



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

MAESTRIA EN CIENCIAS DE LA PRODUCCIÓN Y
DE LA SALUD ANIMAL

EPIDEMIOLOGÍA MOLECULAR DEL VIRUS DE
ESTOMATITIS VESICULAR SEROTIPO NEW JERSEY EN
MÉXICO PERIODO 2005-2009

T E S I S

PARA OBTENER EL GRADO DE MAESTRO EN CIENCIAS

P R E S E N T A

LAURO VELÁZQUEZ SALINAS

TUTOR:

ANTONIO VERDUGO RODRÍGUEZ

COMITÉ TUTORAL:

FRANCISCO JAVIER BASURTO ALCÁNTARA

ALFONSO CLAVIJO

MÉXICO D.F

2011



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Este trabajo está dedicado a todas las personas que en diferentes ámbitos están luchando por ser escuchados y que a pesar de todas las adversidades, trabajan día con día para hacer de este nuestro México un país mejor.

A mis padres Lauro Velázquez Dávila y María del Socorro Salinas Iñiguez, ustedes son mi fuerza e inspiración para seguir adelante en esta aventura llamada vida.

A mi hermano Ramiro Velázquez Salinas, admiro lo que haces hermano. A mis hermanos Fernando Velázquez Salinas y Luis Velázquez Salinas.

Así mismo quiero dedicar este escrito a 3 científicos que admiro mucho. El conocerlos cambio mi perspectiva de la vida y me hizo tomar la decisión de realizar este programa de maestría:

-Dr. Alfonso Clavijo. (Texas A & M University). Gracias por todo tu apoyo y consejos para tomar la decisión de dar este paso en mi vida, por todas tus enseñanzas durante mis 3 meses de estancia en tu laboratorio de enfermedades vesiculares en Winnipeg Canadá, admiro en verdad toda tu trayectoria profesional y una de mis metas después de conocerte es la de llegar tan lejos como tú.

-Dr. José López (Canadian Food Inspection Agency). Mi estimado Brother, Gracias por tu amistad y todos tus consejos para seguir adelante. Te admiró por ser mexicano y por llegar hasta donde has llegado, a base de esfuerzo y honestidad. Conocerte en Canadá en el 2005 cambio mi vida y pude darme cuenta de que es lo que tenía que hacer en esta vida para superarme.

-Dr Luis Rodríguez (Agricultural Research Service USDA). Gracias por todo su apoyo y por creer en mí. Admiro profundamente su trayectoria profesional en el área de estomatitis vesicular. He tenido la fortuna de leer todos sus trabajos en esta materia y gran parte de ellos están referenciados en este escrito. Mi trabajo de tesis está inspirado en sus investigaciones en la epidemiología molecular de este virus. Un privilegio conocerlo en persona en febrero del 2011 durante mi estancia en Plum Island Animal Disease Centre

Agradecimientos:

-Dr Igor Romero Sosa (Director de la CPA) y al Dr Pedro Paz Ramírez por permitirme hacer realidad este proyecto y por las facilidades de financiamiento y tiempo para realizar el mismo.

- Mi Comité tutorial, por apoyarme para la realización de este trabajo. Muy en especial al Dr Antonio Verdugo Rodríguez mi tutor principal por elegir el camino más conveniente para mi formación académica durante la maestría. Agradezco sus consejos para ser cada día una mejor persona y no ponerme límites para conseguir mis metas.

-M en C Belem de Jesús Huerta Lozano (Laboratorio de la CPA), por enseñarme a realizar la técnica de secuenciación utilizada en este trabajo, así como en el diseño de los iniciadores.

-MVZ Elizabeth Ramírez Medina, por todo su apoyo en la realización de las pruebas de RT-PCR, así como en la recolección de las muestras epiteliales utilizadas en este trabajo.

-MVZ Francisco Liljehut Fuentes, por su amistad incondicional, por todo su apoyo y porque juntos iniciamos con el área de enfermedades vesiculares en el laboratorio de la CPA. Gracias amigo.

-Dra Selene Zarate Guerra (UACM), por todo su apoyo para la corrección de este manuscrito, por enseñarme a hacer los análisis filogenéticos, pruebas de selección y el análisis de escalamiento multidimensional utilizados en este trabajo. Mi admiración para ti por tu excelente en tu trayectoria profesional. Gracias por compartir conmigo un poco de lo mucho que sabes.

-Dr Francisco Lobo (EMBRAPA), Por todos sus consejos sobre la metodología a seguir en este trabajo. Por enseñarme los análisis de uso de codones, dinucleotidos, y ENC plot. Te agradezco mucho eres un excelente maestro y gracias aunque tu estas en Brasil por todo el tiempo que me dedicaste para corregir mis resultados y enseñarme un poco de lo mucho que sabes.

Resumen.

El análisis filogenéticos del virus de estomatitis vesicular (VEV) ha mostrado una gran variabilidad genética, sin embargo linajes genéticos asociados a zonas geográficas específicas han permanecido genéticamente estables hasta por 30 años. Así mismo se ha propuesto la hipótesis, de que factores ecológicos más que temporales dominan la evolución del VEV. El objetivo de este trabajo fue determinar los patrones filogenéticos y evolutivos del virus de estomatitis vesicular New Jersey, analizando por medio de herramientas bioinformáticas la región hipervariable del gen P de 140 aislamientos del virus de estomatitis vesicular New Jersey (VEVNJ), representativos de México en el periodo 2005-2009, con la finalidad de describir la epidemiología molecular de este virus en México. Los resultados muestran que existe selección positiva de mutaciones sinónimas, las cuales están influenciando el uso de codones entre los diferentes linajes establecidos en América. El uso de codones en este virus parece estar influenciado tanto por la composición de nucleótidos de la secuencia así como presiones selectivas que favorecen la optimización de la traducción de los linajes establecidos en diferentes nichos ecológicos. Con base en estos hallazgos, describimos 2 zonas endémicas en México en las cuales circulan linajes con características de uso de codones similares, posiblemente como respuesta a circular en zonas ecológicas diferentes. Un hallazgo significativo en este periodo de estudio, fue la aparición de un linaje que se diseminó del sur al norte de México del 2006 al 2009 y lo que resultó en un aumento de 400% de casos positivos en México, basado en su característica de uso de codones, fue posible predecir el origen geográfico del mismo.

Palabras claves: uso de codones; selección negativa; mutaciones sinónimas, dinucleótidos; análisis filogenéticos; evolución.

Abstract.

Phylogenetic analysis of vesicular stomatitis virus (VSV) have shown the presence of a large genetic variability, with genetic lineages associated to specific geographic zones where they have remained genetically stable for 30 years. There is the hypothesis that ecological factors rather than temporal ones are dominating the evolution of VSV. The objective of this research was to determinate the phylogenetic and evolutionary patterns of vesicular stomatitis virus New Jersey serotype VSV NJ by analyzing the hypervariable region of the phosphoprotein gen of 140 VSV NJ isolates representative of Mexico from 2005-2009, in order to describe the molecular epidemiology of VSV NJ in Mexico. The sequences showed evidence of positive selection of synonymous mutations, which results in codon usage bias among different lineages in America. In addition codon usage bias in this virus seems to be influenced by both nucleotide composition and selective pressures maybe to improve the translation of different lineages through different ecologic zones where they are established. Based on these findings, we were able to describe the presence of two endemic zones in México, where lineages with similar codon usage patterns are circulating, possibly as a result of their adaptation to different ecological zones. From 2006 to 2009 a lineage circulating in Mexico, was responsible for increasing the incidence of this disease in 400%, using the specific of codon usage bias, we could detect the geographical origin of this lineage.

Key words: codon usage bias; negative selection; synonymous mutations; dinucleotide, phylogenetic analysis; evolution.

Contenido:

Resumen.....	III
Abstract.....	IV
1. Introducción	
1.1. Generalidades.....	1
1.2. Aspectos epidemiológicos y filogenéticos	
1.2.1 Distribución geográfica.....	1
1.2.2 Ciclo natural de la infección.....	4
1.2.3 Evolución natural y filogenia.....	6
2. Justificación.....	12
3. Hipótesis.....	13
4. Objetivos	
4.1 Objetivo general.....	14
4.2 Objetivos específicos.....	14
5. Material y Métodos.	
5.1 Diseño de iniciadores, extracción de RNA viral y RT-PCR.....	15
5.2 Purificación de los productos de amplificación y secuenciación.....	16
5.3 Análisis filogenético.....	16
5.4 Análisis evolutivo.....	16
5.4.1 Detección de la selección.....	17
5.4.2 Uso relativo de codones sinónimos (RSCU).....	17
5.4.3 Análisis de escalamiento multidimensional.....	18
5.4.4 Calculo de la frecuencia de dinucleótidos.....	18
5.4.5 Numero efectivo de codones (ENC).....	19
5.5 Epidemiología molecular del VEV NJ en México (20052009).....	19

6. Resultados	
6.1 Análisis filogenético.....	20
6.2 Análisis Evolutivo.	
6.2.1 Detección de la selección.....	24
6.2.2 Diferencias en el uso de codones.....	26
6.2.3 Frecuencias de dinucleótidos.....	29
6.2.4 Efecto del patrón de dinucleótidos en el uso de codones.....	29
6.2.5 Numero efectivo de codones.....	33
6.3 Epidemiología molecular del VEV NJ en México.....	35
7. Discusión.	
7.1 Detección de la selección.....	39
7.2 Uso de codones.....	40
7.3 Frecuencia de dinucleótidos.....	42
7.4 Epidemiología molecular del VEV NJ en México.....	44
8. Conclusiones.....	46
9. Perspectivas.....	47
10. Bibliografía.....	48
11. Anexos	
1 Muestras.....	56
2 Información complementaria de la figura 5.....	60
3 Zonas ecológicas de México.....	62
4 Definiciones.....	63

Lista de figuras.

