b>	<b>CTT</b>	<b>CTC</b>	<b>CTA</b>	<b>CTG</b>	<b>ATT</b>	<b>ATC</b>	<b>ATA</b>
<b>Aminoacids</b>	<b>F</b>	<b>F</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>I</b>	<b>I</b>	<b>I</b>
GU131955_HS_2004_Mexico	0.67	1.33	1.30	1.04	0.00	1.04	0.78	1.83	0.45	1.35	1.20
MYF1_AE_2012_Mexico	1.33	0.67	1.30	1.04	0.00	1.04	0.78	1.83	0.45	1.35	1.20
MYF2_AE_2012_Mexico	0.89	1.11	1.30	1.04	0.00	1.04	0.78	1.83	0.30	1.50	1.20

**Figura 6.10.** Fragmento de matriz de RSCU para tres de las secuencias analizadas en este estudio.

## 6.6 Filogeografía

Se realizaron análisis filogeográficos con la intención de identificar las relaciones evolutivas de las secuencias analizadas en su contexto geográfico. Para ello se emplearon los programas BEAST 2.4.3, BEAUti y TreeAnnotator (Drummond *et al.*, 2012) y Tracer 1.6 (Rambaut & Drummond, 2007).

Se analizaron las 20 secuencias del fragmento del gen E obtenidas en el presente estudio y 154 secuencias adicionales descargadas de la base de datos ViPR del continente americano (genotipo Asiático-Americano) y de países asiáticos (genotipo Selvático). La figura 6.11 muestra los países representados en las secuencias empleadas.

Los nombres de las secuencias con las que se construyó el árbol contenían los datos de año de colecta y ubicación geográfica (latitud y longitud en coordenadas UTR) para poder integrar la información temporal y espacial a la reconstrucción filogeográfica.

Se empleó el modelo de sustitución HKY (el modelo óptimo según el análisis realizado previamente con Datamonkey (<http://www.datamonkey.org>), un reloj molecular relajado para la localidad y un reloj molecular estricto para las secuencias. Se empleó una longitud de cadena de 20,000,000 y se obtuvieron 10,000 muestras. El resultado de cada análisis es un árbol filogenético MCC (*Maximum Clade Credibility Tree*). Los análisis se hicieron a dos escalas.

### 6.6.1 Escala global

Se consideraron todas las secuencias descargadas de ViPR para el genotipo en cuestión (154), e incluyendo las generadas en este estudio (20).

### 6.6.2 Escala local

Se consideraron únicamente las secuencias provenientes del estado de Guerrero (20, tanto de mosquito como de suero humano).

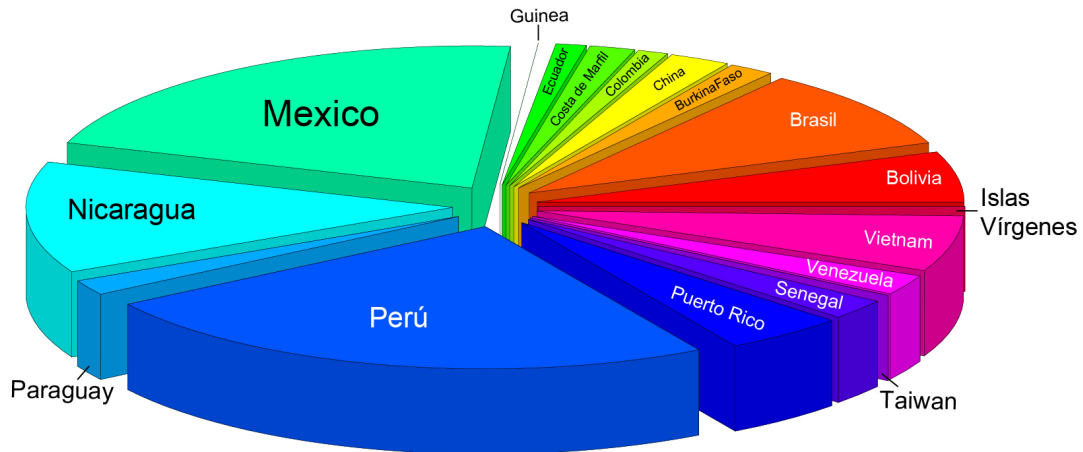


Figura 6.11. Países representados en las secuencias con las que se realizó el análisis filogeográfico.

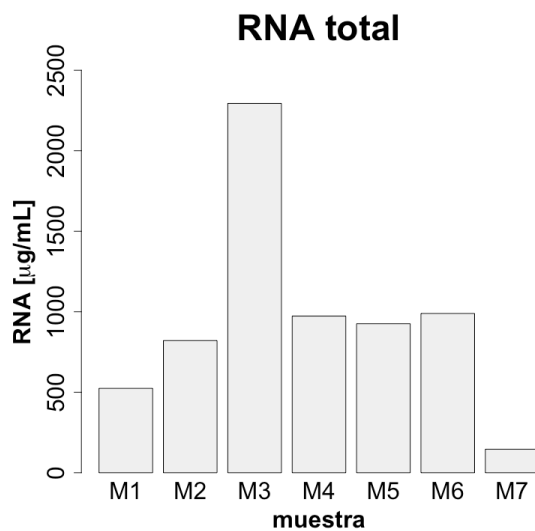
### 6.7 Estructura poblacional

Para determinar si la población estudiada (las secuencias de DENV analizadas) se estructura de acuerdo con la geografía y/o a la temporalidad, se aplicó la prueba de Slatkin & Maddison (Slatkin & Maddison, 1989) empleando el software HyPhy. La prueba determina el número mínimo de eventos de migración entre las poblaciones separadas consistente con la estructura del árbol filogenético reconstruido. El análisis se hizo a partir de un árbol *neighbor joining* construido (también con HyPhy) con las secuencias del fragmento del gen E empleadas para el análisis filogeográfico.

## 7. Resultados

### 7.1 Extracción de RNA

La concentración del RNA total extraído de cada muestra cuantificada por espectrofotometría se presenta en la figura 7.1. Las concentraciones obtenidas tienen un rango de 146.3  $\mu\text{g/mL}$  (M7) a 2,294.2  $\mu\text{g/mL}$  (M3).



**Figura 7.1.** Cuantificación de la concentración de RNA total extraído de cada muestra mediante espectrofotometría.

### 7.2 RT-PCR *One step singleplex*, serotipificación

La tabla 7.1 presenta los resultados de las RT-PCR *One step singleplex* realizados a cada muestra. Se observa que las muestras M2 y M7 no amplificaron con las sondas TaqMan® de DENV-2, por lo que fueron descartadas del estudio.

Con respecto al resto de las muestras, se corrobora que corresponden al serotipo DENV-2.

**Tabla 7.1.** Resultados de la amplificación por RT-PCR *One step singleplex*. Como controles se emplearon, para DENV-2: extracto de RNA total de células infectadas con dicho serotipo, y para DENV-1: muestra de suero de paciente confirmado para dicho serotipo.

Muestra	DENV-1	DENV-2
M1	-	+
M2	-	-
M3	-	+
M4	-	+
M5	-	+
M6	-	+
M7	-	-
Control	+	+

La tabla 7.2 presenta los valores de Ct (“*Threshold cycle*”) de la RT-PCR *One step singleplex* con sondas TaqMan® de DENV-2 para las muestras que sí amplificaron. Se observa que MYF3 y MYF4 son las muestras que presentan menor concentración de RNA.

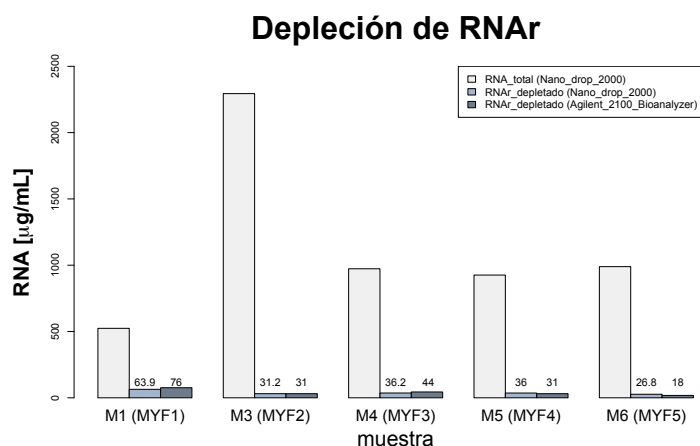
De aquí en adelante, las muestras que sí amplificaron se designan como MYF1 a MYF5.

**Tabla 7.2.** Valores de Ct (“*Threshold cycle*”) para la RT-PCR *singleplex* (la relación entre Ct y concentración de RNA es inversa).

Muestra	Ct (RT-PCR <i>singleplex</i> )
M1(MYF1)	18.58
M3(MYF2)	29.43
M4(MYF3)	37.89
M5(MYF4)	32.89
M6(MYF5)	18.39
Control	36.64

### 7.3 Depleción de RNAr

La figura 7.2 presenta la concentración de RNA total de cada muestra en una cuantificación previa y posterior al tratamiento con el *NEBNext® rRNA depletion kit*. Como se esperaba, la reducción en la concentración de RNA fue muy notoria, oscilando entre 87.8% (muestra MYF1) y 98.6% (MYF3).

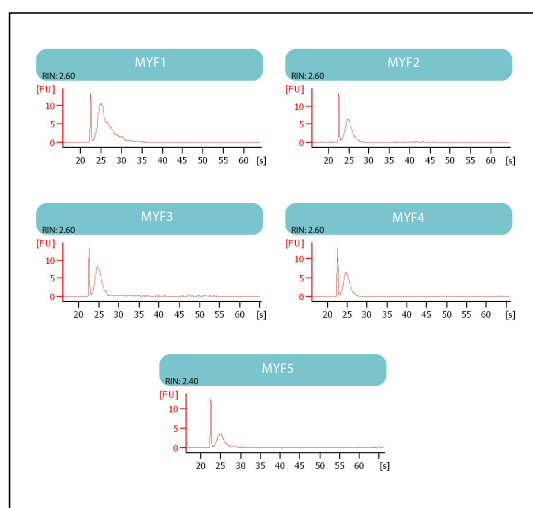


**Figura 7.2.** Comparación de la concentración de RNA total cuantificada por espectrofotometría de manera previa y posterior al tratamiento con el kit de depleción de RNAr *NEBNext® rRNA depletion kit*.

## 7.4 Verificación de la integridad del RNA / RIN

La figura 7.3 muestra los electroferogramas generados a partir de los datos del *Agilent 2100 Bioanalyzer* para cada una de las muestras analizadas en la USMB.

A pesar de que los valores de RIN resultantes (de 2.40 a 2.60) podrían considerarse en otras condiciones como reflejo de degradación, en el presente estudio se explican por el tratamiento de depleción de RNAr previo y por tanto no reflejan el grado de degradación que pudiera haber tenido la muestra (el modelo con el que se determina el RIN considera los valores 28S y 18S como parámetros indicadores de integridad; Schroeder *et al.*, 2006).



**Figura 7.3.** Electroferogramas generados a partir de los datos de *Agilent 2100 Bioanalyzer* para cada una de las muestras analizadas. [FU]: unidades de fluorescencia, [s]: segundos (tiempo de migración).

## 7.5 RNA-seq

La figura 7.4 (A) presenta el número de *reads* que reportó la USMB para cada muestra, el valor máximo lo tuvo MYF1 con 13,653,991 x 2 y el mínimo, MYF5 con 7,790,660 x 2.

## 7.6 Análisis de calidad de la secuenciación

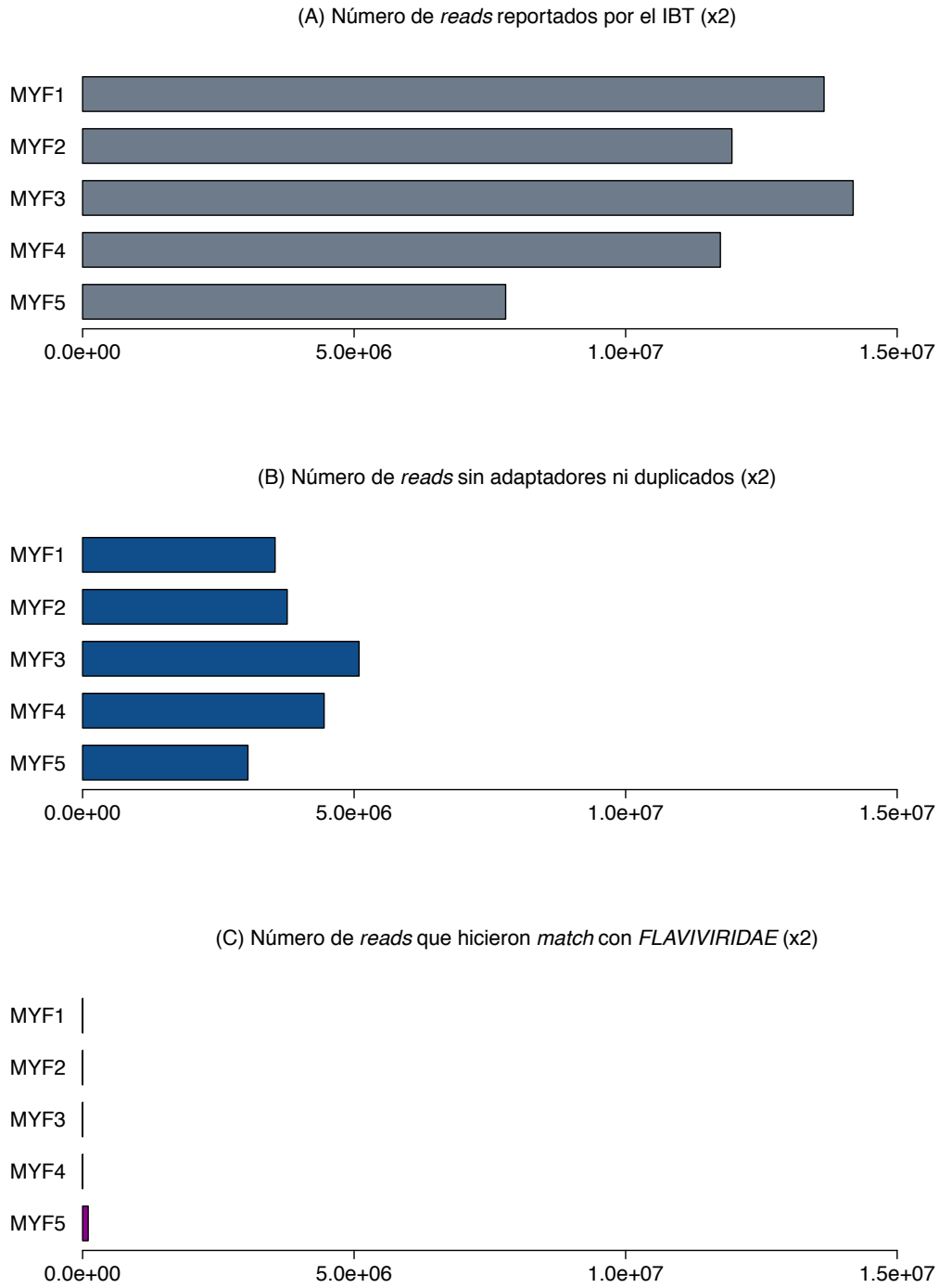
El análisis de calidad realizado con FASTQC indica que para todas las muestras, la calidad de las lecturas es muy buena (promedio de score PHRED entre 30 y 34). En todas se observa que, como se espera en este tipo de secuenciación, la calidad decae conforme aumenta el llamado de bases, siendo el nucleótido 75, en todas las muestras, el que presenta una calidad inferior. Con respecto al %GC, para todas las muestras se presentan múltiples picos y no una distribución normal. Este hecho, más allá de interpretarse como contaminación, refleja la naturaleza diversa del origen de las moléculas de RNA secuenciadas.

## 7.7 Limpieza de adaptadores y eliminación de secuencias redundantes

La figura 7.4 (B) muestra el número de *reads* que quedaron una vez que se hizo la limpieza de adaptadores y secuencias redundantes. Se observa que las secuencias limpias constituyen del 25.9% (MYF1) al 39.1% (MYF5) de las secuencias iniciales.

## 7.8 Binning, reclutamiento de fragmentos

La figura 7.4 (C) muestra el número de *reads* reclutados contra la base de datos de *Flaviviridae* para cada muestra. Los porcentajes de secuencias que hicieron *match* con la base de datos de *Flaviviridae* se encontraron en un rango de 0.0072% (MYF3) a 1.3192% (MYF5) de las secuencias iniciales.



**Figura 7.4.** (A) Número de *reads* obtenidos de la RNA-seq por muestra. (B) Número de *reads* que quedan después de la limpieza de adaptadores y duplicados. (C) Número máximo de *reads* que hicieron *match* con la base de datos de *Flaviviridae*.

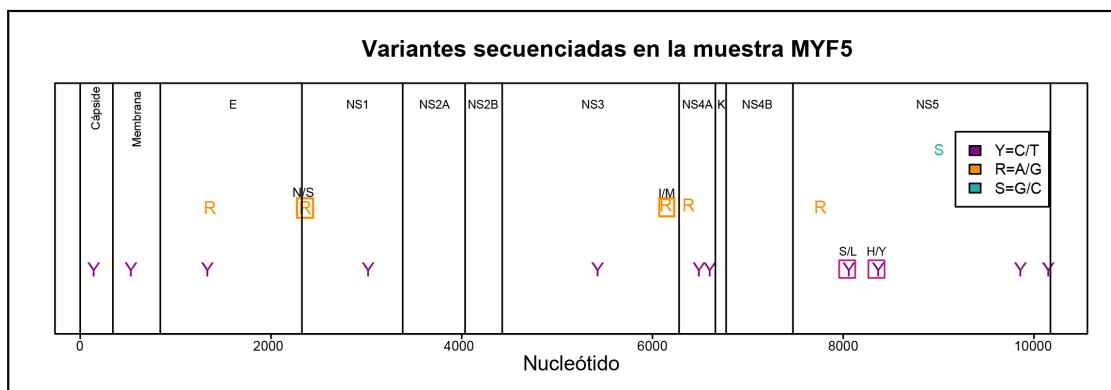
## 7.9 Obtención del genoma consenso

### 7.9.1 Mapeo contra genoma de referencia, variantes poblacionales

Al mapear contra el genoma de referencia se logró obtener genoma consenso para las muestras MYF1, MYF2 y MYF5. Las muestras MYF3 y MYF4 presentaron un número muy reducido de *reads* que mapearon contra el genoma de referencia (menos de 100), por lo que fue imposible obtener una secuencia de éstas. La tabla 7.4 muestra la información de los reportes Qualimap para el mapeo contra referencia de cada una de las muestras. Se observa que la mayor profundidad la tuvo MYF5 (734X). La tabla 7.3 muestra el número de posiciones variables de cada genoma consenso con respecto a la referencia, así como las posiciones que reflejan variabilidad poblacional en cada muestra. La figura 7.5 muestra la distribución de las variantes poblacionales para MYF5 a lo largo del genoma. Se observa que, de las 17 variantes, cuatro corresponden a mutaciones no sinónimas.

**Tabla 7.3.** Posiciones variables con respecto a la referencia, presentes en los archivos vcf correspondientes. Posiciones que reflejan variabilidad poblacional de las secuencias presentes en las muestras de los genomas consenso (Y = C/T, R = A/G, S = G/C).

	Posiciones variables con respecto a la referencia	Posiciones que reflejan variabilidad poblacional de las secuencias por muestra
MYF1	76	ninguna
MYF2	84	1 R, 1 Y
MYF5	93	5 R, 1 S y 11 Y

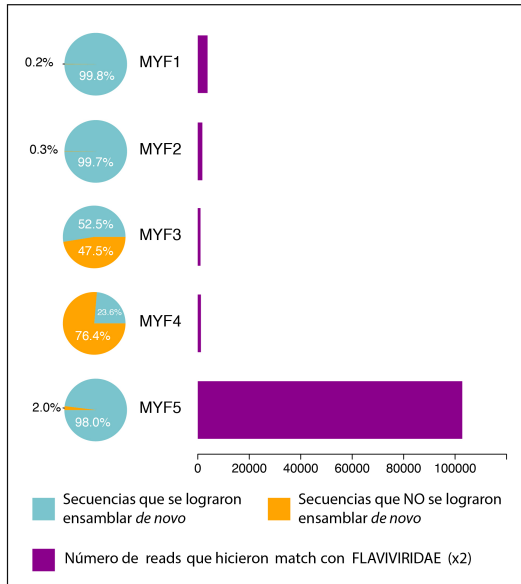


**Figura 7.5.** Distribución de variantes poblacionales de la muestra MYF5 a lo largo del genoma. Las mutaciones no sinónimas se encuentran dentro de recuadros (los aminoácidos para los que codifican se muestran en la parte superior de cada recuadro).

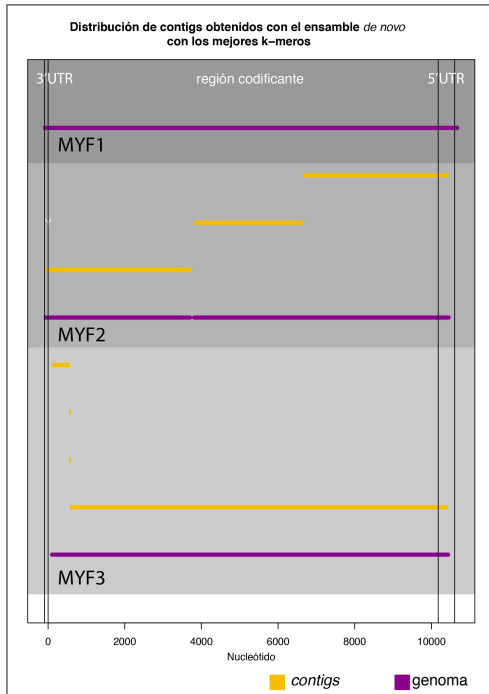
### 7.9.2 Ensamble *de novo*

Se logró ensamblar los genomas completos de MYF1 y el genoma casi completo de MYF2 (hay un *gap* entre los aminoácidos 1245 y 1271) y MYF5 (figura 7.7).

La tabla 7.4 muestra los detalles del ensamble *de novo* para las muestras analizadas. En el Anexo 2 se muestra la información referente a los k-meros empleados y los contigs logrados con éstos. La figura 7.6 muestra la proporción de los *reads* que resultaron del *binning* y lograron ensamblarse *de novo*.



**Figura 7.6.** La gráfica de barras indica el número de *reads* por muestra resultantes del *binning*. Las gráficas de pastel muestran la proporción de las secuencias que se logró ensamblar (verde) y aquellas que no (amarillo).



**Figura 7.7.** Esquema que representa la distribución de los contigs obtenidos mediante ensamble *de novo*. Se observa que para las tres muestras, se obtuvieron también segmentos de las regiones UTR.

**Tabla 7.4. Detalles del mapeo contra genoma de referencia y ensamble *de novo* de las secuencias.**

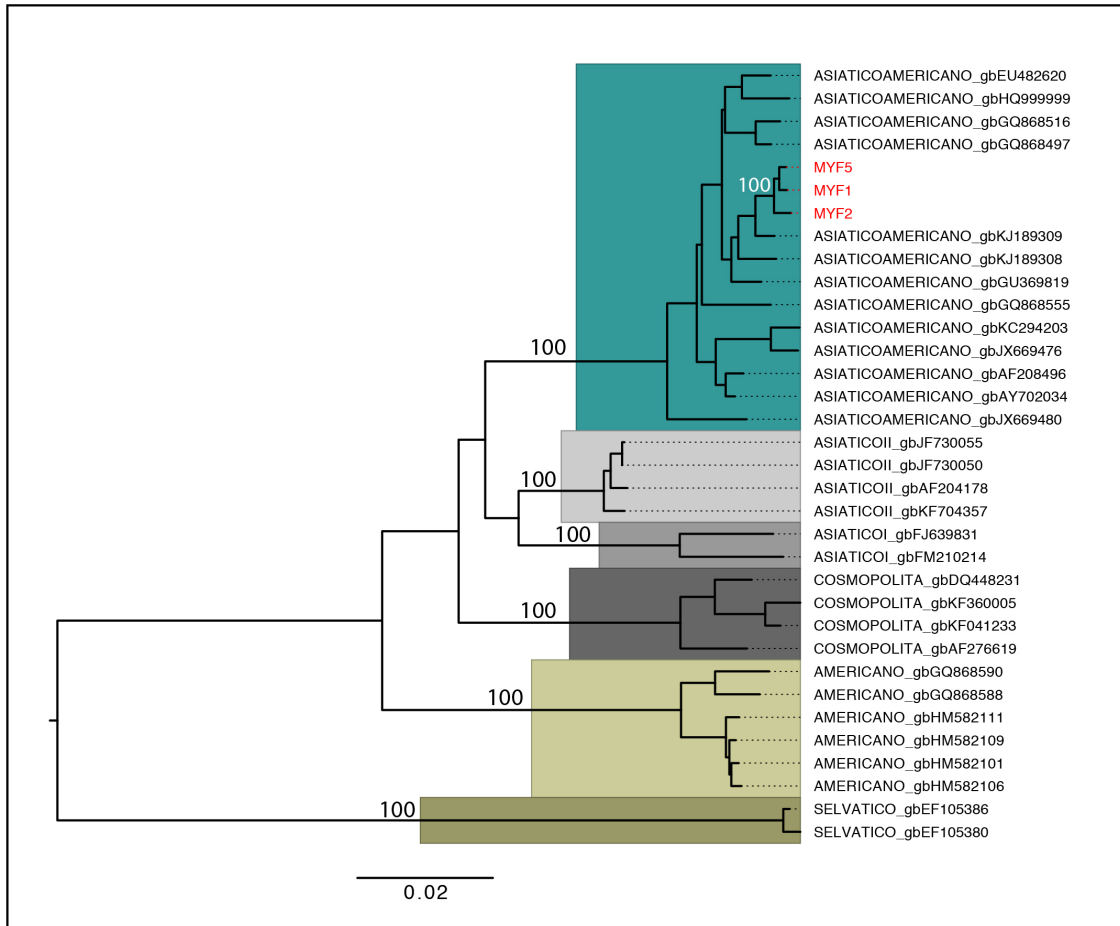
Muestra	Reads derivados de la secuenciación	Reads sin adaptadores ni duplicados	Núm. máximo de <i>match</i> con <i>Flaviviridae</i>	Mapeo contra genoma de referencia (Reporte Quialimap)		Ensamble <i>de novo</i>	
				Reads mapeados	Cobertura (media, desviación estándar)	Referencia para el <i>binning</i>	Resultado del ensamble <i>de novo</i> (Velvet)
MYF1	13,653,991 x 2 = 27,307,982	3,544,613 x 2 = 7,089,226	3,786	3,691	26.6662 12.1585	<i>Binning</i> 1o, BD <i>Flaviviridae</i> 2o, nodos k-amero 17	Se obtuvo un solo contig de 10,648 nucleótidos que comprende a toda la región codificante. N50 = 10,648; Total = 10,743; <i>Reads</i> usados 3780/3786
MYF2	11,957,390 x 2 = 23,914,780	3,767,916 x 2 = 7,535,832	1736	1,711	12.3446 4.9804	Genoma consenso obtenido por mapeo contra referencia	Se obtuvieron dos contigs. Con BLASTx se corroboró que, de la región codificante, fue posible ensamblar los fragmentos correspondientes a los aminoácidos 1 a 1,245 y 1271 a 3391. N50 = 3,745; max = 3,781; total = 10,355; <i>Reads</i> usados=1731/1736 Gap de 72 nt
MYF3	14,190,444 x 2 = 28,380,888	5,091,020 x 2 = 10,182,040	1,026	–	–	<i>Binning</i> <i>Flaviviridae</i>	No se logró ensamblar ningún fragmento considerable del genoma. <sup>6</sup>
MYF4	11,745,024 x 2 = 23,490,048	4,447,670 x 2 = 8,895,340	1,176	–	–	<i>Binning</i> <i>Flaviviridae</i>	No se logró ensamblar ningún fragmento considerable del genoma. <sup>7</sup>
MYF5	7,790,660 x 2 =15,581,320	3,044,691 x 2 = 6,089,382	102,778	101,858	734.1379 221.5021	Genoma consenso obtenido por mapeo contra referencia	Se obtuvieron dos contigs que comprenden toda la región codificante. N50 = 9,840, max = 9,840, total = 10,332; <i>Reads</i> usados = 100756/102778

<sup>6</sup> De los múltiples nodos pequeños ensamblados resultantes, se corroboró con BLASTx que algunos correspondían a DENV-2.

<sup>7</sup> *Ídem.*

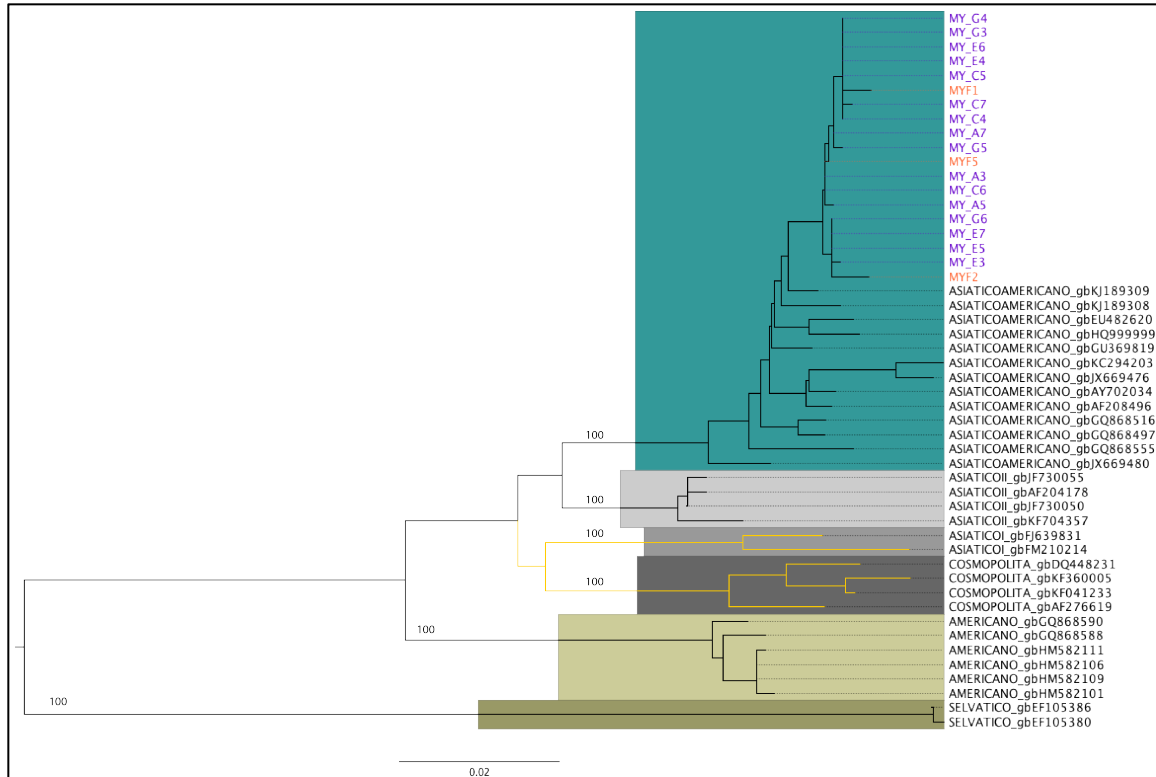
## 7.10 Genotipificación

El genotipo al que pertenecen las secuencias de DENV-2 (MY1, MY2 y MY5) es Asiático-Americano, como lo refleja el árbol filogenético construido con las secuencias de genoma completo (NJ con 1,000 réplicas *bootstrap*) mostrado en la figura 7.8.



**Figura 7.8.** Árbol filogenético (NJ) construido con secuencias de genoma completo, que refleja el genotipo Asiático-Americano de los genomas de DENV-2 provenientes de *Aedes aegypti*.

Al construir el árbol exclusivamente con la región de 861 nucleótidos del gen E e incorporar a las secuencias provenientes de sueros humanos del estado de Guerrero, se observa que su genotipo también es Asiático-Americano (figura 7.9).



**Figura 7.9.** Árbol filogenético (NJ) construido con secuencias de 861 nucleótidos del gen E, que refleja el genotipo Asiático-Americano de las secuencias de DENV-2 extraídas a partir de sueros humanos.