Figura 1. (A) Distribución geográfica y temporal del virus de estomatitis vesicular serotipos New Jersey e Indiana1. (B) Distribución geográfica y temporal de los serotipos de Indiana 2 e Indiana 3.	3
Figura 2. Ciclo de infección natural en el virus de estomatitis vesicular, propuesto en base a la información científica actual.....	5
Figura 3. Distribución filogenética del VEV NJ en América.....	8
Figura 4. Relación entre filogenia y condiciones ecológicas en el VEV NJ.....	11
Figura 5. Representación filogenética del virus de estomatitis vesicular en México entre los años 2005 a 2009.....	22
Figura 6. Relación filogenética, del los linajes encontrados en México (I, II, III, IV, V, VI) en el periodo 2005-2009 con otros previamente publicados en el Genbank.....	23
Figura 7. Comparación de los niveles de selección positiva, neutral y Negativa en el VEV NJ.....	25
Figura 8. Diferencias en la preferencia por el uso de codones entre los linajes de Norteamérica, Centroamérica y Suramérica.....	27
Figura 9. Relación entre uso de codones y adaptación a diferentes zonas ecológicas en el VEV NJ.....	29
Figura 10. Comparación de las frecuencias de dinucleótidos entre los diferentes linajes del VEV NJ estudiados en este trabajo.....	31
Figura 11. Lista ordenada de codones en los diferentes linajes del VEV NJ estudiados en este trabajo en relación a sus valores RSCU.....	32

Figura 12(A). Representación grafica de la distribución geográfica del VEV NJ en México.....**37**

Figura 12(B). Representación grafica de la distribución geográfica del VEV NJ en México.....**38**

Abreviaturas.

Abreviatura	Definición
ApA, ApC, ApG, ApT, CpA, CpC, CpG, CpT, GpA, GpC, GpT, GpG, GpA, GpC, GpT, GpG.	Regiones de la secuencia donde se encuentra un par de adenina (A), citosina (C), guanina (G), timina (T) respectivamente, ligados por un fosfato el cual está representado por la letra p.
CA I	Virus de estomatitis vesicular serotipo New Jersey reportados en el Genbank, los cuales circulan en Centroamérica en los países de Guatemala, Belice y el Salvador.
CA II	Virus de estomatitis vesicular serotipo New Jersey reportados en el Genbank, los cuales circulan en Centroamérica y Suramérica en los países de Honduras, Nicaragua, Panamá, Costa Rica y Ecuador.
dN	Tasa de substitución no sinónima por sitio no sinónimo.
dS	Tasa de substitución sinónima por sitio sinónimo.
ENC	Numero Efectivo de Codones.
RSCU	Uso Relativo de Codones Sinonimos
RT-PCR	Transcriptasa Reversa-Reacción en Cadena de la Polimerasa.
SLAC	Single Likelihood Ancestor Counting.
VEV	Virus de Estomatitis Vesicular.
VEV NJ	Virus de Estomatitis Vesicular serotipo New Jersey.

1 .Introducción.

1.1 Generalidades.

La estomatitis vesicular (EV) es una enfermedad infecciosa, la cual afecta bovinos, porcinos y equinos. Es causada por el virus de la estomatitis vesicular (VEV) el cual pertenece a la familia *Rabdoviridae* y género *Vesiculovirus*, tiene un genoma del tipo RNA de cadena sencilla con una polaridad negativa y una longitud de 11 kb, el cual codifica para la producción de 5 proteínas estructurales: la nucleocapside (N), fosfoproteína (P), matrix (M), glicoproteína (G), y polimerasa (L)(1).

La presentación clínica de la enfermedad se caracteriza por la presencia de lesiones vesiculares y erosivas en la superficie dorsal de la lengua, frecuentemente acompañadas por vesículas en las encías, labios, bandas coronarias, pezones, y prepucio. El impacto económico de esta enfermedad ha sido calculado entre los 100 y 250 dólares por bovino infectado (2). Así mismo, las lesiones producidas en bovinos y cerdos por el VEV no pueden ser distinguidas clínicamente de las producidas por el virus de la fiebre aftosa, una de las enfermedades de mayor impacto económico a nivel mundial, por lo que los hatos afectados por la EV tienen que ser puestos en cuarentenas precautorias hasta poder discriminar la presencia del virus de la fiebre aftosa por medio de pruebas de laboratorio, lo que se traduce en pérdidas económicas por la restricción en la movilización de los animales afectados (3).

1.2 Aspectos Epidemiológicos y Filogenéticos.

1.2.1 Distribución Geográfica.

Por lo menos 28 *vesiculovirus*, han sido identificados serológicamente, tienen la capacidad para infectar vertebrados e invertebrados (4), de estos serotipos los implicados en causar infecciones en animales domésticos son New Jersey (VEV NJ), Indiana (VEV IND-1), Cocal (VEV IND-2), Alagoas (VEV IND-3) Chandipura, Piry y Isfahan (3).

El VEV es endémico del norte de Sudamérica (Colombia, Venezuela, Ecuador, Perú) al sur de México, en estas áreas los brotes clínicos de esta enfermedad se presentan anualmente (Figura 1 A) (5).

En las áreas endémicas, los brotes de EV tienen un patrón de presentación estacional, la mayoría de las veces están asociados a los periodos de transición entre la época de lluvias y la de secas (6). El virus de VEV NJ produce aproximadamente el 80% de los casos clínicos reportados, y el 20% restante es producido por el VEV IND 1 (5). Brasil y Argentina reportan esporádicamente casos de EV, en estos países los serotipos relacionados con la presentación de brotes clínicos son del tipo de Indiana 2 y 3, con una frecuencia de presentación entre 1 y 2 años en Brasil por el serotipo de Indiana 3 y esporádicamente por el serotipo Indiana 2 en ambos países (figura 1B) (7). En el sureste de los Estados Unidos, los brotes causados por el VEV, ocurren esporádicamente en ciclos de aproximadamente 10 años (8).

En México, la presentación de esta enfermedad varía dependiendo de la localización geográfica, en el caso de la parte sur (Veracruz, Chiapas, Tabasco) los brotes clínicos producidos por el VEV se presentan de manera anual, a diferencia de los estados de Jalisco, Guerrero, Morelos y Michoacán donde la detección de casos positivos puede ser por periodos de 5 años consecutivos seguidos de periodos sin actividad de 5 o 6 años aproximadamente. En la parte norte (Sonora, Sinaloa, Chihuahua, Nuevo León) los periodos de actividad pueden ser de 1 o 2 años y no haber evidencia de casos clínicos hasta en periodos de 10 años (8).

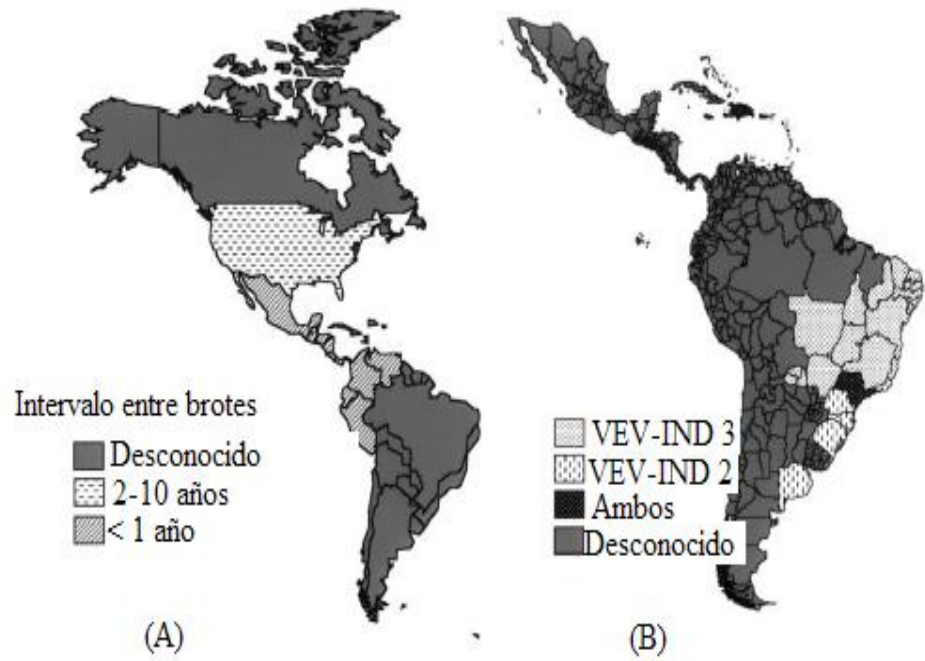


Figura 1. (A) Distribución geográfica y temporal del virus de estomatitis vesicular serotipos New Jersey e Indiana1. (B) Distribución geográfica y temporal de los serotipos de Indiana 2 e Indiana 3.

La información de esta figura fue obtenida del Centro Panamericano de Fiebre aftosa (www.panaftosa.org) y la Organización Mundial de Salud Animal (OIE) (www.oie.int). La figura fue obtenida de Rodriguez L.L.(2002). Emergence and re-emergence of vesicular stomatitis in the United States. *Virus Research* **85**, 211-219.

1.3 Ciclo Natural de Infección.

El ciclo natural de infección del VEV, actualmente es desconocido. Sin embargo existe evidencia de que los insectos son portadores del virus en las zonas endémicas y son capaces de infectar animales domésticos, de laboratorio y presumiblemente infectar humanos (figura 2) (9).

Uno de los primeros reportes documentando la presencia del VEV IND fue encontrado en el díptero del género *Psychodidae*, capturado en las áreas boscosas de Panamá (10). Otros reportes incluyen aislamientos del VEV de otros grupos de dípteros hematófagos del género *Simulidae* (mosca negra), y mosquitos de los géneros *D* *Culicidae* , *Ceratopogonidae* (11). En adición a estos insectos hematófagos, el VEV ha sido encontrado en insectos no hematófagos, como las mosca doméstica (*Musca doméstica*) y *eye gnats* (*Hippelates* sp.), particularmente durante brotes clínicos de la enfermedad en Colorado Estados Unidos en 1982(12).

Experimentalmente, solo en *Lutzomyia shannon* y *Simulium vittatum* (mosca negra), se ha demostrado la capacidad de estos vectores para transmitir el virus a los hospedadores susceptibles y a ellos mismos a través de la ruta transovárica (13).

Resultados de vigilancia serológica, han mostrado la presencia de anticuerpos neutralizantes contra el VEV, no solo en animales domésticos sino en muchas especies de animales salvajes como son: murciélagos, venados y monos que viven en las áreas endémicas de México, Centroamérica y Suramérica (14).

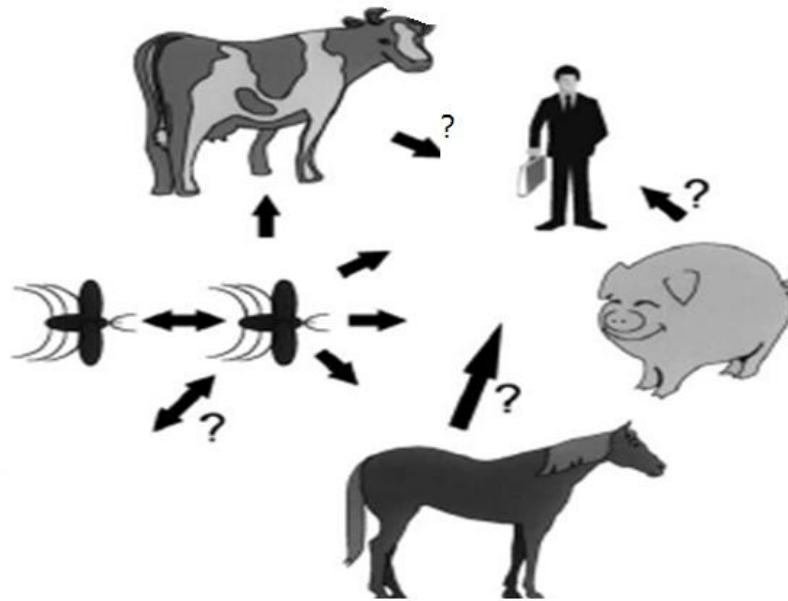


Figura 2. Ciclo de infección natural en el virus de estomatitis vesicular, propuesto en base a la información científica actual. Las flechas indican la dirección de transmisión del virus. Los signos de interrogación, se muestran posibles rutas de transmisión, no documentadas en la literatura. Los animales domésticos parecen ser el hospedador final en el cual el virus no persiste y no retorna al ciclo natural de infección. Figura tomada de: Rodriguez L.L.(2002). Emergence and re-emergence of vesicular stomatitis in the United States. *Virus Research* **85**, 211-219.

Una parte intrínseca en el ciclo de vida de los arbovirus, es la existencia de uno o varios reservorios capaces de producir niveles de viremia sostenidos, así los vectores que se alimentan de la sangre de estos animales son infectados y capaces de transmitir al virus lo que permite completar el ciclo de vida de virus como el dengue, la fiebre amarilla y el virus de encefalitis equina entre otros (15). A pesar de varios intentos para encontrar un reservorio infectado con el virus de EV que sea capaz de mantener niveles constantes de viremia, éste no ha sido encontrado en animales domésticos y salvajes (16).

La inoculación experimental de bovinos y cerdos puede resultar en infecciones clínicas o sub clínicas dependiendo de la dosis y la ruta de inoculación, pero no en la producción de viremia (16). La infección experimental de ratones de laboratorio con el VEV usualmente resulta en viremia seguida de la muerte por encefalitis (17).

Recientemente se describió un mecanismo interesante de transmisión en moscas negras (*S. Vittatum*). Moscas experimentalmente infectadas con el VEV NJ, alimentándose junto con moscas no infectadas de la sangre de un ratón no infectado con el VEV, lograron transmitir horizontalmente el VEV NJ a las moscas sin infección experimental previa (18). Este mecanismo provee una posible explicación, sobre cómo se lleva a cabo el mantenimiento del VEV en condiciones naturales en ausencia de un reservorio capaz de producir un estado virémico sostenido.

1.4 Evolución natural y filogenia.

El VEV ha sido usado como modelo de laboratorio para estudiar la evolución de los virus RNA, esto debido a la incapacidad que tiene para corregir los errores introducidos por su RNA polimerasa en el momento de llevar a cabo la replicación, la tasa de mutación es calculada en el orden de 1 mutación cada 10 kb (19). Por lo tanto la progenie es una mezcla de mutantes genéticamente relacionadas llamada “cuasiespecies”(20), en donde algunos individuos pueden presentar mayor adecuación que otros en condiciones de replicación diferentes (por ejemplo a través de diferentes líneas celulares o alternado el crecimiento entre líneas celulares) y esto representa una ventaja adaptativa para el virus para poder replicarse en diferentes condiciones. Este fenómeno ha sido probado experimentalmente al

replicar constantemente el VEV en células de insecto por 10 meses, resultando en una variante con una baja adecuación a crecer en células de mamífero e inclusive no siendo capaz de infectar experimentalmente ratones. Sin embargo un solo pase en células BHK 21 es capaz de cambiar la adaptabilidad en células de mamífero y la neurovirulencia en los ratones (21) En adición a lo anterior, ha sido posible obtener mutantes de VSV resistentes a la neutralización, cuando este virus es cultivado en presencia de anticuerpos monoclonales e incluso policlonales (22).

Los análisis filogenéticos del VEV en diversas regiones geográficas han mostrado que existe una gran diversidad genética (figura 3), además de la existencia de linajes discretos asociados con una región geográfica específica y ausencia de reloj molecular (figura 3). En otras palabras, se han encontrado linajes del VEV con una estabilidad genética en zonas endémicas hasta por periodos de 30 años, esto contrastando con lo que sucede en otros virus RNA como los virus de influenza y fiebre aftosa, en los cuales existe un patrón notable de evolución temporal (reloj molecular) (23)

En ambos serotipos del VSV (indiana y New Jersey), existe una marcada asociación entre el origen geográfico y su agrupamiento filogenético. Las cepas del VEV IND se separan filogenéticamente en 2 grandes grupos, uno que contiene todas las cepas de México y los Estados Unidos y otro con todas las cepas de centro América (figura 3). En el caso de las cepas del VEV NJ, las cepas que circulan en México y Estados Unidos se agrupan en el mismo clado, separadas de las cepas de origen en Centroamérica (figura 3) (9).

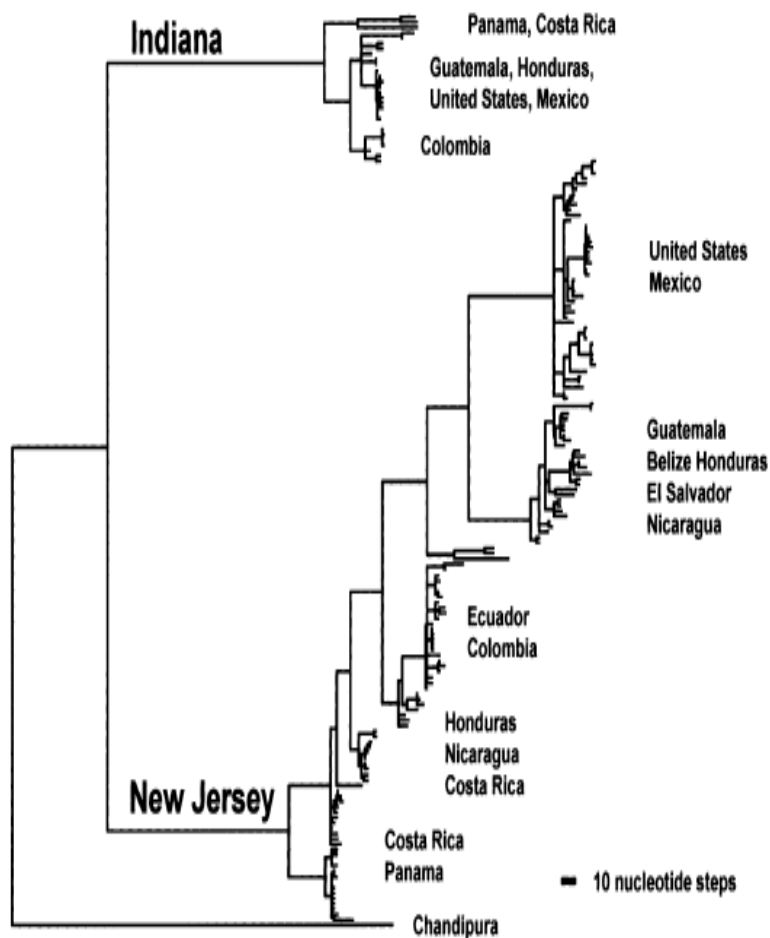


Figura 3 Distribución filogenética del VEV NJ en América. El árbol filogenético del virus de estomatitis vesicular, construido con base en secuencias parciales del gen de la fosfoproteína de 183 virus de estomatitis vesicular New Jersey e Indiana 1, representativos de Norteamérica, Centroamérica y Suramérica. El árbol fue obtenido usando el método de máxima parsimonia como es descrito en el artículo Rodríguez et al, 2000. Figura tomada de : Rodríguez L.L.(2002). Emergence and re-emergence of vesicular stomatitis in the United States. *Virus Research* **85**, 211-219.