### 7.11 Análisis de variabilidad del genoma completo, con respecto al genoma del DENV proveniente de suero humano y de mosquito

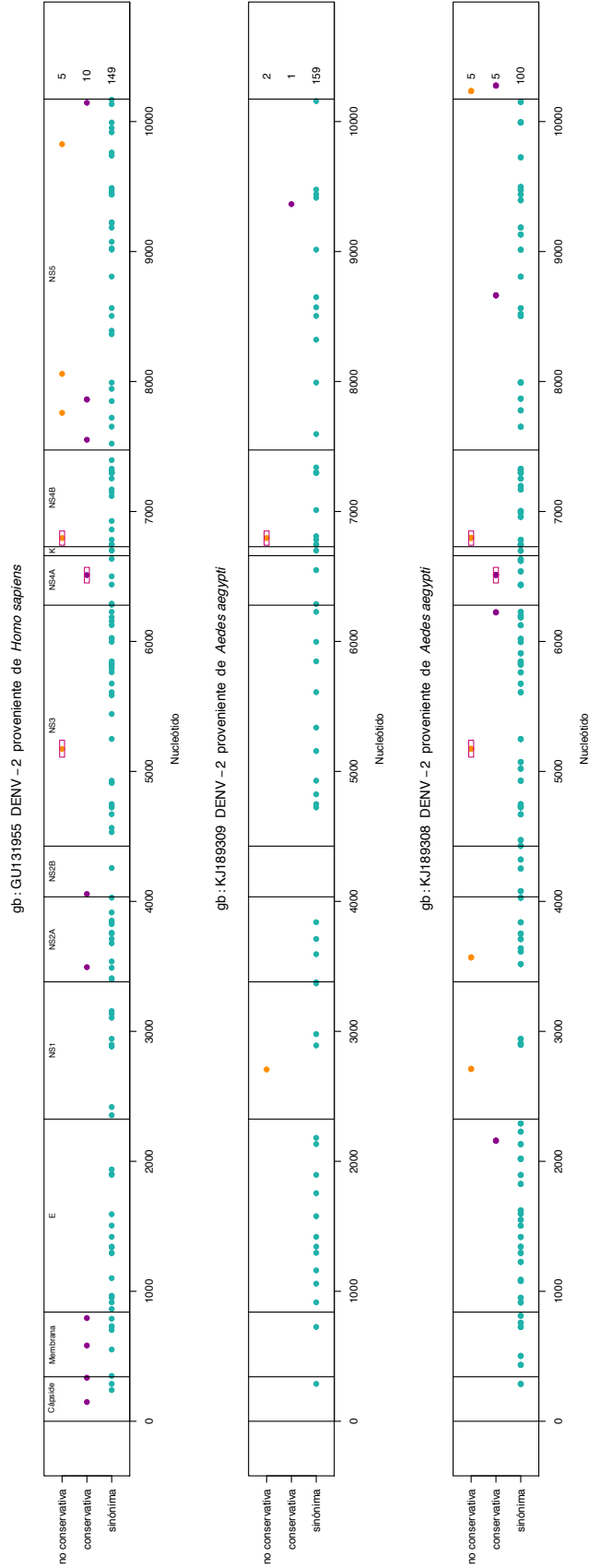
La tabla 7.5 muestra el número de mutaciones sinónimas y no sinónimas (conservativas y no conservativas) comunes a los genomas de MYF1, MYF2 y MYF5 con respecto a tres genomas del DENV de referencia (uno proveniente de suero humano y dos de mosquito). La tabla 7.6 muestra la proporción de dichas mutaciones con relación al tamaño de cada gen. Las variantes de los genomas obtenidos en el presente estudio con respecto a los genomas comparados son preponderantemente sinónimas (90.85% a 98.15%), aunque también se presentan algunas variantes no sinónimas tanto conservativas (6.09% a 0.62%) como no conservativas (3.05% a 1.23%).

**Tabla 7.5.** Número de mutaciones sinónimas y no sinónimas (conservativas y no conservativas) comunes a los genomas de MYF1, MYF2 y MYF5 con respecto a genomas del DENV provenientes de suero humano (GU131955) y *Aedes aegypti* (KJ189309, KJ189308).

	GU131955			KJ189309			KJ189308		
	sinónimas	conser.	no_conser.	sinónimas	conser.	no_conser.	sinónimas	conser.	no_conser.
C	2	2	0	4	0	0	1	0	0
M	6	2	0	8	0	0	5	0	0
E	22	0	0	22	0	0	19	1	0
NS1	13	0	0	11	0	1	3	0	1
NS2A	14	1	0	15	0	0	7	0	1
NS2B	1	1	0	2	0	0	3	0	0
NS3	30	0	1	30	0	0	23	1	1
NS4A	6	1	0	7	0	0	5	1	0
péptido 2K	2	0	0	2	0	0	1	0	0
NS4b	17	0	1	17	0	1	13	0	1
Ns5	36	3	3	41	1	0	20	1	0
<b>Total</b>	<b>149</b>	<b>10</b>	<b>5</b>	<b>159</b>	<b>1</b>	<b>2</b>	<b>100</b>	<b>4</b>	<b>4</b>

**Tabla 7.6.** Proporción de mutaciones por gen en relación con su longitud. Las mutaciones comunes a los genomas de MYF1, MYF2 y MYF5 con respecto a genomas del DENV provenientes de suero humano (GU131955) y *Aedes aegypti* (KJ189309, KJ189308) se clasifican en sinónimas y no sinónimas (conservativas y no conservativas).

	GU131955			KJ189309			KJ189308		
	sinónimas	conserv.	no_conser.	sinónimas	conserv.	no_conser.	sinónimas	conserv.	no_conser.
C	0.58	0.58	0.00	1.17	0.00	0.00	0.29	0.00	0.00
M	1.20	0.40	0.00	1.61	0.00	0.00	1.00	0.00	0.00
E	1.48	0.00	0.00	1.48	0.00	0.00	1.28	0.07	0.00
NS1	1.23	0.00	0.00	1.04	0.00	0.09	0.28	0.00	0.09
NS2A	2.14	0.15	0.00	2.29	0.00	0.00	1.07	0.00	0.15
NS2B	0.26	0.26	0.00	0.51	0.00	0.00	0.77	0.00	0.00
NS3	0.32	0.05	0.00	0.38	0.00	0.00	0.27	0.05	0.00
NS4A	1.57	0.26	0.00	1.84	0.00	0.00	1.31	0.26	0.00
peptido 2K	24.64	0.00	1.45	24.64	0.00	1.45	18.84	0.00	1.45
NS4B	2.28	0.00	0.13	2.28	0.00	0.13	1.75	0.00	0.13
Ns5	1.33	0.11	0.11	1.52	0.04	0.00	0.74	0.04	0.00



**Figura 7.10.** Tipo de variantes comunes a los genomas MYF1, MYF2 y MYF5 al compararlos con genomas del DENV-2 (Asiático-Americano) provenientes de suero de humano (gb:GU131955, Quintana Roo, México, 2004) y de *Aedes aegypti* (gb:KJ189308 y gb:KJ189309, Yucatán, México).

La figura 7.10 muestra la distribución de las variantes a lo largo del genoma que aparentemente es homogénea. Para verificar este hecho, se hizo una prueba de diseño factorial de dos factores<sup>8</sup> (con el paquete R), sin obtener diferencias significativas entre la proporción de mutaciones presentes en los distintos genes ( $p= 0.681$ ). Es posible afirmar que las variantes comunes a los tres genomas obtenidos en el presente estudio se distribuyen de manera homogénea con respecto a los tres genomas con los que fueron comparados.

Son particularmente interesantes los sitios en los que los nucleótidos son comunes a MYF1, MYF2, MYF5, KJ189308 y KJ189309, pero difieren con respecto al genoma del DENV obtenido de suero humano (GU131955). En la tabla 7.7 se muestran las variantes que cumplen los criterios mencionados, éstas podrían ser de carácter mosquito-específico, pudiendo tener papeles importantes en el proceso replicativo dentro del mosquito o favorecer la evasión de su sistema inmune.

**Tabla 7.7.** Posiciones del genoma comunes a MYF1, MYF2, MYF5, KJ189308 y KJ189309, pero variantes con respecto a GU131955. Sólo se muestran las variantes no sinónimas\*, tanto conservativas (C) como no conservativas (NC).

Codón	Aa.	Gen	DENV extraído de <i>Ae. aegypti</i> (MYF1, MYF2, MYF5, KJ189308 y KJ189308)		DENV extraído de suero humano (GU131955)		Tipo*
			Codón	Aa.	Codón	Aa.	
148-150	50	C	CTT	Leucina	TTT	Fenilalanina	C
334-336	112	C	GCG	Alanina	GTG	Valina	C
583-585	195	M	ACC	Treonina	TCC	Serina	C
793-795	265	M	AAG	Lisina	AGG	Arginina	C
3493-3495	1165	NS2A	GTA	Valina	GCA	Alanina	C
4057-4059	1353	NS2B	ATC	Isoleucina	GTC	Valina	C
7552-7554	2518	NS5	GTC	Valina	ATC	Isoleucina	C
7759-7761	2587	NS5	AAT	Asparagina	GAT	Ácido aspártico	NC
8059-8061	2687	NS5	GCA	Alanina	ACA	Treonina	NC
9826-9828	3276	NS5	CCA	Prolina	TCA	Serina	NC
10144-10146	3382	NS5	AGG	Arginina	AAG	Lisina	C

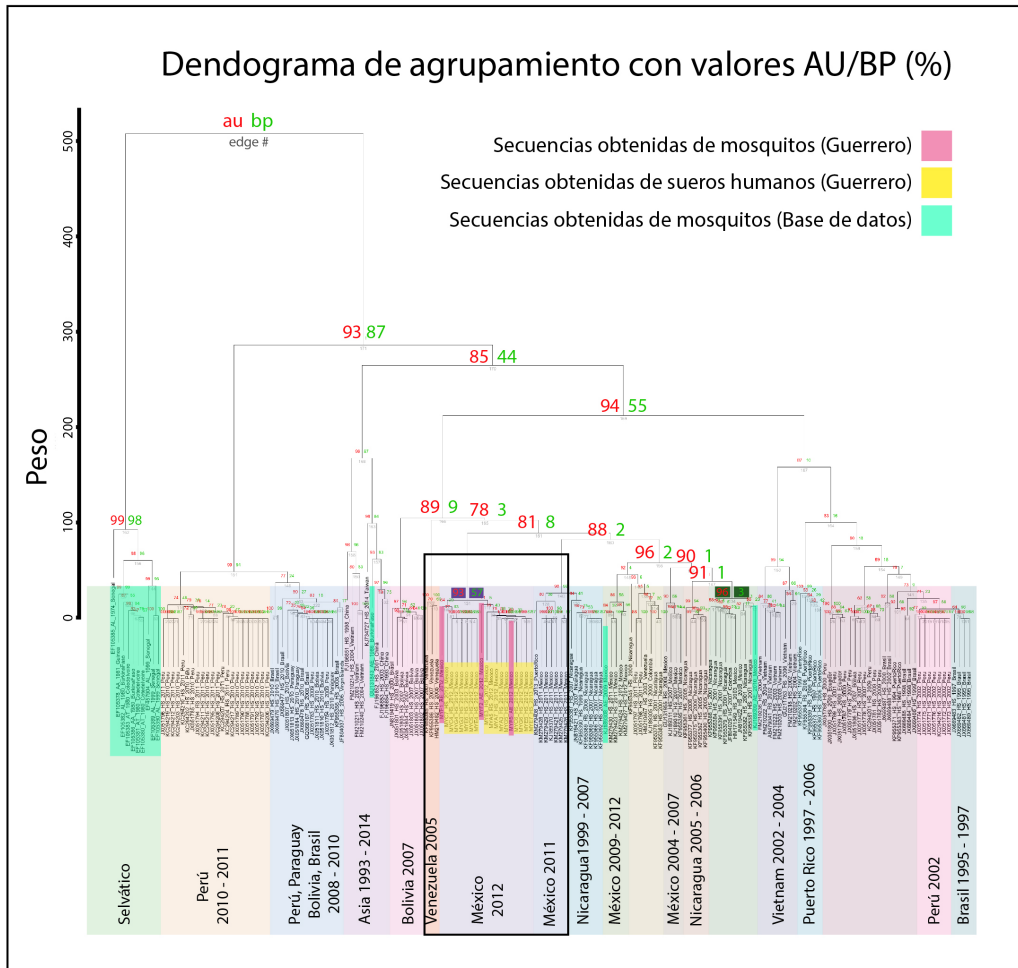
El hecho de que la mayoría de las variantes de los genomas obtenidos en el presente estudio (MYF1, MYF2 y MYF5) sean de carácter sinónimo puede reflejar la preponderancia de la acción de la selección natural negativa. Puede ocurrir también que el blanco de selección sea otro aspecto distinto a la proteína traducida, como el uso relativo de codones sinónimos, secuencias de

<sup>8</sup> Los datos correspondientes al péptido K fueron incorporadas a los del gen NS4A, del que forma parte.

reconocimiento de RNAi del sistema inmune del mosquito o la formación de estructuras secundarias de RNA.

### 7.12 Uso relativo de codones sinónimos

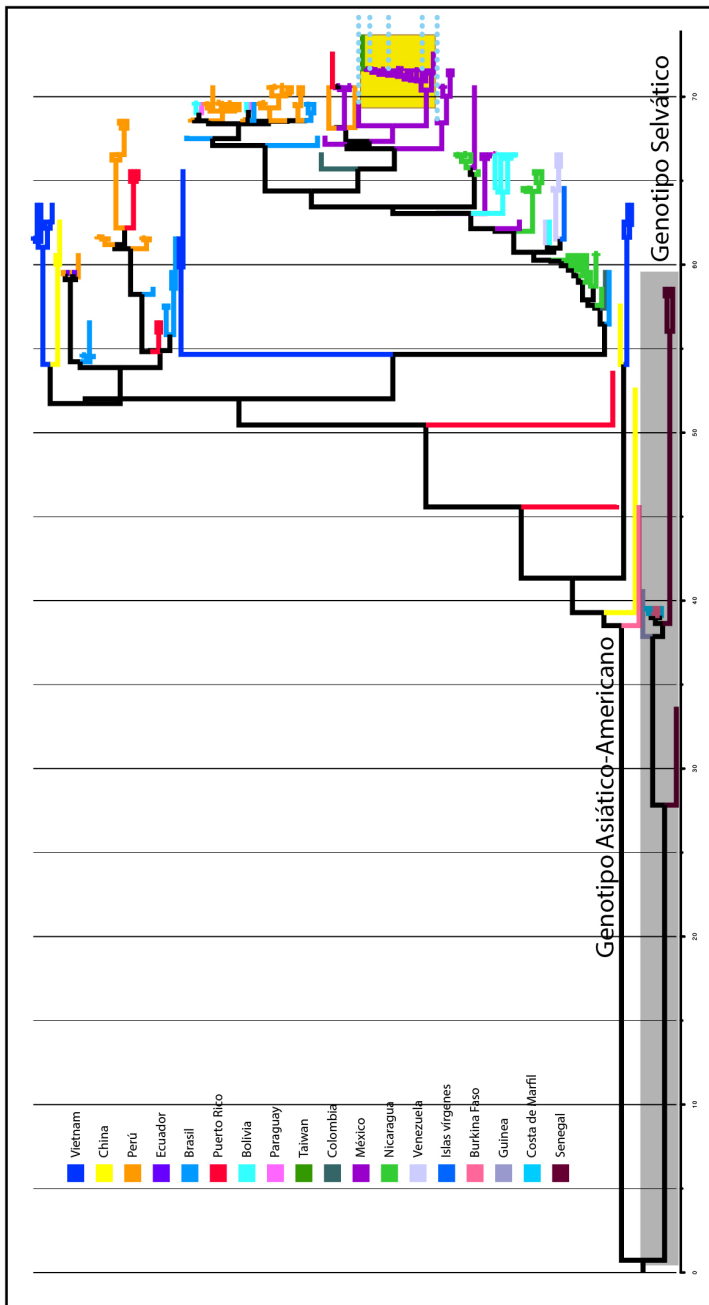
La figura 7.11 muestra la relación entre las fragmentos de las secuencias de E (861 nucleótidos) a partir de su uso relativo de codones sinónimos. Se aprecia que los genomas de DENV-2 extraídos de mosquito se asocian con secuencias extraídas de sueros humanos de la misma distribución geográfica y temporalidad más que con otras secuencias provenientes de mosquito (i.e. KJ189308, KJ189308 o aquellas de genotipo Selvático).