En 1990, Pamuk y colaboradores publicaron un estudio utilizando secuencias completas de los genes N, P y G de 18 aislamientos representativos del virus de estomatitis vesicular NJ en América, como resultado los autores concluyeron que por lo menos estos 3 genes compartían la misma historia evolutiva, al obtener reconstrucciones filogenéticas similares al evaluar los genes por separados y concatenados. Este hallazgo mostro que la recombinación no juega un papel importante en la evolución de este virus, por lo menos en estos 3 genes. Sin embargo se encontraron diferencias en la tasa de fijación de mutaciones en la secuencias de aminoácidos. Las secuencias de aminoácidos de la proteína N fue la más conservada entre todos los aislamientos utilizados, indicando restricciones funcionales de esta proteína en el virus de EV. En contraste la secuencias de aminoácidos de la proteína P fueron altamente variables, explicado posiblemente por presión selectiva o deriva genética. Muchos de estos cambios fueron encontrados a nivel de la región hipervariable localizada entre los aminoácidos 153 y 205 (23). Debido a la acumulación de sustituciones en la región hipervariable del gen P, se ha propuesto que estos cambios podrían reflejar los cambios genéticos en otras partes del genoma, al obtener reconstrucciones filogenéticas similares, cuando se ha comparado esta región contra los otros genes del virus. (24,25).

Las fuerzas que afectan la evolución del VEV en la naturaleza, no han sido descritas. Estudios en las zonas endémicas de centro América y sur América, muestran que entre el 9 y 11 % del ganado de habita esas zonas presenta infecciones clínicas cada año aún cuando estos presentan una tasa elevada de anticuerpos neutralizantes contra esta enfermedad, (26,27). Esto sugeriría que las mutantes resistentes a la neutralización son las responsables de causar la enfermedad en estos animales, sin embargo estudios de poblaciones de ganado en las áreas endémicas de centro América y Sur América, no muestran evidencia de selección inmunológica, al no existir cambios de aminoácidos en los epítopes neutralizantes presentes en la glicoproteína de los virus aislados en ganado de estas zonas. (28).

En 1996 Rodríguez y colaboradores, realizaron un estudio utilizando diferentes linajes del VEV NJ los cuales fueron obtenidos de diferentes lugares endémicos en centro América. Los autores utilizaron linajes aislados en diferentes zonas ecológicas en Costa Rica, Panamá, Nicaragua y Honduras, en altitudes que varían entre 0 y 2500 metros sobre el nivel del mar. Al realizar la reconstrucción filogenética que se muestra la figura 4, se observó que los virus se agruparon en dos distintos clados (figura 4 clados A y B) relacionados con la zona ecológica donde fueron obtenidos. El análisis filogenético mostró que estos linajes han permanecido estables en estas áreas por periodos de hasta 30 años. La relación entre ellos ha sido ecológica, más que geográfica, ya que algunos de estos linajes se agruparon juntos aún cuando fueron aislados de regiones a 800 km de distancia. Esto contrasta con virus que circulan a 25 km de distancia pero en diferentes zonas ecológicas y los cuales presentan una mayor distancia genética entre ellos. Esto sugiere que las condiciones ecológicas, más que las geográficas son las responsables de la evolución del VEV en condiciones naturales (29) (figura 4).

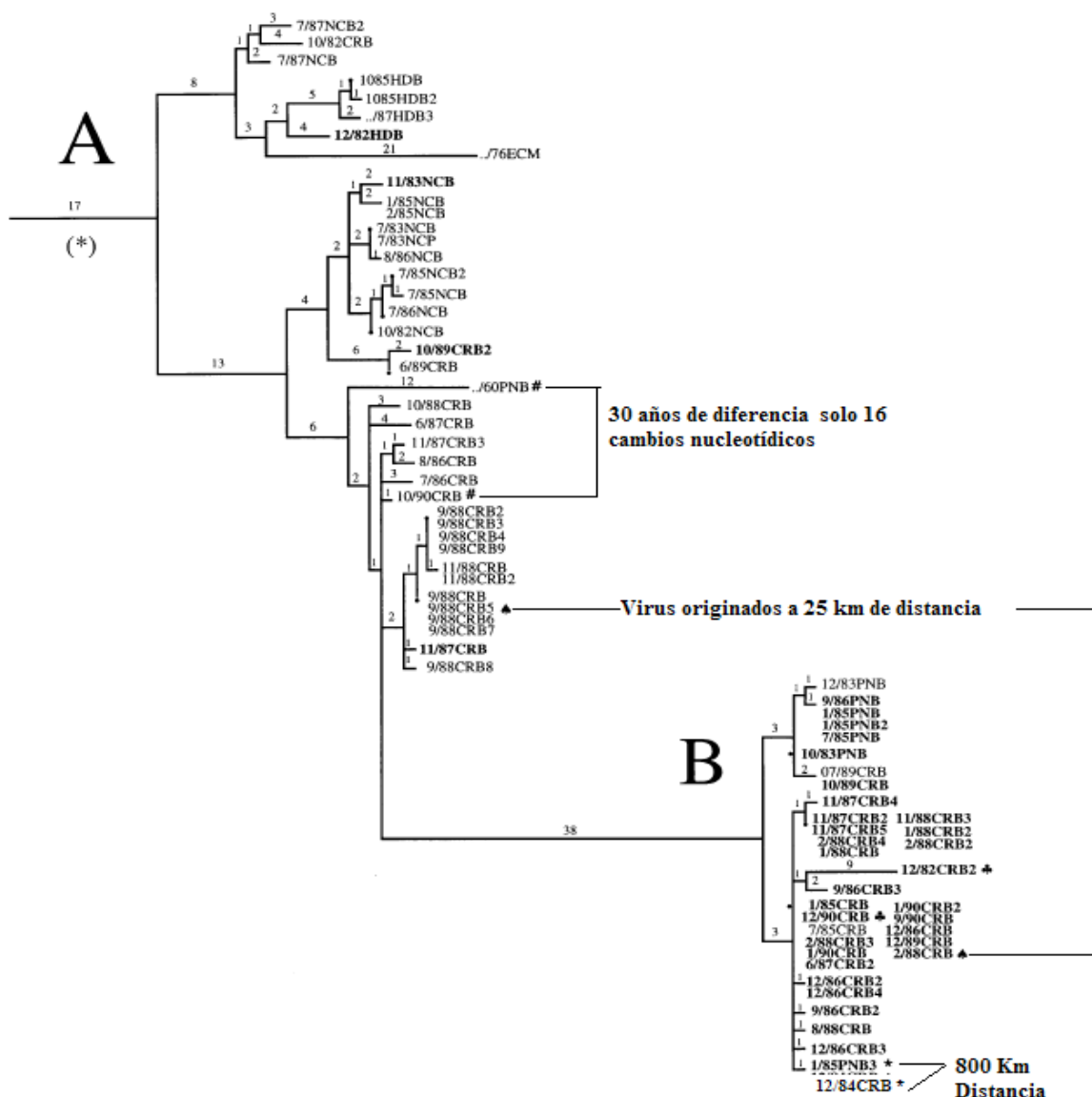


Figura 4. Relación entre filogenia y condiciones ecológicas en el VEV NJ. El árbol filogenético, construido utilizando secuencias parciales del gen de la fosfoproteína del virus de estomatitis vesicular New Jersey, de Honduras (HD), Nicaragua (NC), Costa Rica (CR), y Panamá (PN). El número de sustituciones nucleotídicas entre cada virus están indicadas sobre la líneas de las ramas del árbol. El clado A contiene los virus que se originaron en áreas <600 metros sobre el nivel del mar (Zonas Bajas). El clado B contiene los virus que se originaron de áreas > 600 metros sobre el nivel del mar (Zonas montañosas). Figura tomada de la referencia: Rodriguez L.L.(2002). Emergence and re-emergence of vesicular stomatitis in the United States. *Virus Research* **85**, 211-219.

2. Justificación.

La estomatitis vesicular es una enfermedad de presentación anual en México, la cual produce signos clínicos similares a los producidos por el virus de la fiebre aftosa, una de las enfermedades de mayor impacto económico a nivel mundial, por lo que el diagnóstico en el laboratorio es esencial para discriminar la presencia de fiebre aftosa. Diferentes estudios han sido publicados describiendo la epidemiología molecular del virus especialmente en los Estados Unidos, sin embargo en la actualidad no existe un estudio similar describiendo la epidemiología molecular de este virus en México. Así mismo muchos aspectos relacionados con la evolución en poblaciones naturales permanecen desconocidas. La mayoría de las inferencias evolutivas han sido realizadas a partir de virus crecidos en el laboratorio, por lo que los resultados generados en este trabajo serán de gran utilidad para proponer la manera en que este virus es mantenido en las zonas endémicas México y es capaz de producir brotes en las partes centro y norte del mismo.

3. Hipótesis.

3.1.-Diversos linajes del virus de estomatitis vesicular se encuentran circulando en diferentes zonas de México.

4. Objetivos.

4.1 Objetivo General:

Determinar los patrones filogenéticos y evolutivos del virus de estomatitis vesicular New Jersey, analizando por medio de herramientas bioinformáticas la región hipervariable del gen P de 140 muestras epiteliales positivas al virus de estomatitis vesicular New Jersey, las cuales fueron obtenidas entre los años 2005-2009 con la finalidad de describir la epidemiología molecular del virus en México.

4.2 Objetivos específicos

4.2.1 Estandarizar la prueba de RT-PCR, mediante el diseño de iniciadores específicos que permitan amplificar la región hipervariable del gen P del virus de estomatitis vesicular New Jersey, para obtener un producto de amplificación entre 550 y 600 pb que permita realizar el análisis de secuenciación de esta región.

4.2.2 Analizar 138 muestras positivas al virus de estomatitis vesicular New Jersey recibidas en el laboratorio de la CPA en el periodo 2005-2009, mediante el método de secuenciación automática, para obtener la secuencia de la región hipervariable del gen P de cada una de las muestras.

4.2.3 Realizar el análisis filogenético de las secuencias obtenidas, por el método de distancia de *Neighbor-joining*, para establecer la relación espacio temporal de los diferentes genotipos circulantes del virus de estomatitis vesicular New Jersey en México y su relación con otros genotipos circulantes en América previamente reportados.

4.2.4. Calcular diferentes parámetros evolutivos (presiones selectivas, uso de codones sinónimos, número efectivo de uso de codones, frecuencias de dinucleótidos), por métodos bioinformáticos, con la finalidad de encontrar evidencia sobre los mecanismos evolutivos del virus de EV y proponer una nueva hipótesis para explicar los mecanismos relacionados al mantenimiento de este virus en las zonas endémicas y su circulación fuera de ellas.

5. Material y métodos.

5.1 Diseño de iniciadores, extracción del RNA viral y RT-PCR.

Para el diseño de los iniciadores se utilizaron 18 secuencias completas del gen P del VEV NJ, las cuales fueron obtenidas del Gen Bank y los números de acceso van consecutivos del M31862.1 al M31880.1, las cuales son representativas de Norte-Centro-Suramérica. Estas secuencias fueron alineadas utilizando el programa ClustalW (30), utilizando el software Bioedit (31) y como resultado se obtuvieron los iniciadores sentido 5' AATCGTCAGATTCAGATACTGAC 3', y antisentido 5' TYYGGGCATACTGARTRATA 3, los cuales producen un fragmento de 567 pares de bases que van de las bases 170-192 (sentido) a la 718-739 (antisentido) y flaquean a la región hipervariable del VEV NJ ubicada entre las bases 467 y 625.

Para la extracción de RNA viral se utilizaron 140 suspensiones epiteliales previamente diagnosticadas como positivas al VEV NJ todas ellas por ELISA de antígeno y RT-PCR por el laboratorio de la CPA entre los años 2005 al 2009, las cuales son procedentes de diferentes regiones geográficas de México (anexo 1). Esta se realizó utilizando tiocinato de guanidina siguiendo el método del TRIzol (32). Las extracciones de RNA fueron utilizadas inmediatamente para la reacción de RT-PCR o guardadas a -70°C hasta su uso. La reacción de RT-PCR fue realizada en un solo paso utilizando el kit *QIAGEN one Step RT-PCR* número de catálogo 210210, siguiendo las instrucciones del fabricante para la elaboración de la mezcla de reacción.

(http://www.ebiotrade.com/buyf/productsf/qiagen/1020892HBRTPCR_05202.pdf)

El protocolo de amplificación incluyó 30 minutos a 48°C seguido por 10 minutos a 95°C por un ciclo, 15 segundos a 94°C, 30 segundos a 56°C, 30 segundos a 72°C estos 3 pasos se repitieron por 40 ciclos, finalmente un paso de 7 minutos a 72°C. Los productos de amplificación fueron visualizados en un gel de agarosa al 1.5% teñido con bromuro de etidio.

5.2 Purificación de los productos de amplificación y secuenciación.

Los productos de amplificación fueron purificados directamente de la reacción de RT-PCR utilizando el kit de purificación *QIAquick de Qiagen* siguiendo las instrucciones del fabricante. Los productos de PCR fueron secuenciados en reacciones independientes para los iniciadores sentido y antisentido a una concentración de 3.2 pmol , utilizando el kit de secuenciación *BigDye Terminator v3.1 Cycle sequencing kit de Applied biosystems*. Finalmente la secuencia fue obtenida usando un secuenciador automatizado, de la marca Applied Biosystems modelo 3730 A.

5.3 Análisis filogenético.

Las secuencias de nucleótidos fueron alineadas utilizando el programa ClustalW (30), utilizando el software Bioedit (31). Posteriormente los alineamientos finales fueron editados manualmente para obtener un marco de lectura abierto de 489 nucleótidos. El análisis filogenético se llevó a cabo utilizando la paquetería contenida en el software Mega 5 (33). Para la reconstrucción filogenética del árbol de la figura 5, se utilizó el método de distancia *Neighbor-Joining* (34), considerando como modelo de sustitución nucleotídica el Tamura Nei (35) el cual fue determinado por máxima verosimilitud como el de mejor ajuste de 24 modelos de sustitución nucleotídica disponibles en el programa. Los datos presentaron una distribución gama para la variación entre sitios de 0.53506. Como método de soporte estadístico se utilizaron 1000 replicas de bootstrap (36). Con la finalidad de conocer la relación filogenética de los linajes encontrados en este trabajo con otros previamente publicados de México y otras partes del mundo, se realizó una reconstrucción filogenética 80 secuencias mas procedentes del Genbank, con los siguientes números de acceso: L24994.1, L24965.1 al L25009.1, L25010.1 al L25040.1, FJ595501 al FJ595501. El procedimiento para dicha reconstrucción filogenética, fue el mismo que se indica arriba pero usando la distribución gama con un parámetro α 0.77.

5.4 Análisis evolutivo.

5.4.1 Detección de la Selección.

El análisis de detección de selección fue llevado a cabo comparando la proporción relativa de sustituciones no sinónimas (β) versus sinónimas (α), para inferir el tipo de selección que está actuando sobre cada uno de los codones de los virus considerados en el análisis. La relación $\omega = \beta/\alpha$ (también referida como dN/dS) fue interpretada como $\omega = 1$ evolución neutral, $\omega < 1$ selección negativa y $\omega > 1$ selección positiva. Este análisis fue llevado a cabo utilizando los métodos SLAC (37) y Evolutionary Fingerprint (38), disponibles en el servidor DATAMONKEY (39), y se consideraron para el análisis 163 codones del gen de la fosfoproteína del virus de EV NJ, los cuales codifican para los aminoácidos 62 al 224. Se utilizaron 140 muestras producidas en este trabajo y 18 secuencias completas del gen de la fosfoproteína del VEV NJ, las cuales son representativas de los linajes de Norteamérica, Centroamérica y Suramérica y fueron obtenidas del Genbank con los números de acceso M31862.1 al M31870.1 y M31870.1 al M31880.1. Estas fueron editadas para obtener los mismos 163 codones presentes en las muestras obtenidas en este trabajo.

5.4.2 Uso relativo de codones sinónimos (RSCU).

El valor RSCU de un codón (40) es su frecuencia observada dividida entre su frecuencia esperada en ausencia de una preferencia en el uso de codones, la cual es la frecuencia promedio de todos los codones para ese aminoácido. Los valores RSCU no están afectados por el tamaño de la secuencia ni por la frecuencia de los aminoácidos ya que estos factores son eliminados durante los cálculos. El uso de codones usados por arriba o por abajo del promedio son representados como >1 ó <1 respectivamente (41). Codones con valores de RSCU >1.0 fueron considerados como usados más frecuentemente que lo esperado, mientras que valores < 1.0 fueron considerados como subrepresentados. Los cálculos se llevaron a cabo en el software Mega 5 (33), para este fin se utilizaron las 140 secuencias obtenidas en este trabajo y denominadas como linaje I, II, III, IV, VI, 16 de las denominadas Centro América 1 (CA I) que considera virus colectados en los países Belice, El salvador, Guatemala con números de acceso al Genbank L24965.1, L24966.1, L24972.1, L24974.1, L24985.1, L24993.1, L24995.1, L25008.1, L25009.1, L25022.1, L25031.1, L25034.1, L25035.1, L25040.1, así como 41 del linaje denominado Centroamérica II que

contempla virus circulantes en Costa Rica, Honduras, Nicaragua, Ecuador y Panamá números de acceso en el Gen bank: M31862.1, M31874.1 al M31880.1, U39202-U39225.

5.4.3 Análisis de escalamiento multidimensional.

Para examinar si existían diferencias en el uso de codones entre los diferentes linajes encontrados en este trabajo se utilizó un análisis de escalamiento multidimensional, utilizando los valores RSCU calculados para cada uno de los linajes analizados en este trabajo. Este método ha sido usado en análisis de uso de codones (42) y provee una representación grafica de la mejor comparación simultánea entre grupos utilizando una matriz de datos compuesta por filas y columnas.

Este método calcula la posición de la secuencia en un espacio multidimensional, de acuerdo con el uso de codones y su representación en 2 dimensiones. Por lo tanto secuencias que tienen un uso de codones similar se colocan juntas dentro de la grafica (43). Este análisis se llevó a cabo utilizando el programa estadístico R (<http://www.r-project.org/>).

5.4.4 Calculo de la frecuencia de dinucleótidos.

Este se define como la probabilidad de encontrar un dinucleótido en una secuencia dada dividida por el producto de las probabilidades de encontrar cada nucleótido que forma el par de dinucleótidos en la misma secuencia, este es calculado utilizando la siguiente fórmula:

$$P_{xy} = f_{xy} / f_x f_y$$

Donde f_x y f_y denotan la frecuencia de mononucleótidos X y Y en una secuencia dada y f_{xy} denota la frecuencia de dinucleótidos xy en la misma secuencia. Se utilizó el criterio de que valores de dinucleótidos fuera del rango de valores 0.78-1.25 fueran considerandos como dinucleótidos con baja o alta abundancia respectivamente (45).

Para este fin se utilizaron las mismas secuencias descritas en el punto 5.4.2, excepto las secuencias de CA I, ya que se uso el linaje VI como representativo de esa zona.