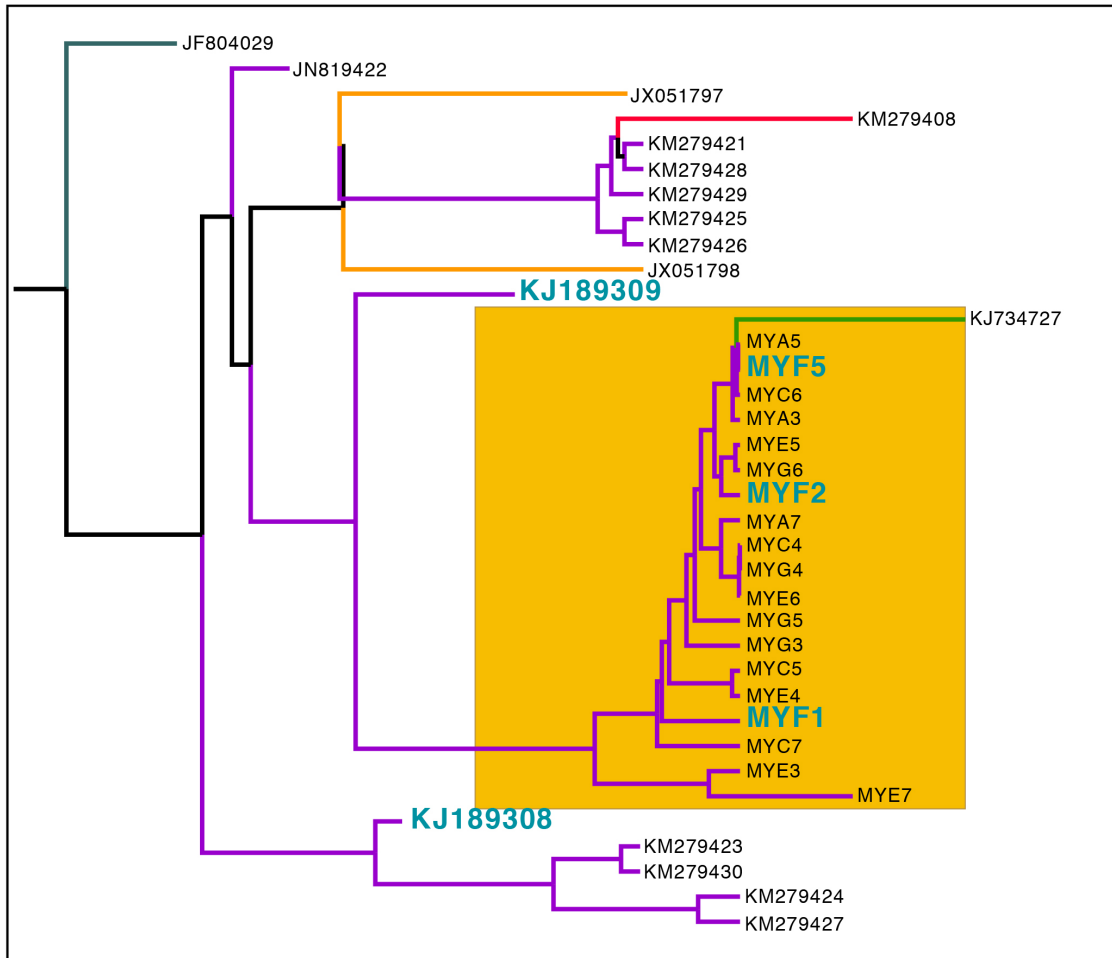
**Figura 7.11.** Dendrograma de agrupamiento de las secuencias según su RSCU. Se observa que, al igual que la mayoría de las secuencias, se asocian por geografía y temporalidad. AU = *approximately unbiased*, BP = *bootstrap probability* (valor de *bootstrap* para cada *cluster*).

### 7.13 Filogeografía

El análisis filogeográfico se refleja en los árboles MCC (Maximum Clade Credibility). En la figura 7.12 se observa que cada una de las secuencias MYF se agrupa con un conjunto de secuencias extraídas de sueros, lo que permite visualizar el dinamismo de los genomas entre humano y mosquito. Las secuencias provenientes de México se van sucediendo entre secuencias de otros países (Nicaragua, Bolivia, Perú, Taiwán, Puerto Rico), lo que refuerza el escenario de introducciones recurrentes de “nuevos” virus.

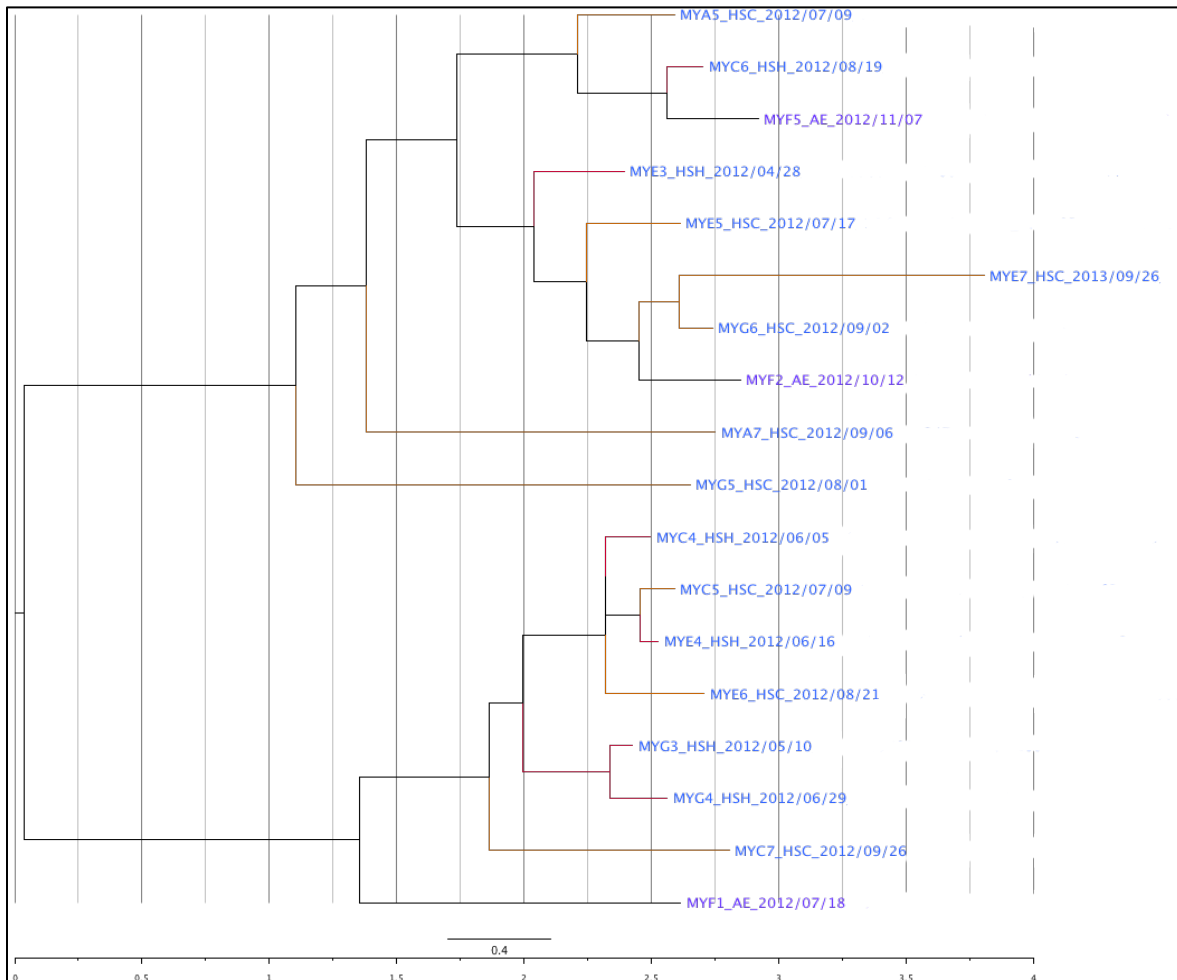


**Figura 7.12.** MCC (*Maximum Clade Credibility Tree*) construido a partir de 174 secuencias del DENV de genotipo Asiático-Americano y Selvático con 861 nucleótidos del gen E (incluidas las generadas en el presente estudio).



**Figura 7.13.** Acercamiento del MCC construido a partir de 174 secuencias del DENV de genotipo Asiático-Americano y selvático con 861 nucleótidos del gen E (incluidas las generadas en el presente estudio).

La figura 7.14 muestra el árbol MCC construido exclusivamente con las secuencias del gen E (861 nucleótidos) del estado de Guerrero obtenidas en este estudio. Se observa que a escala local, no es evidente la estructuración de la población por geografía. Tampoco se encuentra correlación entre el grado de severidad de la enfermedad y la temporalidad o distribución geográfica.



**Figura 7.14.** MCC (*Maximum Clade Credibility Tree*) construido a partir de las secuencias de 861 nucleótidos del gen E obtenidas de muestras de sueros de humanos y mosquitos en el presente estudio. Las leyendas moradas corresponden a las secuencias obtenidas de mosquitos, las azules a aquellas obtenidas de sueros de humanos infectados con DENV. Las líneas amarillas señalan a los pacientes que padecieron dengue clásico y las rojas a los que padecieron dengue hemorrágico.

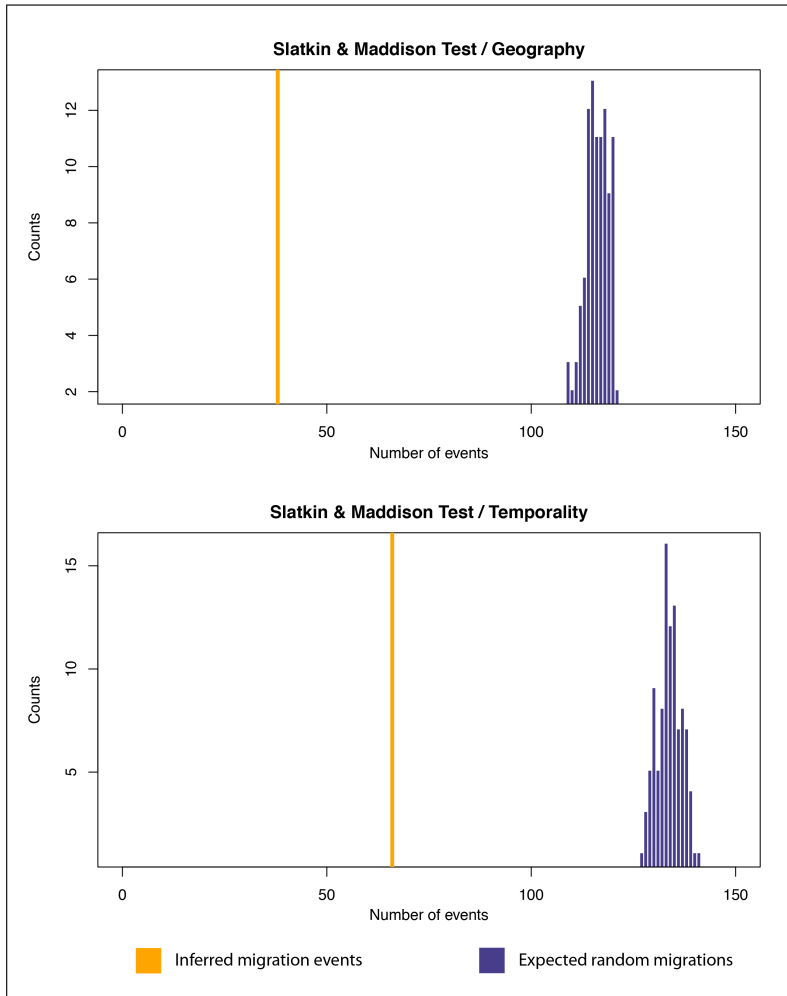
## 7.14 Estructura poblacional

Las poblaciones naturales presentan estructura: algunos individuos están más estrechamente relacionados que otros y uno de los factores que favorecen dicha estructuración es la migración no uniforme de sus elementos (Petkova *et al.*, 2016).

Al aplicar la prueba de Slatkin & Maddison a las secuencias analizadas bajo el criterio de agrupación por geografía, se infirieron 38 eventos de migración con una probabilidad de que este

arreglo se haya obtenido al azar ( $p = 0$ ); al considerar la temporalidad se infirieron 66 eventos de migración con un valor  $p = 0$  (figura 7.15).

Lo anterior refleja el hecho de que la población se estructura tanto por geografía como por temporalidad.



**Figura 7.15.** Resultados de la prueba Slatkin & Maddison considerando estructuración poblacional (Arriba) por geografía y (Abajo) por temporalidad. Gráficas construidas con R a partir de los datos generados por HyPhy.

## 8. Discusión

### 8.1 Genomas obtenidos: profundidades, coberturas, longitud de contigs

Tradicionalmente la detección de virus en mosquitos se ha enfocado a la búsqueda de arbovirus de importancia médica usando PCR y PCR en tiempo real (que no requiere replicación *in vitro*). A pesar de que estos métodos son eficientes y específicos, no permiten obtener mucha información sobre la diversidad viral ni de sus posibles efectos en la evolución (Shi *et al.*, 2015). Con la secuenciación masiva ha sido posible superar estas limitantes, y en los últimos 10 años el estudio del viroma de los mosquitos ha cobrado auge debido al dramático incremento en el interés por el descubrimiento de nuevos virus específicos de insectos (Bolling *et al.*, 2015). Sin embargo, la secuenciación masiva de DENV provenientes de mosquitos no ha sido suficientemente explorada.

A la fecha, no existen reportes que hayan logrado la obtención de genomas completos de DENV en las condiciones empleadas en el presente estudio, con muestras consistentes de *pools* de mosquitos colectados en campo y sin ser replicados en cultivos celulares<sup>9</sup>.

Se ha observado que tanto las transiciones entre hospederos, como el paso de un tipo de tejido a otro (por ejemplo del abdomen hacia las glándulas salivales del mosquito) repercute directamente en la diversidad del DENV (Sessions *et al.*, 2015; Lin *et al.*, 2004). Dado que la intención del presente estudio es analizar la diversidad del virus presente en mosquitos que forman parte de las poblaciones naturales, se prefirió no replicar a los virus en un cultivo celular, pues ello podría haber significado la pérdida de variantes por la transición obligada y por la generación de secuencias clonales en el cultivo (Kunin *et al.*, 2008).

Se logró obtener dos secuencias de genoma consenso completo (MYF1 y MYF5) y una de genoma consenso parcial (MYF2, que presenta un gap de tan sólo 78 nucleótidos) con profundidades de cobertura de 26.6, 734.1 y 12.3 respectivamente.

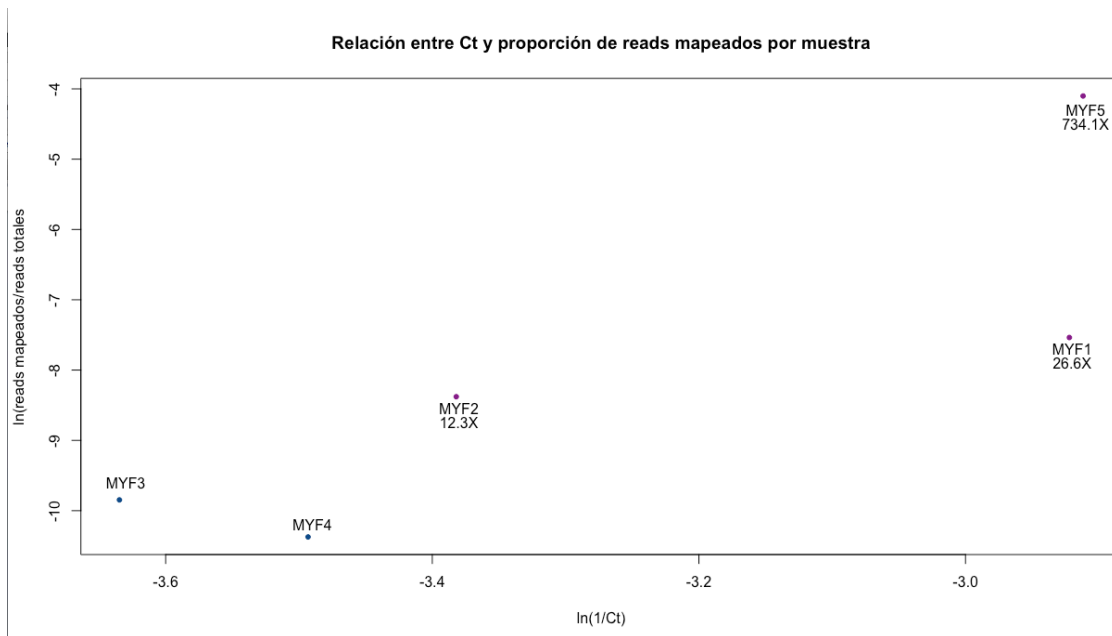
---

<sup>9</sup> Tradicionalmente se hace en células C6/36 de *Aedes albopictus*.

Se observó que la posibilidad de obtener coberturas importantes del genoma (como ocurrió con las muestras MYF1, MYF2 y MYF5) está en estrecha relación con el valor de Ct obtenido en la RT-PCR (tabla 7.2). A pesar de que la relación de las profundidades obtenidas y los valores de Ct es inversa pero no lineal (figura 8.1), resulta claro que el límite para la obtención de genomas completos en las condiciones de este estudio se encuentra cercano al valor  $Ct = 30$ . Se sabe que la proporción de *reads* que mapean contra el genoma de referencia y los *reads* totales (o los contigs que se logran ensamblar *de novo*) están en estrecha relación con la cantidad de partículas virales en la muestra (Marston *et al.*, 2013; Malboeuf *et al.*, 2013). Por los resultados obtenidos podemos considerar que en las condiciones de este estudio, el valor de Ct es un indicador del número de partículas virales.

Los altos valores de Ct de MYF3 y MYF4 pueden ser resultado de una baja concentración de partículas virales en los mosquitos muestreados y/o de degradación posiblemente debida a un mal manejo (ambas muestras fueron colectadas el mismo día y en la misma colonia, tabla 7.3).