5.4.5 Numero efectivo de Codones.

El número efectivo de codones (ENC), es una medida para cuantificar la preferencia de un gen para usar cierto codón en la síntesis de un aminoácido, y es una manera de probar si las substituciones sinónimas son neutrales. El ENC es definido como el número de codones que producirían el nivel observado de uso de codones si todos los codones fueran igual de frecuentes. Los valores ENC van de 20 (cuando solo se usa un codón por aminoácido) a 61 (cuando se usan todos los codones sinónimos para sintetizar un aminoácido, no existe ninguna preferencia de uso), si el valor de un gen es 35 o menos, se considera que existe una preferencia muy fuerte por el uso de ciertos codones, los genes con valores por arriba de 50, se consideran con una baja preferencia por el uso de ciertos codones (46). Estos valores fueron calculados utilizando el software CodonW <http://mobyli.pasteur.fr/cgi-bin/portal.py?form=codonw> (44). Existe una correlación entre el valor de ENC con el porcentaje de G+C en la posición número 3 de los codones de un gen (GC3%), por lo que para comprobar esto, los valores ENC y GC3% fueron graficados junto con una línea de tendencia previamente calculada utilizando la siguiente formula $NC=2 + fCG + (29 / f^2CG + (1-fCG)^2)$, la cual asume los valores esperados de ENC a causa del contenido de GC3% (46).

5.5. Epidemiología molecular del VEV NJ en México (2005-2009)

Los datos de la reconstrucción filogenética, los aspectos relacionados con la evolución de este virus, así como la base de datos del área de enfermedades vesiculares del laboratorio de la CPA, fueron usados para describir la epidemiología molecular del VEV NJ en México. La localización geográfica de cada virus fue representada gráficamente utilizando el software Google Earth (<http://www.google.com/intl/es/earth/index.html>). Para la localización geográfica de cada secuencia se utilizo únicamente los datos de Estado y municipio, al no tener en la mayoría de los casos las coordenadas del lugar. Las zonas ecológicas fueron clasificadas de acuerdo al Centro de Ecología de la UNAM (50)

6. Resultados.

6.1 Análisis filogenético.

El árbol filogenético construido utilizando la región hipervariable del gen P del VEV NJ muestra la existencia de 2 nodos ancestrales de donde emergen por lo menos 6 diferentes linajes los cuales fueron determinados en relación a los ancestros comunes que compartían entre ellos (figura 5 y anexo 2).

El 87% de los virus secuenciados se alojaron dentro del linaje I, es cual forma un grupo monofilético con una distribución temporal entre el 2006 al 2009 y una distribución geográfica primordialmente en la parte centro norte de México, diferente a los linajes II, III, IV, V y VI, los cuales se mantuvieron circulando básicamente en la zona sur de México. (anexo 2).

La reconstrucción filogenética utilizando estos linajes junto con otras secuencias de virus previamente reportados, mostró básicamente 3 grandes clados en los que se agrupan los diferentes linajes del VEV NJ circulantes en Norteamérica, Centroamérica y Sudamérica.

Los linajes I, II, III y IV se agruparon en el clado de Norteamérica, mientras que los linajes V y VI que se agrupan en el clado de Centroamérica junto con virus que circulan en Guatemala, Belice y Honduras (figura 6). El hecho de encontrar los linajes V y VI circulando en México, es un hallazgo importante en este trabajo, ya que este hecho nunca había sido descrito en trabajos anteriores, probando la hipótesis planteada en este trabajo que supone que diversos linajes del virus de estomatitis vesicular se encuentran circulando en diferentes zonas de México

La prueba por el método de máxima verosimilitud para determinar la presencia de reloj molecular en las topologías obtenidas en los arboles filogenéticos de las figuras 5 y 6, mostró que las secuencias evolucionan bajo ausencia de reloj molecular. Esto es evidente al analizar la reconstrucción filogenética en la figura 5, donde la secuencias 116/Chiapas/06 se coloca muy cercana de la secuencia nj1096cpb2 ambas circulando en Chiapas con 10 años de diferencia indicativo de una estabilidad genética. El mismo caso aparece 19/Chiapas/07 con las secuencias de Chiapas nj1296cpb y nj0789cpb4 con esta última una diferencia de 18 años. En el caso del linaje VI que se encuentra actualmente circulando en

México este es similar a la cepa nj0287GM circulante en Guatemala desde 1987 la cual se mantiene genéticamente estable y está circulando en México 20 años después.

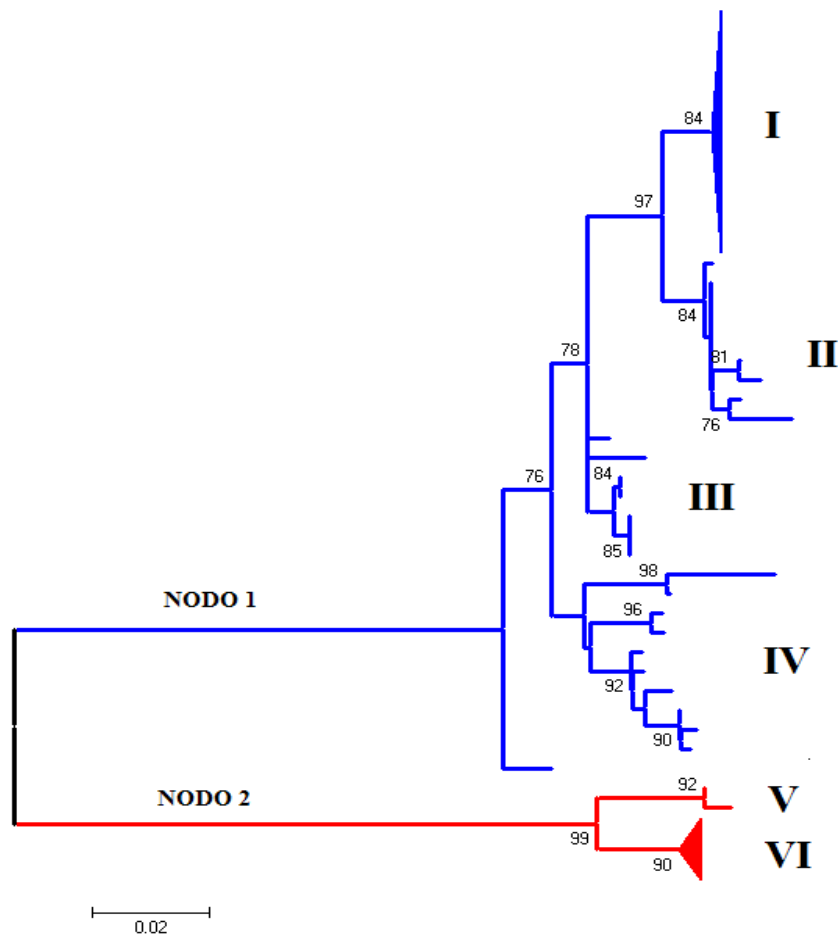
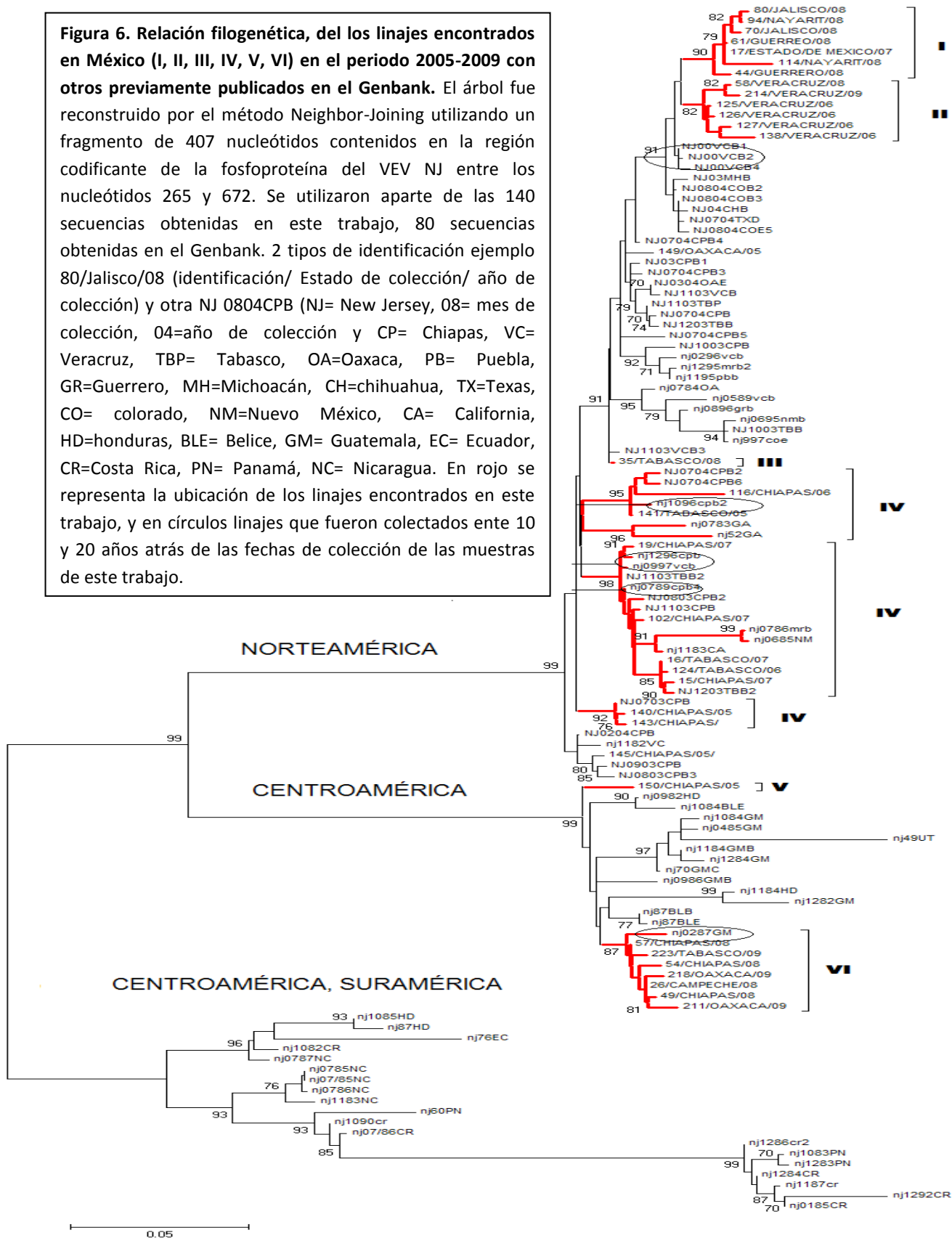