El número de partículas virales presentes en los mosquitos que conforman cada muestra depende de múltiples factores, como el número de mosquitos infectados contenidos en cada *pool*, el grado de viremia de la(s) persona(s) de la(s) que se alimentó cada mosquito, el periodo de infección en que se encontraba el mosquito cuando fue capturado, la presencia de variantes del DENV que pudieran favorecer o desfavorecer la replicación en el mosquito, e incluso la cepa del mosquito que, se ha visto, influye en su competencia como vector (Salazar *et al.*, 2010). Ya que generalmente no se cuenta con elementos para poder conocer esa información, es importante para estudios posteriores tener en consideración el valor de Ct de las muestras para la elección de aquellas que se desee secuenciar masivamente.



**Figura 8.1.** Relación entre Ct y profundidad de secuenciación obtenida por mapeo contra genoma de referencia para las distintas muestras.

El hecho de que para la muestra MYF2 exista una región sin *reads* es consistente con la naturaleza de las muestras; se sabe que las secuencias metagenómicas se caracterizan por ser fragmentadas y no totalmente solapantes (Taboada *et al.*, 2014).

Aunque Shi y cols. (2015) no estudiaron el DENV, utilizaron un método muy similar al empleado en el presente estudio. Dichos autores hicieron un análisis metagenómico para tres muestras de mosquitos (de entre 350 y 440 mosquitos cada una) colectados en una zona residencial de Hubei, China. Emplearon un sistema de secuenciación Roche 454 GS FLX que generó lecturas con una longitud promedio de 247 nt y mediante ensamble *de novo* obtuvieron contigs de 327 a 2,083 nucleótidos. Por muestra, el número de *reads* virales que obtuvieron se encontró entre 170 y 28,717 (80,769 nt a 7,093,099 nt).

En el presente estudio se logró obtener contigs de 3,781 nt (MYF2) a 10,648 nt (MYF1), mientras que el número de *reads* (longitud promedio de aprox. 74 nt) virales se encontró entre 1,711 y 101,858 (126,614 nt a 7,537,492 nt). Estos resultados parecen ser comparables a los obtenidos por Shi y cols. (2015).

A pesar de que el análisis metagenómico viral se ha generalizado para muestras de muy diverso tipo, como de tracto respiratorio, heces fecales, sangre, tejidos animales y plantas, entre otros, o biotopos específicos (Shi *et al.*, 2015), es necesario considerar que el rendimiento y las

secuencias que puedan obtenerse dependen en gran medida de una adecuada extracción y que no existe un método de extracción "*one-size-fits-all*" para muestras de tan distinta naturaleza (Kunin *et al.*, 2008). Por los resultados obtenidos de las muestras MYF1, MYF2 y MYF5, es posible afirmar que el método de extracción empleado funciona de manera adecuada para la obtención de genomas completos de DENV a partir de mosquitos adultos capturados en campo.

## 8.2 Genomas “consenso” y variantes

La muestra MYF5 presentó 17 sitios que reflejan variabilidad poblacional de las partículas virales contenidas en los mosquitos que lograron ser secuenciadas, la muestra MYF2 sólo presentó dos sitios variables y MYF1 no presentó ninguna (tabla 6.7). Este hecho, sin embargo, es un reflejo de su baja profundidad, más que de una menor variabilidad.

Tradicionalmente, las secuencias genómicas virales han sido reportadas como genomas únicos por humano o mosquito muestreado, lo que asume implícitamente que un único genoma del DENV representa adecuadamente a la infección dada. Debido a su excepcionalmente alta tasa de mutación (aproximadamente un millón de veces más grande que en los vertebrados), los virus de RNA típicamente existen dentro de sus hospederos como poblaciones heterogéneas de variantes genéticas relacionadas. Esta abundante diversidad de las poblaciones del DENV, acoplada a los tiempos de generación cortos, alimentan su rápida evolución a lo largo de una única infección (Lambrechts & Lequime, 2016; Wash & Díaz, 2015).

Los avances que han tenido las tecnologías de secuenciación en la última década han permitido el análisis de la heterogeneidad y las subpoblaciones de las muestras, y con ello se ha favorecido un cambio de paradigma de una perspectiva de “secuencia consenso” a una “perspectiva poblacional”. El alto rendimiento de la secuenciación masiva paralela permite el muestreo incluso de subpoblaciones raras, y aunque los errores de secuenciación complican la determinación de variantes verdaderas en la población, se han desarrollado métodos que permiten la identificación de variantes de baja frecuencia (Wilm *et al.*, 2012).

El límite de detección de variantes depende principalmente de la profundidad de secuenciación, con mayores profundidades, resultando en mayor sensibilidad (Lambrechts & Lequime, 2016).

Por lo anterior, los genomas de DENV-2 derivados del presente estudio deben tratarse con ciertas consideraciones:

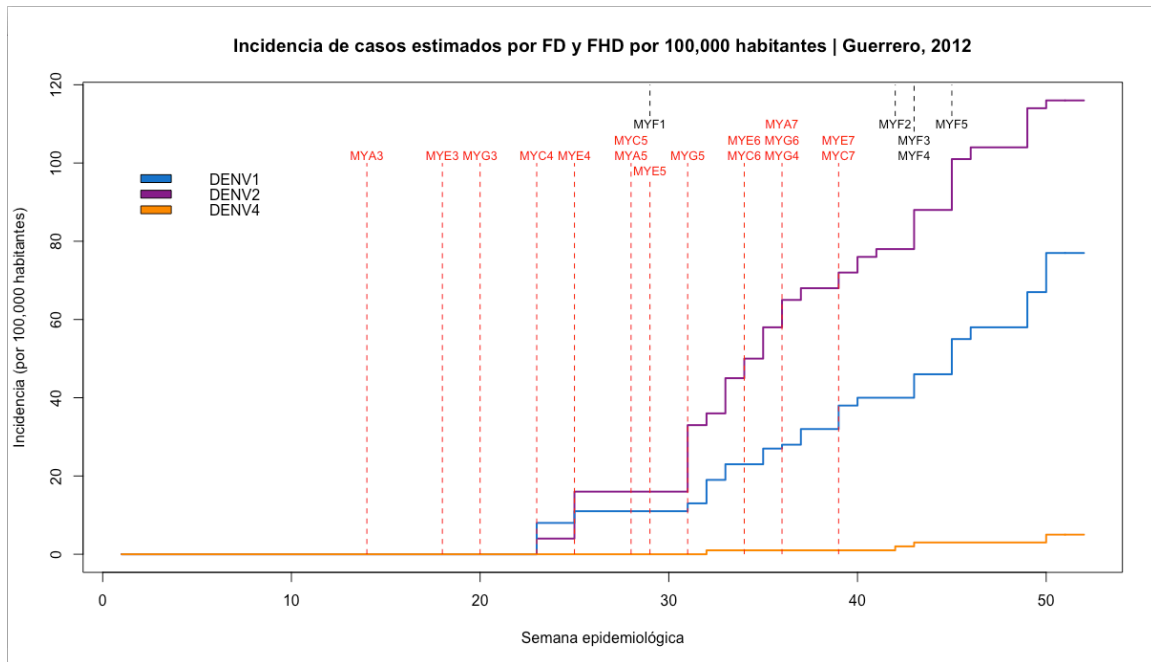
- Los contigs ensamblados son fragmentos genómicos compuestos, pues cada *read* proviene potencialmente de un individuo (virión) distinto dentro de la población del DENV presente en cada mosquito.
- Dada la naturaleza de la muestra, se están considerando no sólo los DENV presentes en un mosquito, sino probablemente los que están presentes en varios.
- Muchas variantes son posiblemente enmascaradas por las secuencias presentes en las poblaciones genómicas dominantes.
- Las muestras MYF1 y MYF2 no necesariamente son menos variables que MYF5; simplemente no tuvieron la profundidad de cobertura necesaria para poder distinguir las variantes de posibles errores de secuenciación.

### **8.3 Serotipo y genotipo reportado, consistente con la diversidad reportada para Guerrero**

Todas las secuencias analizadas en el presente estudio, obtenidas tanto de mosquitos como de sueros humanos, corresponden a DENV-2 de genotipo Asiático-Americano.

El InDRE reporta la presencia de los serotipos DENV-2, DENV-1 y DENV-4 durante el 2012. La figura 8.2 presenta los registros de incidencia de cada uno de los serotipos durante las 52 semanas epidemiológicas de ese año, así como la semana de colecta de las muestras analizadas. Se aprecia que, a pesar de que DENV-1 y DENV-2 tienen registros a partir de la semana epidemiológica número 23, su incidencia en el estado es distinta, siendo mayor para el DENV-2.

Los mosquitos fueron colectados en el interior y en las cercanías de los hogares con casos probables o confirmados de dengue. A pesar de que no se cuenta con la información correspondiente al serotipo y genotipo de dichos pacientes, se observa que presentan el serotipo del DENV de mayor incidencia.



**Figura 8.2.** Incidencia de casos estimados por FD y FHD por 100,000 habitantes durante las 52 semanas epidemiológicas de 2012 en el estado de Guerrero (a partir de los datos del InDRE, 2012). Las líneas punteadas indican las semanas en las que se colectaron los mosquitos (color negro) o en las que se tomaron las muestras de suero de los humanos afectados (color rojo).

#### 8.4 Un nucleótido con variación poblacional relevante

Díaz y cols. (2006) describen que (al igual que en otros países del continente americano) en México la circulación del DENV se ha caracterizado por episodios sucesivos de introducción, expansión, extinción aparente y reintroducción del mismo o de distintos genotipos virales.

El desplazamiento progresivo del antiguo genotipo Americano por virus que se originaron en Asia ocurrió en el Caribe durante la década de 1980, en el norte de Sudamérica y en Brasil en la década de 1990, y en América Central, Bolivia y Perú a finales de 1990. En México, el último reporte del genotipo Americano fue en 1995 (Díaz y cols., 2006) y los primeros reportes del genotipo Asiático-Americano corresponden a aislados de Yucatán en 2001 y 2002 (Loroño-Pino *et al.*, 2004) y de Oaxaca en 2001 (Cisneros *et al.*, 2006). Con respecto a los otros genotipos de DENV-2, el número de reportes es muy escaso. Díaz y cols. (2006) reportan una muestra de genotipo Cosmopolita de 1996 y dos muestras de genotipo Asiático 2 de 1997.

Hasta donde fue posible investigar en este estudio, la extensión actual que tiene el genotipo Asiático-Americano en nuestro país no ha sido documentada; sin embargo, parece que se ha convertido en el genotipo predominante, si no es que el único. La base de datos

especializada ViPR tiene 127 registros de DENV-2 de México, de los cuales 56.7% corresponde a Asiático-Americano (la distribución de genotipos se muestra en la tabla 8.1).

**Tabla 8.1.** Registros de DENV-2 presentes en la base de datos especializada ViPR<sup>10</sup> para México.

Número de secuencias	Genotipo	Fechas de colecta
5	Americano	1983-1994
1	Cosmopolita	2006
72	Asiático-Americano	2002-2012
49	N/A	2002-2012

Se ha considerado que la introducción del genotipo Asiático-Americano fue responsable de los primeros brotes epidémicos de DHF en el continente (Cuba, 1981; Venezuela, 1989 y Brasil, 1990), donde potencialmente desplazó al genotipo nativo Americano. En México, la introducción del genotipo Asiático-Americano en 2001 fue seguida de un incremento en la proporción de casos de las formas severas de la enfermedad (DHF/DSS) (Díaz *et al.*, 2006)<sup>11</sup>.

Salazar y cols. (2010) encontraron que una de las posibles razones que favorecieron el desplazamiento del genotipo Americano por el Asiático-Americano es la competencia diferencial en el vector. Aunque ambos genotipos presentan capacidades equivalentes para infectar los *midguts* de los mosquitos, los virus del genotipo Asiático-Americano se diseminan en forma mucho más eficiente a las glándulas salivales de *Aedes aegypti*, en comparación con los del genotipo Americano (con tasas de diseminación de 60-80% en contraste con ~19% respectivamente). Este hecho se traduce en mayores tasas de replicación y una posible mayor virulencia (Rivera-Osorio *et al.*, 2011).

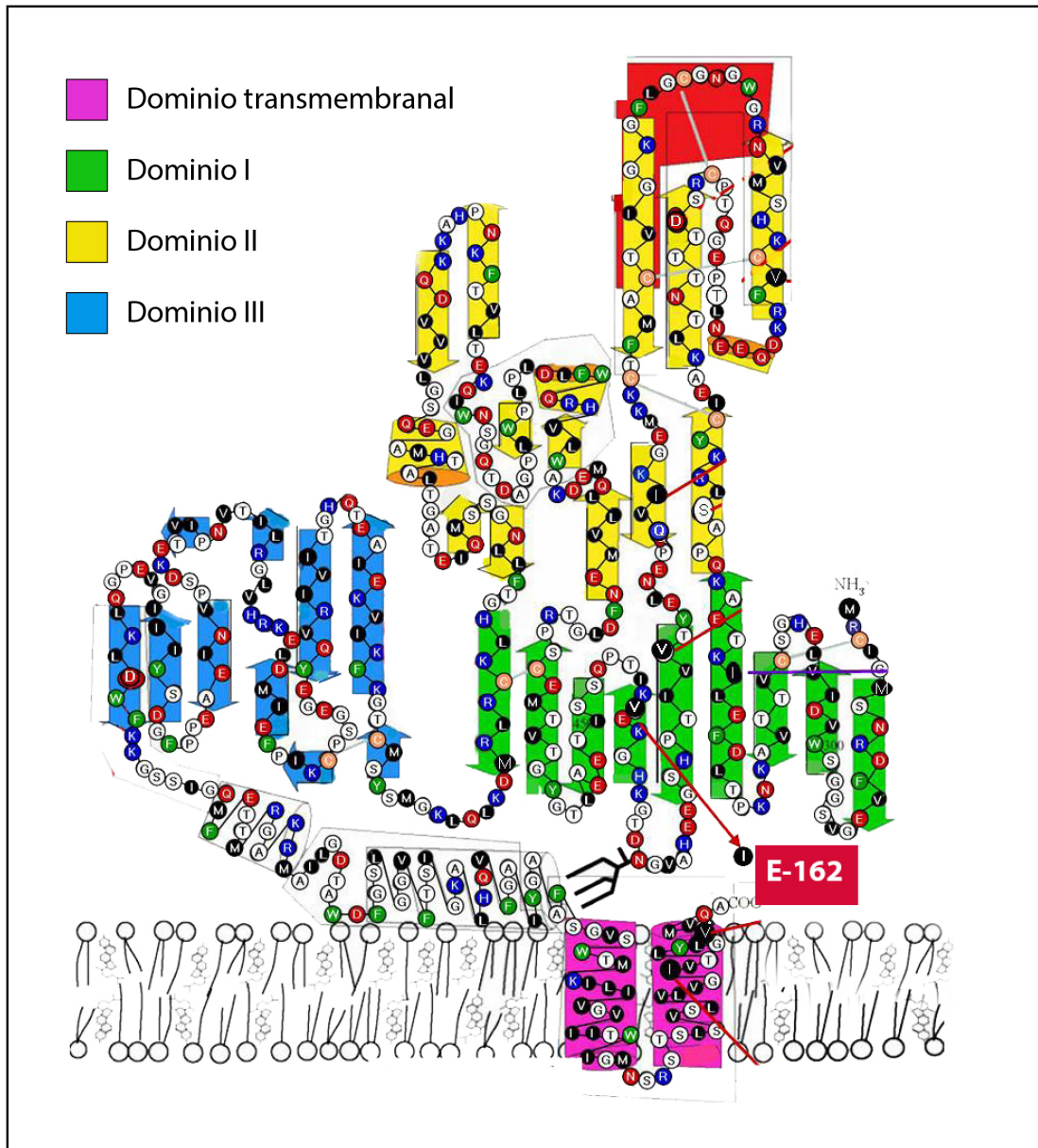
Al comparar las secuencias del gen E para los genotipos Americano y Asiático-Americano, Salazar y cols. (2010) encuentran 14 sustituciones no sinónimas distribuidas en los tres dominios de la proteína, que se asocian con la mayor competencia del genotipo no nativo para replicarse en el mosquito. De las secuencias analizadas en el presente estudio, se comprobó que sólo la correspondiente al aminoácido 162 de la proteína E presenta alguna variación con respecto a lo encontrado por Salazar y cols. (2010). La variante es una mutación en la tercera posición del codón (sinónima) que codifica para el aminoácido isoleucina presente en el dominio I (figura 8.3). La relevancia de esta variante radica en la distribución poblacional que tiene en las muestras analizadas. El 50% de las secuencias presenta el codón ATT y el 50% restante presenta el codón ATC. La figura 8.4 muestra cómo las variantes se presentan incluso entre aquellas

<sup>10</sup> Consultada en diciembre de 2016.

<sup>11</sup> Un primer incremento en la proporción de casos severos se dio a mediados de la década de 1990 y se asocia con la introducción de DENV-3 y/o con la cocirculación de genotipos no americanos (Díaz y cols., 2006).

muestras que fueron colectadas en la misma localidad y el mismo día, hecho que se confirma totalmente con la muestra extraída de mosquito MYF5 que contiene ambas variantes.

Conocer el papel específico que tiene esta transición en la competencia y/o virulencia del virus requerirá de investigaciones posteriores. Rivera-Osorio y cols. (2011) reportan una sustitución de adenina por guanina que también es silenciosa y que se traduce en el aminoácido 296 (lisina) de la proteína E de DENV-1, presente en 33.3% de las secuencias asociadas con DHF y sólo en 8.5% de las secuencias asociadas con DF que analizan. Los autores también consideran que son necesarios estudios posteriores para comprender la relevancia de las mutaciones silenciosas y su posible efecto en la virulencia.

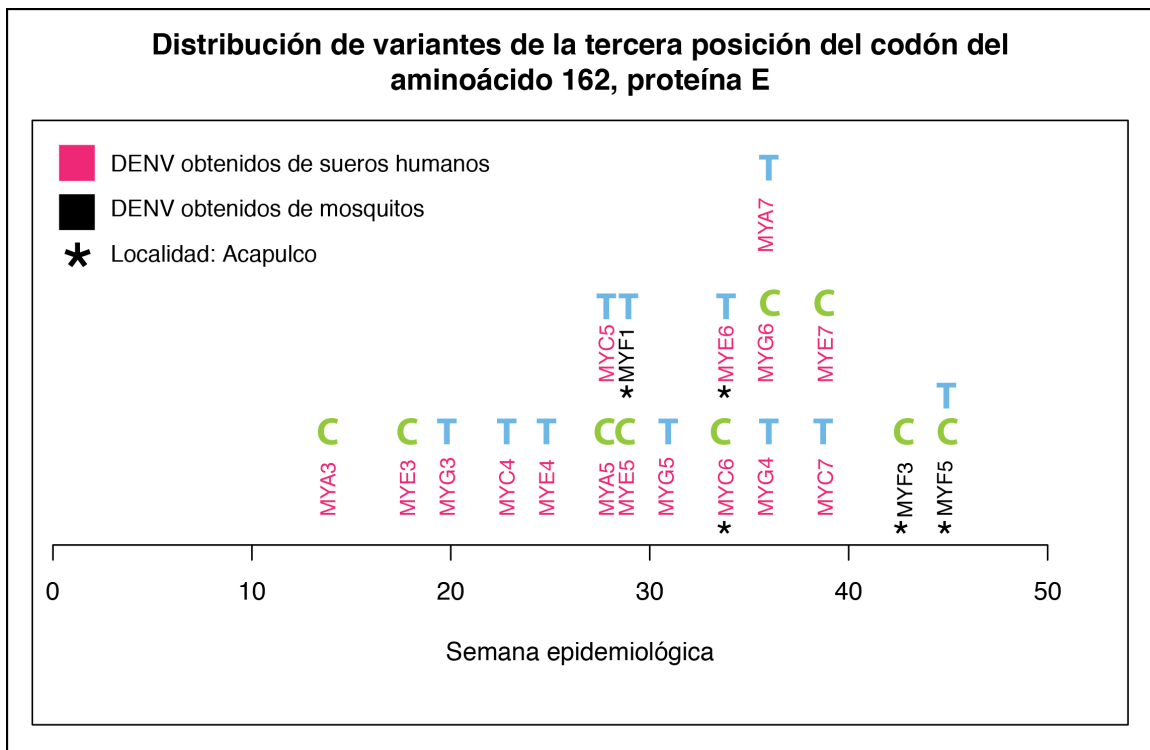


**Figura 8.3.** Sitios variantes que favorecen la replicación del genotipo Asiático-Americano en las glándulas salivales de los mosquitos, incluida la isoleucina 162, variante poblacional de las muestras analizadas en el presente estudio. (Tomada de Salazar y cols., 2010.)

Uno de los mecanismos que podrían estar regulando la heterogeneidad del codón en cuestión es la inmunidad del mosquito.

Se sabe que los RNAi son parte de la principal respuesta antiviral inmune de los mosquitos contra infecciones de arbovirus. Debido a que el RNAi resulta en la degradación de RNA viral a manera de secuencia específica, es un modulador potencial de la dinámica de coevolución intrahuésped entre las poblaciones virales. Lambrechts & Lequime (2016) proponen

que variantes virales raras son selectivamente ventajosas si su secuencia es suficientemente divergente para escapar a la respuesta de RNAi que tiene como blanco la secuencia viral dominante. Aunque el papel que tiene el RNAi en la infección del mosquito por DENV aún no se ha descrito a profundidad, se sabe que porciones del genoma de DENV-2 (Asiático-Americano) desencadenan la respuesta de RNAi (Sánchez-Vargas *et al.*, 2009). Por otra parte, para el virus de West Nile se ha demostrado que, a pesar de que ciertas regiones (particularmente las UTR tanto 5' como 3') son blancos preferentes de las respuesta de RNAi, existen regiones blanco distribuidas a lo largo de todo su genoma (Brackney, 2009).



**Figura 8.4.** Distribución de variantes de la tercera posición del codón del aminoácido 162 de la proteína E en las muestras de DENV de Guerrero. Se observa que es un sitio “muy variable”.

De acuerdo con el escenario planteado por Lambrechts & Lequime (2016), a medida que las variantes de escape aumentan en frecuencia, son cada vez más objetivo de la respuesta de RNAi y permiten que otras variantes raras sean seleccionadas. Si la región que contiene al codón que codifica para el aminoácido 162 de E fuera blanco de un RNAi (p ej., en su forma replicativa de dsRNA), la presión selectiva del RNAi y de la importancia que aparentemente tiene la isoleucina en cuestión podrían explicar la variabilidad nucleotídica que presenta en las secuencias consenso analizadas en este estudio.

Además de la isoleucina 162 existen otros dos codones que presentan variabilidad notoria en las secuencias analizadas (tabla 8.2).

**Tabla 8.2.** Variantes poblacionales de las muestras analizadas.

Aminoácido de la proteína E	Dominio	Variantes, proporción en la población	Codones	Aminoácido(s)
162	I	T-52.6% / C-47.3%	ATT/ATC	Isoleucina
225	II	C-57.9% / T-42.1%	GAC/GAT	Ácido aspártico (D)
360	III	A-73.7% / G-26.3%	GAA/GGA	Ácido glutámico (E) / Glicina (G)

### 8.5 Uso relativo de codones

Velázquez-Salinas y cols. (2016) describen cómo el hecho de que los arbovirus sean totalmente dependientes de los tRNAs de los hospederos para la traducción de su genoma, hace que el uso de codones sea blanco de selección. Los autores demuestran que el Uso Relativo de Codones Sinónimos, RSCU (Relative Synonymous Codon Usage) de distintos arbovirus refleja el de sus hospederos naturales. En contraste con lo anterior, el uso relativo de codones sinónimos de las secuencias analizadas el presente estudio, no se agrupan en función de los hospederos de los que fueron extraídos sino que se agrupan de acuerdo a su espacialidad y temporalidad. Esto refleja, para el caso de los DENV-2 extraídos de mosquito, que no son genomas que hayan estado mucho tiempo aislados en la población de artrópodos, sino que refleja la dinámica característica del ciclo de transmisión humano.

### 8.6 Estructura poblacional

Las poblaciones naturales presentan estructura: algunos individuos están más estrechamente relacionados que otros y uno de los factores que favorecen dicha estructuración es la migración no uniforme de sus elementos (Petkova *et al.*, 2016).

Al aplicar la prueba de Slatkin & Maddison a las secuencias analizadas bajo el criterio de agrupación por geografía, se infirieron 38 eventos de migración con una probabilidad de que este arreglo se haya obtenido al azar ( $p = 0$ ); al considerar la temporalidad se infirieron 66 eventos de migración con un valor  $p = 0$ .

Lo anterior refleja el hecho de que la población se estructura tanto por geografía como por temporalidad.

Al aplicar la prueba de genética de poblaciones Slatkin & Maddison se observó que las secuencias analizadas se estructuran tanto por geografía como por temporalidad. Ambos escenarios han sido descritos, aunque a distintas escalas. Allicock y cols. (2012) indican que el DENV a escala continental (América) presenta una estructura poblacional en función de su distribución geográfica. Describen una dinámica caracterizada por eventos de invasión, dispersión geográfica y flujo génico entre localidades cercanas. Por otra parte, Díaz y cols. (2006) y Carrillo-Valenzo y cols. (2010) encuentran que a escala nacional (México), las secuencias se agrupan por temporalidad. Describen la dinámica del DENV caracterizada por la introducción de múltiples linajes virales, su frecuente reemplazo y extinción, y una baja diversidad en cada brote.

Es posible que esta situación refleje procesos de invasión que agrupan temporalmente y procesos de diversificación que agrupan espacialmente. También es posible que la introducción de linajes virales sea menos frecuente de lo que se ha considerado hasta ahora y que sea la diversidad generada y mantenida en la población de mosquitos y en los individuos asintomáticos (que hasta ahora no ha sido considerada con la relevancia que merece) la que sostiene este comportamiento.

## **8.7 Selección natural**

Se ha demostrado (a partir de la comparación del número de mutaciones sinónimas y no sinónimas) que la diversidad genética del DENV en el hospedero mosquito está determinada esencialmente por la selección natural negativa (purificadora) y la deriva génica (Lequime *et al.*, 2016; Díaz *et al.*, 2010; Lin *et al.*, 2004). También se ha señalado que la selección natural negativa es una fuerza evolutiva importante para los DENV derivados de sueros humanos, especialmente en los genes M, E y NS1, por el hecho de que sus productos de traducción son blancos de la respuesta inmune (Sim & Hibberd, 2016).

La preponderancia de las mutaciones sinónimas sobre las no sinónimas obtenidas a lo largo de los genomas de DENV derivados de mosquitos analizados en este estudio puede reflejar la prevalencia de la selección natural negativa. Aunque se verificó que el uso relativo de codones no es blanco de presiones de selección en nuestras muestras, es posible que las mutaciones aparentemente neutrales tengan en realidad un efecto en el *fitness* viral, como podría ocurrir en el codón 162 de la proteína E. La neutralidad aparente de las mutaciones ha sido documentada para los virus de RNA y DNA por Cuevas y cols. (2012).

En cuanto a la ausencia de selección positiva en las secuencias analizadas, éste es un hecho común en los estudios de secuencias de pacientes infectados con DENV. Thai y cols. (2012) proponen que es consecuencia de que el dengue sea una infección autolimitante, en la que los mecanismos innatos, humorales y celulares remueven a la población viral antes de que alguna evidencia de selección positiva pueda ser detectada.

## 9. Conclusiones

Se logró obtener el genoma consenso de DENV de tres de las cinco muestras que dieron positivo para DENV-2.

Al mapear contra genoma de referencia se logró una cobertura total para los tres genomas, y al ensamblar *de novo* se logró la cobertura total de uno de ellos (MYF1) y casi total de otros dos (MYF2 Y MYF5). Las muestras de las que no se logró obtener ningún contig considerable (MYF3 y MYF4) se caracterizaron por tener Ct mayores a 30 en la RT-PCR.

Las muestras de las que se logró obtener mayor profundidad de cobertura (MYF5) y mayor cobertura al ensamblar *de novo* (MYF1) se caracterizaron por tener en la RT-PC los menores valores de Ct, de aproximadamente 18.

La detección de variantes está en función directa de la profundidad de cobertura obtenida. Fue posible detectar variación poblacional de DENV sólo en la muestra que presentó mayor profundidad de cobertura (MYF5 con 734X). Se detectaron 17 transiciones y una transversión. De las variantes identificadas, cuatro constituyen mutaciones no sinónimas que están presentes en las proteínas no estructurales (NS1, NS3 y NS5) y el resto son mutaciones sinónimas.

El genotipo de todas las muestras analizadas, provenientes tanto de mosquito como de suero humano, es Asiático-Americano, el más reportado para secuencias de México desde 2002 en la base de datos especializada ViPR.

Al considerar sólo aquellas posiciones nucleotídicas que son comunes a los tres genomas derivados de mosquito y compararlas tanto con los otros dos genomas completos obtenidos de mosquito reportados en las bases de datos (KJ189308 y KJ189309), como con un genoma obtenido de suero humano (GU131955), se identificaron 149 posiciones que podrían ser específicas de DENV obtenido de mosquito. Los sitios variables se distribuyen homogéneamente a lo largo de todo el genoma y de aquellos que constituyen mutaciones no sinónimas, ocho son conservativas y tres son no conservativas. Notablemente, todas las no conservativas se encuentran en el gen de la RNA polimerasa (NS5).

Al analizar la diversidad de la región de 861 nucleótidos del gen E tanto de las muestras provenientes de humano como de mosquito, también se encontró que la mayoría de las

mutaciones son sinónimas, lo que puede reflejar la acción de la selección natural negativa reportada por diversos autores para el DENV. Sin embargo, una de las posiciones variables (codón 162 de la proteína E) que ha sido previamente reportada por la relevancia que tiene en la virulencia del genotipo Asiático-Americano presenta un polimorfismo poblacional, con aproximadamente 50% de las secuencias analizadas con la variante T y el otro 50% con la variante C. Se identificaron otras dos posiciones (codones 225 y 360 de la proteína E, con características semejantes) que, aunque no han sido reportadas como relevantes, podrían tener un papel similar.

Se ha propuesto que las mutaciones sinónimas pueden jugar un papel importante para el *fitness* de los DENV, lo que facilita, por ejemplo, la evasión del sistema inmune dentro del mosquito (mediante la interacción o no con RNAi) o favorece el uso de codones de los distintos hospederos. Las posiciones variables mencionadas podrían ser un ejemplo de ello, aunque para corroborarlo sería necesario llevar a cabo más estudios. Según los análisis realizados en el presente estudio, el uso relativo de codones sinónimos no constituye un blanco de selección para las secuencias de DENV analizadas.

A escala continental (analizando secuencias del continente americano de la base de datos ViPR), gracias a los análisis filogeográficos y las pruebas de genética de poblaciones realizadas (Slatkin & Maddison, 1989), fue posible determinar que las secuencias se agrupan tanto por su cercanía geográfica como temporal. Este hecho puede reflejar un número de invasiones menor a lo que hasta ahora se ha considerado que agrupa a las secuencias temporalmente con la colonización y el desplazamiento del clado dominante; y, a la vez, procesos de diversificación dentro de las poblaciones establecidas, en los que el papel de los humanos infectados asintomáticos y los hospederos mosquitos (con capacidad de transmisión transovárica) tiene más relevancia de lo que hasta ahora se ha estimado.

Al comparar el fragmento del gen E de las secuencias del DENV obtenidas de mosquito y de suero humano, se observó que se agrupan de forma que reflejan la dinámica de transmisión horizontal característica del ciclo humano. Los DENV extraídos de mosquito reflejan la diversidad presente en los humanos infectados, por lo que la hipótesis de trabajo es corroborada.

## 10. Perspectivas

Sería importante analizar el papel que tienen las posiciones reportadas en el presente estudio como posibles sitios específicos de DENV derivados de mosquitos, así como aquellas que constituyen sitios polimórficos poblacionales en las secuencias de DENV derivados tanto de mosquitos como de humanos, en la respuesta inmune del artrópodo hacia el virus (RNAi); además, analizar las regiones no traducidas de los genomas que se lograron ensamblar y realizar un análisis de búsqueda de otros virus con los *reads* generados en la secuenciación que no hicieron *match* con *Flaviviridae*.

Es necesario continuar con el estudio de la diversidad del DENV y de otros arbovirus en sus hospederos artrópodos, pues esto constituye una importante aproximación para monitorear la introducción, el movimiento y el tráfico de virus, así como para predecir potenciales consecuencias epidemiológicas de dichos eventos.