Figura 5. Representación filogenética del virus de estomatitis vesicular en México entre los años 2005 a 2009. El árbol fue construido utilizando un fragmento de 489 nucleótidos, comprendido entre los nucleótidos 184 y 672 de la región codificante del gen de la fosfoproteína del VEV NJ. Se utilizaron 140 secuencias y como método de reconstrucción filogenética se utilizó el método de distancia Neighbor Joining. Los números sobre los nodos representan el valor de bootstrap utilizado como soporte estadístico. Los números romanos representan los diferentes linajes encontrados en el periodo de estudio y fueron determinados en base al nodo ancestral más cercano que da origen a cada uno de ellos. El linaje I contiene virus circulantes en Aguascalientes, Chihuahua, Colima, Estado de México, Guanajuato, Guerrero, Jalisco, Michoacán, Morelos, Nayarit, Oaxaca, Puebla, Querétaro, Sinaloa y Sonora. Linaje II Veracruz y Tabasco, Linaje III Tabasco, Veracruz, Chiapas, Oaxaca, Linaje IV Chiapas y Tabasco, Linaje V Chiapas y Veracruz, y linaje VI Campeche, Chiapas, Oaxaca, Tabasco y Veracruz.

Figura 6. Relación filogenética, del los linajes encontrados en México (I, II, III, IV, V, VI) en el periodo 2005-2009 con otros previamente publicados en el Genbank. El árbol fue reconstruido por el método Neighbor-Joining utilizando un fragmento de 407 nucleótidos contenidos en la región codificante de la fosfoproteína del VEV NJ entre los nucleótidos 265 y 672. Se utilizaron aparte de las 140 secuencias obtenidas en este trabajo, 80 secuencias obtenidas en el Genbank. 2 tipos de identificación ejemplo 80/Jalisco/08 (identificación/ Estado de colección/ año de colección) y otra NJ 0804CPB (NJ= New Jersey, 08= mes de colección, 04=año de colección y CP= Chiapas, VC= Veracruz, TBP= Tabasco, OA=Oaxaca, PB= Puebla, GR=Guerrero, MH=Michoacán, CH=chihuahua, TX=Texas, CO= colorado, NM=Nuevo México, CA= California, HD=honduras, BLE= Belice, GM= Guatemala, EC= Ecuador, CR=Costa Rica, PN= Panamá, NC= Nicaragua. En rojo se representa la ubicación de los linajes encontrados en este trabajo, y en círculos linajes que fueron colectados ente 10 y 20 años atrás de las fechas de colección de las muestras de este trabajo.



6.2 Análisis evolutivo.

6.2.1 Detección de la selección.

Los resultados utilizando la prueba *Evolutionary finger print* en busca de selección positiva no muestran evidencia de selección positiva en los 163 codones contenidos en el fragmento parcial del gen P utilizado en este trabajo (figura 7). Con la finalidad de comprobar estos resultados, y ver que no fueran producto tanto del tamaño de la secuencia como del gen utilizado, se utilizaron 19 secuencias completas del gen de la fosfoproteína (825 nucleótidos-275 codones) números de acceso consecutivos M31862.1 al M31870.1 y M31870.1 al M31880.1, 15 secuencias completas del gen de la nucleoproteína (1263 nucleótidos 423 codones) números de acceso M31845.1 al M31847.1, M31850.1 al M31853.1, M31860.1 y M3186, 32 secuencias del gen de la glicoproteína (1554 nucleótidos-518 codones) números de acceso M214316 al M21437, M21558 al M21568 y M23450, no mostraron evidencia de selección positiva.

El valor global de dN/dS utilizando el método *SLAC* fue de 0.219459, indicando que la selección negativa juega un papel importante en la evolución de este virus. Sin embargo, al calcular los valores dN/dS por separado, se encontró que los linajes II, III y IV presentan valores los cuales sugieren que están bajo un proceso de selección negativa (0.2156, 0.0964, y 0.401267 respectivamente), mientras que los linajes I y VI mostraron valores mucho más cercanos a la neutralidad 0.7 y 0.8 respectivamente así como evidencia de selección positiva en 2 codones en el linaje I y uno en el linaje VI.

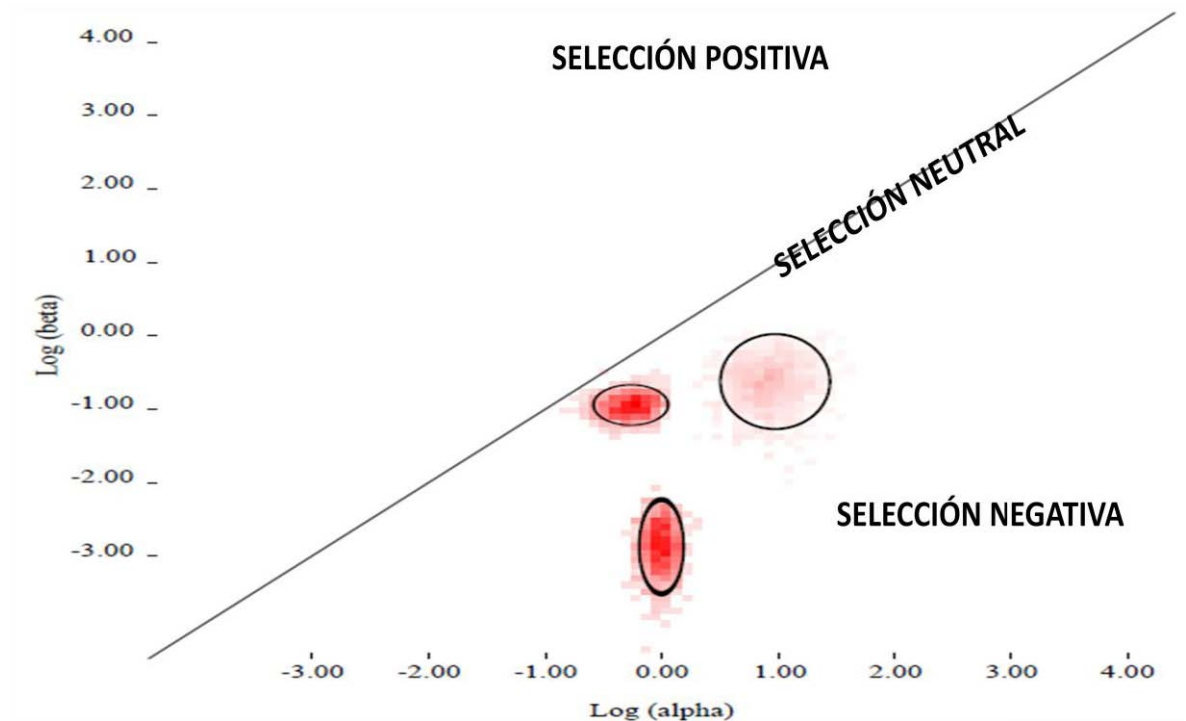


Figura 7. Comparación de los niveles de selección positiva, neutral y Negativa en el VEV NJ. La grafica muestra la distribución estimada del valor ω (dN/dS), realizada en 163 codones del gen de la fosfoproteína del VEV NJ, obtenido por el método *Evolutionary Fingerprint*, disponible en el software DataMonkey (<http://www.hyphy.org/cgi-bin/Datamonkey2007/finishUpload.pl>). La línea diagonal representa un estado idealizado de evolución neutral ($\omega=1$), los puntos ubicados por arriba de esta línea corresponden a codones bajo selección positiva ($\omega>1$), mientras que los puntos por debajo de la línea representan selección negativa ($\omega<1$).

6.2.2 Diferencias en el uso de codones.

El resultado del análisis de los valores RSCU por el método de escalamiento multidimensional, mostró que existen diferencias en el uso de codones sinónimos entre los diferentes linajes analizados, agrupando a linajes de Norte (I, II, III, IV), Centro (V,VI, CAI) y Centro-Sur América (CAII), en cuadrantes específicos (figura 8), similar a la topología mostrada en la reconstrucción filogenética mostrada en la figura 6 elaborada con nucleótidos. Con base en este análisis, en el cuadrante de Norte América, los linajes circulantes en México se dividen básicamente en 2 grupos, en uno de los cuales se agrupan juntos los linajes I y II y en el otro los linajes III y IV. Los linajes V y VI se colocaron en el cuadrante de los virus circulantes en Centro América, coincidiendo con la reconstrucción filogenética de la figura 6.

Con la finalidad de comprobar si las diferencias en el uso de codones podrían estar relacionadas con la adaptación de los virus a diferentes zonas ecológicas, se utilizaron las secuencias publicadas por Rodríguez et al en 1996 (29), donde el autor concluye que factores ecológicos más que temporales están dominando la evolución del virus de EV NJ, y utilizando los valores RSCU de estas secuencias se construyó una matriz de distancia y esta fue representada en un dendograma figura 9. Los resultados mostraron que existe una correlación entre el uso de codones sinónimos y las diferentes zonas ecológicas previamente determinadas por el autor, al obtener una topología similar a la obtenida utilizando nucleótidos.