## 11. Referencias

1. Aaskov J., Buzacott K., Thu H. M. *et al.* (2006) Long-term transmission of defective RNA viruses in humans and *Aedes* mosquitoes. *Science*. 311:236-238.
2. Agilent (2005) Agilent 2100 Bioanalyzer User's guide for molecular assays. <http://ipmb.sinica.edu.tw/microarray/index.files/Agilent%202100%20Bioanalyzer%20user%20guide.pdf>
3. Allicock, O. M., Lemey, P., Tatem, A. J. *et al.* (2012) Phylogeography and population dynamics of dengue viruses in the Americas *Mol. Biol. Evol.* 29(6):1533-1543.
4. Andrews S. (2010) FastQC: a quality control tool for high throughput sequence data. Disponible en: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
5. Awaidi, S. T. A., Obeidani, I. A., Bawikar, S. *et al.* (2014) Dengue epidemiological trend in Oman: a 13-year national surveillance and strategic proposition of imported cases. *Tropical Doctor*. 44(4): 190–195.
6. Azhar, E. I., Hashem, A. M., El-Kafrawy, S. *et al.* (2015) Complete genome sequencing and phylogenetic analysis of dengue type 1 virus isolated from Jeddah, Saudi Arabia. *Virology Journal*. 12:1.
7. Beltramello M., Williams K. L., Simmons C. P. *et al.* (2010) The human immune response to Dengue Virus is dominated by highly cross-reactive antibodies endowed with neutralizing and enhancing activity. *Cell Host Microbe*. 8(3):271-83.
8. Bennett K. E., Olson K. E., Munoz M. *et al.* (2002) Variation in vector competence for dengue 2 virus among 24 collections of *Aedes aegypti* from Mexico and the United States. *Am J Trop Med Hyg*. 67:85–92.
9. Bhatt, S., Gething, P. W., Brady, O. J. (2013) The global distribution and burden of dengue. *Nature*. 496(7446): 504–507.
10. Bolling, B. G., Weaver, S. C., Tesh, R. B. *et al.* (2015) Insect-Specific Virus Discovery: Significance for the Arbovirus Community. *Viruses* 7: 4911-4928.
11. Brackney D. E., Beane J. E. & Ebel G. D. (2009). RNAi targeting of West Nile virus in

- mosquito midguts promotes virus diversification. *PLoS Pathog* 5(7):e1000502.
12. Carrillo-Valenzo, E., Danis-Lozano, R. *et al.* (2010) Evolution of dengue virus in Mexico is characterized by frequent lineage replacement. *Arch. Virol.* 155(9): 1401-12.
  13. Carrington L. B., Nguyen H. L., Nguyen N. M., Duong T. H. K., Tuan T. V., Giang N. T. *et al.* (2015) Naturally-acquired dengue virus infections do not reduce short-term survival of infected *Aedes aegypti* from Ho Chi Minh City, Vietnam. *Am J Trop Med Hyg.* 92:492–6.
  14. Chao D. Y., King C. C., Wang W. K. *et al.* (2005) Strategically examining the full-genome of dengue virus type 3 in clinical isolates reveals its mutation spectra. *Virol J.* 2:72.
  15. Cisneros, A., Díaz-Badillo, A., Cruz-Martínez, G. *et al.* (2006) Dengue 2 genotypes in the state of Oaxaca, Mexico. *Archives of virology.* 151:113-125.
  16. Coffey L. L., Beeharry Y., Bordería A. V., Blanc H., Vignuzzi M. (2011) Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc Natl Acad Sci U S A.* 108:16038–43.
  17. Coffey L. L., Vignuzzi M. (2011) Host alternation of Chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. *J Virol.* 85:1025–35.
  18. Cruz, L. C., Pereira S., O., Leal-Santos, F. A. *et al.* (2015) Natural transovarial transmission of dengue virus 4 in *Aedes aegypti* from Cuiabá, State of Mato Grosso, Brazil. *Revista da Sociedade de Brasileira de Medicina Tropical.* 48(1):18-25.
  19. Cuevas J. M., Domingo-Calap P. & Sanjuán R. (2012) The fitness effects of synonymous mutations in DNA and RNA viruses. *Mol Biol Evol.* 29: 17–20.
  20. Danecek P., Auton A., Abecasis G., *et al.* (2011) The variant call format and VCFtools. *Bioinformatics.* 27(15):2156-8. BCFtools, disponible en <https://samtools.github.io/bcftools/>
  21. De Coffey, L. L. Page, B. L., Greninger, A. L. *et al.* (2014) Enhanced arbovirus surveillance with deep sequencing: identification of novel rhabdoviruses and bunyaviruses in Australian mosquitoes. *Virology.* 448:146-158.
  22. Descloux E., Cao-Lormeau, V. M., Roche C. *et al.* (2009) Dengue 1 diversity and microevolution, French Polynesia 2001-2006: connection with epidemiology and clinics. *PLoS Negl. Trop. Dis.* 3:e493.
  23. Deubel, V., Kinney, R. M. & Trent, D. W. (1988) Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins of dengue type 2 virus, Jamaica genotype: comparative analysis of the full-length genome. *Virology.* 165: 234-244.
  24. De Thoisy, B., Lacoste, V., Germain, A. *et al.* (2009) Dengue infection in neotropical forest mammals. *Vector-borne and zoonotic diseases.* 9(2): 157- 169.

25. Diallo, M., Ba, Y., Sall, A. A. *et al.* (2003) Amplification of the Sylvatic Cycle of Dengue Virus Type 2, Senegal, 1999–2000: Entomologic Findings and Epidemiologic Considerations. *Emerg Infect Dis.* 9(3): 362–367.
26. Díaz, F. J., Black IV, W. C. *et al.* (2006) Dengue virus circulation and evolution in Mexico: a phylogenetic perspective. *Archives of Medical Research.* 37:760-773
27. Drummond A. J., Suchard M. A., Xie D. & Rambaut A. (2012) Bayesian phylogenetics with BEAUti and the BEAST. *Molecular Biology And Evolution.* 29: 1969-1973. Disponible en <http://beast2.org/>
28. Dugan, V., Saira, K. & Ghedin, E. (2012) Large-scale sequencing and the natural history of model human RNA viruses. *Future Virol.* (6):563-573.
29. Dzul-Manzanilla, F., Martínez N., Cruz-Nolasco, M. *et al.* (2015) Arbovirus surveillance and first report of Chikungunya Virus in wild populations of *Aedes aegypti* from Guerrero, Mexico. *J Am Mosq Control Assoc.* 31(3):275-7.
30. Espinosa, M., Giamperetti, S., Abril, M. & Seijo, A. (2014) Vertical transmission of dengue virus in *Aedes aegypti* collected in puerto Iguazú, Misiones, Argentina. *Rev. Int. Med. Trop. Sao Paulo.* 56(2):165-167.
31. Espinoza-Gómez, F., López-Lemus, A. U., Rodríguez-Sánchez, I. P. *et al.* (2011) Detection of sequences from a potentially novel strain of cell fusing agent virus in Mexican *Stegomyia (Aedes) aegypti* mosquitoes. *Archives of virology* 156(7):1263-1267.
32. Farci, P., A. Shimoda, A. Coiana, G. Diaz, G. Peddis, J. C. Melpolder, A. Strazzer, D. Y. Chien, S. J. Munoz, A. Balestrieri, R. H. Purcell, and H. J. Alter. (2000) The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science.* 288:339–344.
33. Fernandez-Garcia, M.D., Mazzon, M., Jacobs, M. & Amara, A. (2009) Pathogenesis of flavivirus infections: using and abusing the host cell. *Cell host microbe* 5:318-328.
34. Fitzpatrick K. A., Deardorff E. R., Pesko K., Brackney D. E., Zhang B., Bedrick E. *et al.* Population variation of West Nile virus confers a host-specific fitness benefit in mosquitoes. *Virology.* 2010; 404: 89–95.
35. Fu, J., Tan, B. H., Yap, E. H. *et al.* (1992) Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90). *Virology.* 188: 953-958.
36. García-Rejón, J. E., Loroño-Pino, M. A., Farfán-Ale, J. A. *et al.* (2011) Mosquito infestation and Dengue Virus infection in *Aedes aegypti* females in schools in Mérida, México. *Am. J. Trop. Med. Hyg.* 84(3):489–496.
37. Gardella-García, C. Pérez-Ramírez, G. *et al.* (2008) Specific genetic markers for detecting subtypes of dengue virus serotype-2 in isolates from the states of Oaxaca and Veracruz,

- Mexico. *BMC Microbiology*. 8:117.
38. Gómez-Dantés, H., Farfán-Ale. J. A. & Sarti, E. (2014) Epidemiological trends of dengue disease in Mexico (2000-2011): a systematic literature search and analysis. *PLOS Neglected Tropical Diseases*. 8(11): e3158.
  39. Gouy M., Guindon, S. & Gascuel, O. (2010) SeaView version 4 : a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* 27(2):221-224. SeaView disponible en <http://www.molrev.org/software/alignment/seaview>
  40. Guzman, M. G., Halstead, S. B., Artsob, H. *et al.* (2010) Dengue: a continuing global threat. *Nature Reviews Microbiology*. S7-S16.
  41. Holland, J. J., J. C. De la Torre, and D. A. Steinhauer (1992) RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* 176:1–20.
  42. Holmes. E. C. (2008) Evolutionary History and phylogeography of human viruses. *Annu. Rev. Microbiol.* 62:307-328.
  43. Holmes, E. C. & Twiddy, S. S. (2003) The origin, emergence and evolutionary genetics of dengue virus. *Infection, Genetics and Evolution*. 3:19-28.
  44. Hull, B., Tikasingh, E., De Souza, M & Martínez, R. (1984) Natural transovarial transmission of dengue 4 virus in *Aedes aegypti* in Trinidad. *American Journal of Tropical Medicine and Hygiene*. 33:1248-1250.
  45. Ibáñez-Bernal, S. (1987) Nuevo Registro Altitudinal de *Aedes (Stegomyia) aegypti* (Linnaeus, 1762) (Diptera: Culicidae) en México. *Folia Entomol Mex.* 72:163-164.
  46. Ibáñez-Bernal, S. & Gómez, H. (1995) Los vectores del dengue en México: una revisión crítica. *Salud Pública de México*. 37(Su1):53-63.
  47. InDRE. (2012) Panorama epidemiológico de fiebre por dengue y fiebre hemorrágica por dengue en entidades federativas. Sistema Nacional de Vigilancia Epidemiológica.
  48. Kunin, V., Copeland, A., Lapidus, A. *et al.* (2008). A bioinformatician's guide to metaenomics. *Microbiology and molecular biology reviews*. 72(4): 557-578.
  49. Kyle, J. L. & Harris, E. (2008). Global spread and persistence of dengue. *Annu. Rev. Microbiol.* 62:71-92.
  50. Lambrechts, L. & Lequime, S. (2016) Evolutionary dynamics of dengue virus populations within the mosquito vector. *Current Opinion in Virology*. 21:47-53.
  51. Lambrechts, L., Fansiri, T., Pongsiri, A. *et al.* (2011) Dengue-1 virus clade replacement in Thailand associated with enhanced mosquito transmission. *Journal of Virology*. 86(3):1853-61.

52. Langmead B. & Salzberg S. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 9:357-359. Bowtie2 disponible en <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>
53. Le Goff, G., Revollo, J., Guerra, M. *et al.* (2011) Natural vertical transmission of dengue viruses by *Aedes aegypti* in Bolivia. *Parasite* 18:277-280.
54. Lequime S., Fontaine A., Ar Gouilh M. *et al.* (2016) Genetic drift, purifying selection and vector genotype shape dengue virus intra-host genetic diversity in mosquitoes. *PLoS Genet* 12:e1006111.
55. Li H., Handsaker B., Wysoker A. *et al.* (2009) The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*. 25 (16): 2078-9. Samtools, disponible en: <http://samtools.sourceforge.net/>
56. Li, W. & Godzik. A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 22:1658-9. Cd-hit, disponible en: <http://weizhongli-lab.org/cd-hit/download.php>
57. Li, Y-L., Weng, J-C., Hsiao, C. *et al.* (2015) PEAT: an intelligent and efficient paired-end sequencing adapter trimming algorithm. *BMC Bioinformatics*. 16(Suppl 1):S2.
58. Lindenbach, B. D., & C. M. Rice. (2001) Flaviviridae: the viruses and their replication, p. 991–1041. En: D. M. Knipe, D. M., Howley, P. M. *et al.* (eds.), *Fields virology*, 4th ed. Lippincott Raven, Philadelphia.
59. Lin, S., Hsieh, S., Yueh, Y. *et al.* (2004) Study of sequence variation of Dengue Type 3 virus in naturally infected mosquitoes and human hosts: implications for transmission and evolution. *Journal of Virology*. 78(22): 12717-12721.
60. Loroño-Pino M. A., Farfan-Ale J. A., Zapata-Peraza A. L. *et al.* (2004) Introduction of the American/Asian genotype of dengue 2 virus into the Yucatan State of Mexico. *Am J Trop Med Hyg*. 71: 485-92.
61. Lozano-Fuentes, S., Fernández-Salas, I., Muñoz, M. L. *et al.* (2009) The neovolcanic axis is a barrier to gene flow among *Aedes aegypti* populations in Mexico that differ in vector competence for dengue 2 virus. *PLoS Negl Trop Dis*. 3:e468.
62. Machow, E., Makino, Y., Zhao, B. *et al.* (1987) The nucleotide sequence of dengue type 4 virus: analysis of genes coding for nonstructural proteins. *Virology*. 159:217-228.
63. Malboeuf, Ch. M., Yang, X., Charlebois, P. *et al.* (2013) Complete viral RNA genome sequencing of ultra-low copy samples by sequence-independent amplification. *Nucleic Acids Research*. 41(1) e13.
64. Martell, M., Esteban, J. I., Ouer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J., and

- Gomez, J. (1992) Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J. Virol.* 66:3225–3229.
65. Marston, D. A., McElhinney, L. M., Ellis, R. J. *et al.* (2013) Next generation sequencing of viral RNA genomes. *BMC Genomics* 14:444.
66. Martínez, N. E., Dzul-Manzanilla, F., Gutiérrez-Castro, C. *et al.* (2014) Natural vertical transmission of Dengue-1 virus *Aedes aegypti* populations in Acapulco. Mexico. *J Am Mosq Control Assoc* 30:143-146.
67. Martinis V. E. P., Alencar, C. H., Kamimura, M. T. *et al.* (2012) Occurrence of natural vertical transmission of dengue-2 and Dengue-3 viruses in *Aedes aegypti* and *Aedes albopictus* in Fortaleza, Ceará, Brazil. *PLoS ONE*. 7(7):e41386.
68. Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2015) Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32(2):292-4. Qualimap 2, disponible en [http://qualimap.bioinfo.cipf.es/doc\\_html/intro.html](http://qualimap.bioinfo.cipf.es/doc_html/intro.html)
69. Osatomi, K. & Sumiyoshi, H. (1990) Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology*. 176: 643-647.
70. Parameswaran, P., Charlebois, P., Tellez, Y. *et al.* (2012) Genome-wide patterns of intrahuman Dengue Virus diversity reveal associations with viral phylogenetic clade and interhost diversity. *Journal of Virology*. 86(16): 8546–8558.
71. Pérez-Ramírez, G., Díaz-Badillo, A., Camacho-Nuez, M. *et al.* (2009) Multiple recombinants in two dengue virus, serotype-2 isolates from patients from Oaxaca, Mexico. *BMC Microbiology*. 9:260.
72. Petkova D., Novembre J. & Stephens, M. 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nat Genet.* 48(1): 94–100.
73. Rambaut, A. & Drummond, J. (2007) *Tracer v1.4* and higher. Disponible en: <http://tree.bio.ed.ac.uk/software/tracer/>
74. Rico-Hesse, R. (2003) Microevolution and virulence of dengue viruses. *Adv Virus Res.* 59:315-341.
75. Rivera-Osorio, P., Vaughan, G., Ramírez-González, J. E. *et al.* (2011) Molecular epidemiology of autochthonous dengue virus strains circulating in Mexico. *Journal of Clinical Microbiology*. 49(9):3370-3374.
76. Salazar, M. I., Loroño-Pino, M. A., Farfán-Ale, J. A. *et al.* (2010) American and American/Asian genotypes of dengue virus differ in mosquito infection efficiency: candidate molecular determinants of productive vector infection. *Rev Biomed.* 21:121-135.
77. Salazar, M. I., Richardson, J. H. *et al.* (2007) Dengue virus type 2: replication and tropisms in

- orally infected *Aedes aegypti* mosquitoes. *BMC Microbiology*, 7:9.
78. Sánchez-Vargas, I., Scott, J. C., Poole-Smith, B. K. *et al.* (2009) Dengue virus type2 infections of *Aedes aegypti* are modulated by the mosquito's RNA interference pathway. *PLoS Pathog.* 5(2): e1000299.
  79. Sessions O. M., Wilm A., Kamaraj U. S. *et al.* (2015) Analysis of dengue virus genetic diversity during human and mosquito infection reveals genetic constraints. *PLoS Negl Trop Dis* 9:e0004044.
  80. Sharpton, T. J. (2014) An introduction to the analysis of shotgun metagenomic data. *Frontiers in plant science.* 5(209):1-14.
  81. Shi, C., Liu, Y., Hu, X. *et al.* (2015) A Metagenomic Survey of Viral Abundance and Diversity in Mosquitoes from Hubei Province. *PLoS ONE.* 10(6): e0129845.
  82. Sim, S. & Hibberd, M. L. (2016) Genomic approaches for understanding dengue: insights from the virus, vector and host. *Genome Biology.* 17:38.
  83. Sim, S., Aw, P. P., Wilm A. *et al.* (2015) Tracking Dengue Virus Intra-host Genetic Diversity during Human-to- Mosquito Transmission. *PLoS Negl Trop Dis.* 9(9): e0004052.
  84. Sim S., Jupatanakul N., Dimopoulos G. (2014) Mosquito Immunity against Arboviruses. *Viruses.* 6:4479–4504.
  85. Sim, S., Jupatanakul, N., Ramirez, J. L. *et al.* (2013) Transcriptomic profiling of diverse *Aedes aegypti* strains reveals increased basal-level immune activation in dengue virus-refractory populations and identifies novel virus–vector molecular Interactions. *PLoS Negl Trop Dis.* 7:e2295.
  86. Sinave, DGE & Salud (2012) Perfil Epidemiológico del dengue en México. México.
  87. Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 123:603-613.
  88. Stucky B. J. (2012) SeqTrace: a graphical tool for rapidly processing DNA sequencing chromatograms. *J Biomol Tech.* 23(3):90-3.
  89. Sylla, M., Bosio, C., Urdaneta-Marquez, L. *et al.* (2009) Gene flow, subspecies composition, and dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *PLoS Negl Trop Dis.* 3:e408.
  90. Taboada, B., Espinoza, M. A., Isa, P. *et al.* (2014) Is There Still Room for Novel Viral Pathogens in Pediatric Respiratory Tract Infections? *Plos One.* 9(11): e113570.
  91. Talavera, G. & Manrique-Saide, P. (2014) Natural vertical transmission of Dengue-1 virus in *Aedes aegypti* populations in Acapulco, Mexico. *Journal of the American Mosquito Control Association.* 30(2):143-146

92. Thai, K., Henn, M. R., Zody, M. *et al.* (2012) High-Resolution analysis of intrahost genetic diversity in dengue virus serotype 1 infection identifies mixed infections. *Journal of Virology*. 86: 835–843.
93. Torres-Galicia, I., Cortés-Poza, D. & Becker, I. (2014) Dengue en México: análisis de dos décadas. *Gaceta Médica de México*. 150:122-127.
94. Twiddy, S., Farrar, J. J. *et al.* (2002) Phylogenetic Relationships and Differential Selection Pressures among Genotypes of Dengue-2 Virus. *Virology*. 298: 63-72.
95. Vasilakis, N., Cardoso, J., Hanley, A., Holmes, E. & Weaver, S. C. (2011) Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health. *Nat. Rev. Microbiol.* 9(7):532-541.
96. Vecna Technologies (2016). Virus genotyping: standard operating procedure. [https://www.viprbrc.org/brcDocs/documents/ViPR\\_Genotype\\_SOP.pdf](https://www.viprbrc.org/brcDocs/documents/ViPR_Genotype_SOP.pdf)
97. Velandia, M. L. & Castellanos, J. E. (2011) Virus del dengue: estructura y ciclo vital. *Infectio*. 15(1):33-43.
98. Velázquez-Salinas, L., Zárate, S., Eschbaumer, M. *et al.* (2016) Selective Factors Associated with the Evolution of Codon Usage in Natural Populations of Arboviruses. *Plos One*. 11(7):e0159943.
99. Wang, W. K., Lin S. R., Lee C. M. *et al.* (2002a) Dengue type 3 virus in plasma is a population of closely related genomes: quasispecies. *J. Virol.* 76:4662– 4665.
100. Wang W. K., Sung T. L., Lee C. N. *et al.* (2002b) Sequence diversity of the capsid gene and the nonstructural gene NS2B of dengue-3 virus *in vivo*. *Virology*. 303:181–191.
101. Wash, R., & Díaz S., C. (2015) True Blood: dengue virus evolution. *Nature Reviews Microbiology*. 13(662).
102. Wellcome Trust Sanger Institute (2012) SMALT Disponible en <http://www.sanger.ac.uk/science/tools/smalt-0>
103. Westaway, E. G. & Blok, J. (1997) Taxonomy and evolutionary relationships of flaviviruses. En: Gubler, D. J. & Kuno, G. (eds). *Dengue and dengue hemorrhagic fever*. Cab International. UK. Pp. 147-173.
104. Wilm, A., Kim A., P. P., Bertrand, D. *et al.* (2012) LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Research*. 40 (22): 11189–11201.
105. Zerbino, D. R. & Birney, E. (2008) Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18(5):821-9. Velvet, disponible en: <http://www.ebi.ac.uk/~zerbino/velvet/>

106. Zhao, B., Mackow, E., Buckler-White, A. *et al.* (1986) Cloning full-length dengue 4 viral DNA sequences: analysis of genes coding for structural proteins. *Virology*. 155: 77-88.
107. Zhang, M., Zheng X., Wu Y. *et al.* (2010) Quantitative analysis of replication and tropisms of dengue virus type 2 in *Aedes albopictus*. *Am. J. Trop. Med. Hyg.* 83(3): 700-707.

## 12. Anexos

### 12.1 Anexo 1

Virus representados en la base de datos de *Flaviviridae*.

- |                                  |  |                                      |
|----------------------------------|--|--------------------------------------|
| 1. Aedes flavivirus              | 29. Flavivirus White Kaiya duck                  | 52. Mosquito flavivirus              |
| 2. Alfuy virus                   | 30. Hepatitis C virus                            | 53. Murray Valley encephalitis virus |
| 3. Alkhurma virus                | 31. Hepatitis G virus                            | 54. Naranjal virus                   |
| 4. Apoi virus                    | 32. Hepatitis GB virus A                         | 55. Negishi virus                    |
| 5. Aroa virus                    | 33. Hepatitis GB virus B                         | 56. Nhumirim virus                   |
| 6. Bagaza virus                  | 34. Hog cholera virus                            | 57. Nienokoue virus                  |
| 7. Bainyik virus                 | 35. Human pegivirus genotype 2                   | 58. Non-primate hepacivirus NZP1     |
| 8. Batu Cave virus               | 36. Iguape virus                                 | 59. Ntaya virus                      |
| 9. Border disease virus          | 37. Ilheus virus                                 | 60. Omsk hemorrhagic fever virus     |
| 10. Bovine viral diarrhea virus  | 38. Israel turkey meningoencephalomyelitis virus | 61. PWARPT Powassan virus            |
| 11. Bussuquara virus             | 39. Japanese encephalitis virus                  | 62. Paraiso Escondido virus e        |
| 12. Cacipacore virus             | 40. Jutiapa virus                                | 63. Parramatta River virus           |
| 13. Cell fusing agent virus      | 41. Kamiti River virus                           | 64. Pestivirus Burdur                |
| 14. Chaoyang virus               | 42. Karshi virus                                 | 65. Phnom Penh bat virus             |
| 15. Classical swine fever virus  | 43. Kedougou virus                               | 66. Porcine pestivirus               |
| 16. Culex flavivirus             | 44. Kokobera virus                               | 67. Powassan virus                   |
| 17. Culiseta flavivirus          | 45. Kunjin virus                                 | 68. Pronghorn antelope pestivirus    |
| 18. Deer tick virus              | 46. Kyasanur forest disease virus                | 69. Quang Binh virus                 |
| 19. Dengue Virus 1               | 47. Langat virus                                 | 70. Recombinant Hepatitis C Virus    |
| 20. Dengue Virus 2               | 48. Louping ill virus                            | 71. Rio Bravo virus                  |
| 21. Dengue Virus 3               | 49. Mercadeo virus                               | 72. Rocio virus                      |
| 22. Dengue Virus 4               | 50. Modoc virus                                  | 73. Rodent pegivirus                 |
| 23. Donggang virus               | 51. Montana myotis leukoencephalitis virus       |                                      |
| 24. Duck Tembusu virus           |  |                                      |
| 25. Duck egg-drop syndrome virus |  |                                      |
| 26. Duck flavivirus              |  |                                      |
| 27. Entebbe bat virus            |  |                                      |
| 28. Equine Pegivirus             |  |                                      |

- 74. Sepik virus
- 75. Simian pegivirus
- 76. Sitiawan virus
- 77. Sokoluk virus
- 78. Spanish goat encephalitis virus
- 79. Spondweni virus
- 80. St. Louis encephalitis virus
- 81. Stratford virus
- 82. Tamana bat virus
- 83. Theiler's disease-associated virus
- 84. Tick-borne encephalitis virus
- 85. Torres virus
- 86. Tyuleniy virus
- 87. Usutu virus
- 88. Vaccinia virus
- 89. Wesselsbron virus
- 90. West Nile virus
- 91. Xishuangbanna aedes flavivirus
- 92. Yellow fever virus
- 93. Yokose virus
- 94. Zika virus

## 12.2 Anexo 2

Detalle de los ensamblajes generados con distintos k-meros (Velvet).

### MYF1

Ensamble	k-mero	Nodos	N50	max	total	Reads usados
2	9	1317	15	66	8768	3403/3464
3	11	92	287	1042	9780	3364/3464
1	13	10	6379	6379	10632	3414/3464
4	15	5	6417	6417	10531	3390/3464
<b>5</b>	<b>17</b>	<b>4</b>	<b>6613</b>	<b>6613</b>	<b>10697</b>	<b>3411/3464</b>
6	19	4	6507	6507	10535	3379/3464

Utilizando los nodos del k-mero 17 como referencia para el *Binning*.

k-mero	Nodos	N50	max	total	Reads usados
15	3	6467	6467	10604	3758/3786
<b>17</b>	<b>2</b>	<b>10648</b>	<b>10648</b>	<b>10743</b>	<b>3780/3786</b>
19	2	10646	10646	10741	3777/3786
21	2	10644	10644	10743	3776/3786

### MYF2

k-mero	Nodos	N50	max	total	Reads usados
15	20	1845	2831	10484	1685/1776
17	20	2796	2829	10668	1683/1776
19	13	1584	2827	10668	1620/1776
21	15	1499	2825	10632	1621/1776

Utilizando los nodos del k-mero 19 como referencia para el *Binning*.

k-mero	Nodos	N50	max	total	Reads usados
15	9	2831	3044	10273	1728/1758
17	8	2829	3042	10537	1742/1758
19	9	1025	2827	10106	1799/1758

Utilizando el genoma originado de mapear contra la referencia (MYF2), como referencia para el *binning*.

k-mero	Nodos	N50	max	total	Reads usados
17	3	3745	3781	10355	<b>1731/1736</b>

### MYF3

k-mero	Nodos	N50	max	total	Reads usados
13	306	28	108	2696	363/1026
15	238	35	135	2663	341/1026
17	186	47	133	2303	302/1026
19	43	57	131	1270	<b>539/1026</b>
21	100	55	129	2001	368/1026

### MYF4

k-mero	Nodos	N50	max	total	Reads usados
13	280	69	215	4689	376/1176
15	221	64	298	4581	352/1176
17	186	63	296	4331	348/1176
<b>19</b>	<b>57</b>	<b>110</b>	<b>294</b>	<b>2251</b>	<b>278/1176</b>

### MYF5

k-mero	Nodos	N50	max	total	Reads usados
13	4484	10	88	25727	63083/91126
15	2939	15	91	24846	75435/91126
17	34	2078	5414	11107	89977/91126
19	14	3436	3848	10736	89603/91126
21	17	6312	6312	10865	90093/91126

Utilizando los nodos del k-mero 19 como referencia para el *Binning*.

k-mero	Nodos	N50	max	total	Reads usados
13	5055	9	49	27651	62049/101122
15	458	129	681	12359	99999/101122
17	2521	17	143	24171	80560/101122

Utilizando el genoma originado de mapear contra la referencia (MYF5), como referencia para el *Binning*.

k-mero	Nodos	N50	max	total	Reads usados
17	2516	17	143	23668	81237/102778
19	302	158	1212	11920	102111/102778
21	256	221	1241	11829	101472/102778
23	2	9676	9676	9800	95412/102778
<b>25</b>	<b>4</b>	<b>9840</b>	<b>9840</b>	<b>10332</b>	<b>100756/102778</b>

### 12.3 Anexo 3

Secuencias empleadas para la genotipificación, obtenidas de ViPR (todas corresponden a genomas completos)

Genotipo	GenBank	Año	País	Hospedero
Asiático-Americano	JX669480	1995	Humano	Brasil
	JX669476	2010	Humano	Brasil
	GU369819	2002	Humano	México
	AY702034	1997	N/A	Cuba
	AF208496	N/A	Humano	Martinica
	GQ868555	2205	Humano	Colombia
	HQ999999	2009	Humano	Guatemala
	GQ868497	2006	Humano	México
	GQ868516	2007	Humano	México
	KJ189308	2009	Mosquito (AE)	México
	KJ189309	2010	Mosquito (AE)	México
	EU482620	2006	Humano	Nicaragua
	KC294203	2010	Humano	Perú
Asiático 2	AF204178	1987	Humano	China
	KF704357	1981	Humano	Cuba
	JF730050	2007	Humano	USA
	JF730055	2009	<i>Mus musculus</i>	USA
Americano	HM582106	1972	N/A	Samoa americana
	HM582101	1971	N/A	Fiji
	HM582109	1971	N/A	Polinesia francesa
	HM582111	1974	N/A	Tonga
	GQ868588	1983	Humano	México
	GQ868590	1992	Humano	México
Asiático 1	FM210214	2004	Humano	Vietnam
	FJ639831	2001	Humano	Thailandia
Selvático	EF105380	1980	N/A	Costa de marfil
	EF105386	1980	N/A	Burkina Faso
Cosmopolita	KF041233	2011	Humano	Pakistan
	KF360005	2010	Humano	Pakistan
	AF276619	N/A	N/A	China
	DQ448231	2001	Humano	India

## **13. Resumen del trabajo presentado en el IX Congreso Nacional de Virología, septiembre de 2015**

**IX Congreso Nacional de Virología**

**Morelos, México del 22 al 25 de septiembre de 2015**

### **FILOGEOGRAFÍA DE DENV-1 Y DENV-2 EN EL CONTINENTE AMERICANO.**

**Hernández-Pérez, F. & Zárata S.**

Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México (UACM). San Lorenzo #290, Col. Del Valle. CP 03100, Distrito Federal. México. Tel. 54886661 ext 15318. [prenaufrago@hotmail.com](mailto:prenaufrago@hotmail.com)

Existen cuatro serotipos del virus de dengue (DENV-1 a DENV-4), cada uno de los cuales se divide en subtipos (genotipos) distintos. En el continente americano, se ha observado que estacional y regionalmente, el DENV muestra predominio de uno de los serotipos, y que a escalas epidemiológicas pequeñas, el movimiento del DENV entre regiones deriva en el establecimiento de ciertos linajes dentro de las poblaciones, acompañado de la extinción de los genotipos de circulación local.

A escala global se ha observado que, como ocurre en general con los virus de RNA, las poblaciones de algunos genotipos de DENV muestran una subdivisión espacial fuerte. Diferentes localidades geográficas se caracterizan por la presencia de subtipos virales filogenéticamente distintos, aunque algunos genotipos muestran una distribución cosmopolita. En general, esto se ha observado a partir de filogenias construidas con regiones específicas del genoma (particularmente la proteína E).

Con la intención de determinar si en el continente americano los distintos serotipos del DENV

presentan una estructura poblacional determinada por la localidad, se realizaron pruebas de genética de poblaciones desde una perspectiva filogeográfica, comparando árboles filogenéticos, contruidos a partir de las secuencias del genoma completo de DENV, con su distribución geográfica.

Se realizó un análisis con más de 300 secuencias de DENV-1 y DENV-2 obtenidas de la base de datos pública Virus Pathogen Resource. Los serotipos DENV-1 y DENV-2 fueron elegidos por ser los que presentan un mayor número de secuencias en la base de datos para la región estudiada (continente americano). Para cada serotipo, se descartaron aquellas secuencias que fueran registro único para su país de procedencia; de las secuencias restantes se recortaron las regiones 5'UTR y 3'UTR y se realizó un alineamiento únicamente con la región codificante. Se construyó un árbol filogenético por el método de Neighbor joining y empleando el software HyPhy se realizó una prueba de Slatkin y Maddison.

La prueba de Slatkin y Maddison permitió corroborar que las secuencias de DENV-1 y DENV-2 se agrupan según su geografía, lo que refleja el carácter local de los brotes. Este resultado indica que, independientemente de la fuerza evolutiva que module predominantemente la evolución del DENV, existen firmas específicas en el genoma que se asocian a su ubicación geográfica. Se realizó la búsqueda de aquellas regiones o motivos comunes a las secuencias características de una región geográfica determinada con la intención de encontrar indicadores de la procedencia de un genotipo a partir de la secuencia de un aislado viral.

El jurado designado por la Universidad Autónoma de la Ciudad de México aprobó esta tesis el día 24 de enero del 2017, en la Ciudad de México, para optar al Grado de Maestra en Ciencias Genómicas, a la Biól. Cindy Fabiola Hernández Pérez.

Dra. Claudia Selene Zárate Guerra

---

Dra. Lilia López Cánovas

---

Dra. Rosa Martha Eugenia  
Yocupicio Monroy

---

Dra. Sofía Lizeth Alcaraz Estrada

---

Dra. Blanca Itzelt Taboada Ramírez

---