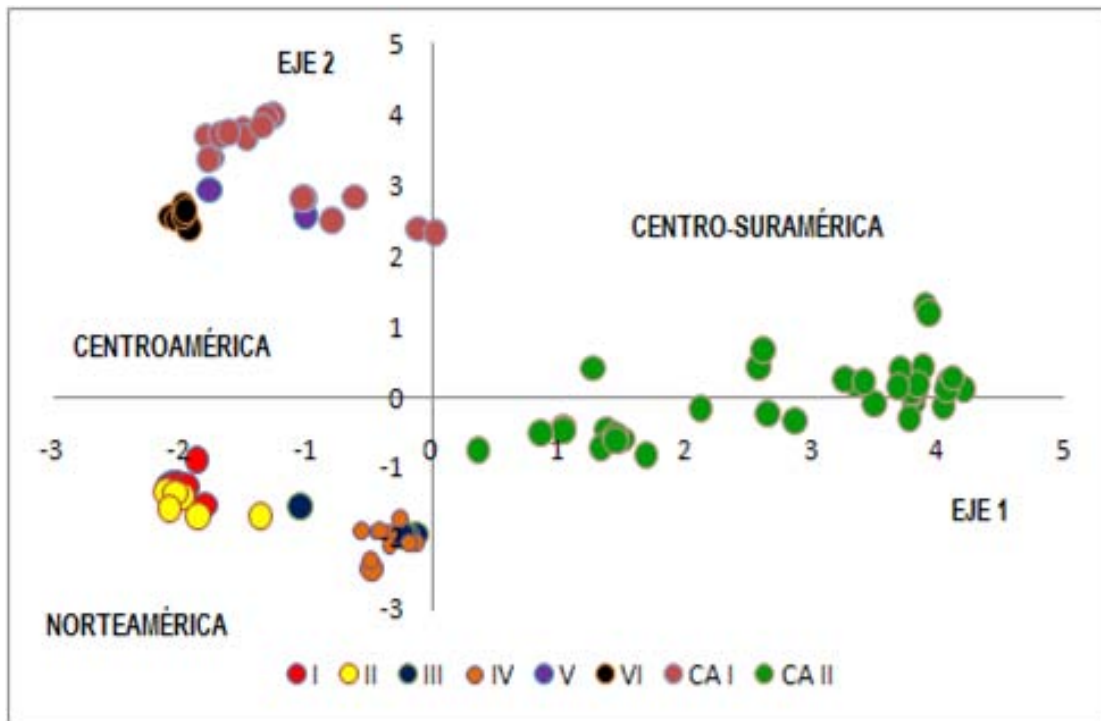


Figura 8. Diferencias en la preferencia por el uso de codones entre los linajes de Norteamérica, Centroamérica y Suramérica. La grafica muestra los resultados del análisis de escalamiento multidimensional realizado a partir de los valores RSCU (Uso relativo de codones sinónimos) de 196 secuencias parciales del gen de la fosfoproteína del VEV NJ, representativas de Norte-Centro y Suramérica. A partir de 64 codones de cada secuencia el análisis de escalamiento multidimensional, reduce estos valores en 2 dimensiones los cuales pueden ser graficados en una grafica de 2 ejes. La ubicación de los puntos en diferentes cuadrantes representa sus diferencias en uso de codones.

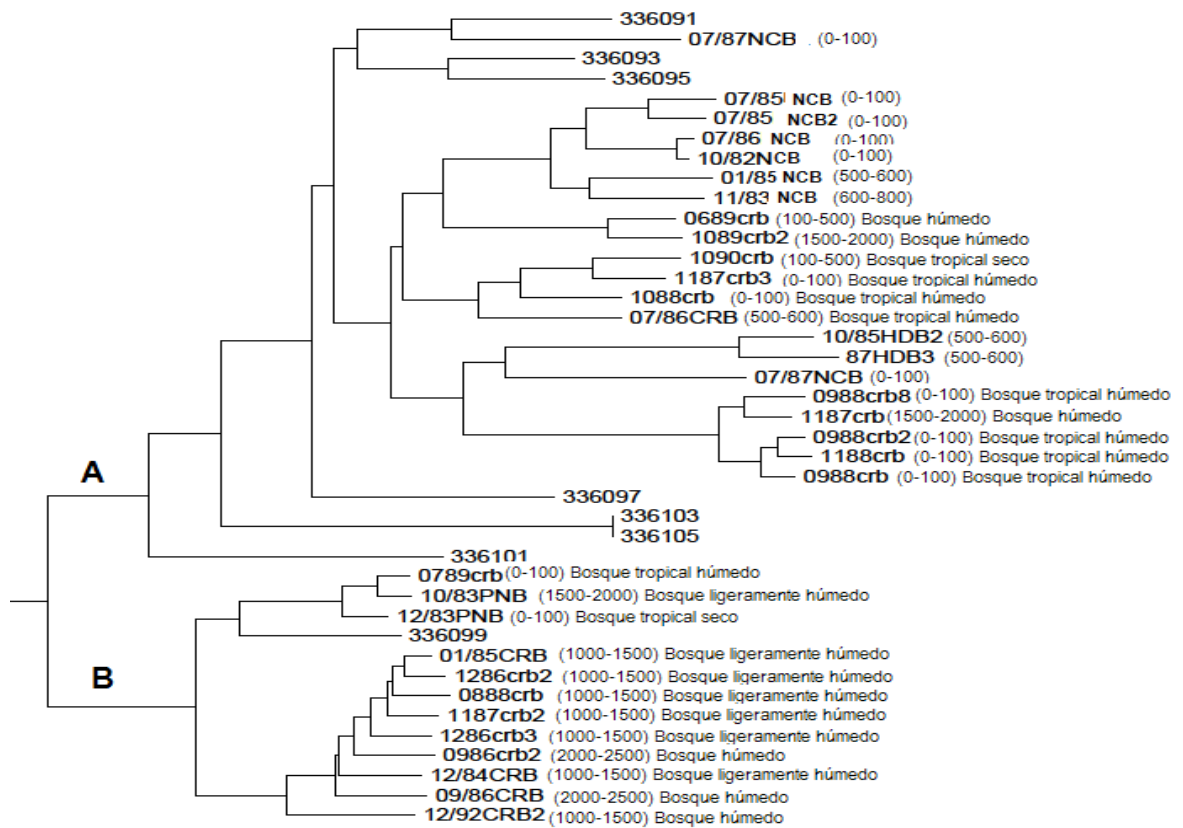


Figura 9. Relación entre uso de codones y adaptación a diferentes zonas ecológicas en el VEV NJ. A partir de los valores RSCU obtenidos de las secuencias parciales del gen de la fosfoproteína, obtenidas de Nicaragua (NC) Honduras (HD) Costa Rica (CR) y Panamá (PN), las cuales fueron publicadas en el artículo Rodríguez L.L., Fitch W.M., Nichol S.T. (1996) Ecological factors rather than temporal factors dominate the evolution of vesicular stomatitis virus. Proc. Natl. Acad. Sci. 93,13030-13035, se obtuvo una matriz de distancias Utilizando el programa estadístico R, la cual fue usada para construir este dendograma. Se muestran 2 nodos marcados como A donde se alojan secuencias con origen geográfico entre 0 y 800 metros sobre el nivel del mar, y el nodo B con secuencias con origen geográfico entre 1000 y 2500 metros sobre el nivel del mar. Esta topología es similar a la obtenida por los autores usando secuencias de nucleótidos. Sugiriendo que el uso de codones en el VEV NJ es filogenéticamente informativo y podría estar relacionado con la adaptación del virus a diferentes nichos ecológicos.

6.2.3 Frecuencia de dinucleótidos.

La distribución de la frecuencia de dinucleótidos ha sido reconocida como un buen estimador para determinar presiones selectivas, así mismo se ha asociado con influenciar el uso de codones (47). Los valores obtenidos son usualmente diversos entre especies y muy invariantes a lo largo de un genoma individual, por estas razones el patrón de dinucleótidos de una secuencia es referido como su firma genética (48).

Como lo muestra la figura 10, existen diferencias en el patrón de distribución dinucleotídica entre los diferentes linajes. Todos presentan una sobrerrepresentación de ApG y GpA así como una subrepresentación CpG, GpC, GpT, TpA, TpT. Sin embargo se diferencian ya que en los linajes representativos de Norteamérica (I, II, III y IV) presentan una sobrerrepresentación de ApA y CpA, los linajes de Centroamérica (VI) una sobrerrepresentación de ApA y finalmente los linajes de Centro-Suramérica (CAII) una sobrerrepresentación de CpA y TpC y una subrepresentación de GpG.

6.2.4 Efecto del patrón de dinucleótidos en el uso de codones.

Para establecer el efecto del patrón de dinucleótidos en el uso de codones sinónimos, se seleccionaron aminoácidos que fueran codificados por más de 3 codones y que tuvieran dinucleótidos sobre y subrepresentado como serina (UCU, UCC, UCA, UCG), prolina (CCU, CCC, CCA, CCG), treonina (ACU, ACC, ACA, ACG), alanina (GCU, GCC, GCA, GCG), glicina (GGU, GGC, GGA, GGC), Arginina (CGU, CGC, CGA, CGG, AGA, AGG).

Los resultados muestran (figura 11) que los linajes de Norteamérica (I, II, III, y IV) tienen una tendencia por usar más frecuentemente codones que contienen algún dinucleótido sobrerrepresentado como CpA o GpA al utilizar para codificar estos aminoácidos los codones UCA, CCA, ACA, GCA, GGA, AGA, sobre codones que tienen dinucleótidos subrepresentado como ACG, GCG, CGU, CGC, CGA, CGG.

En el caso del linaje VI representante de Centroamérica, la selección del uso de codones, no parece estar influenciada por la composición de dinucleótidos, ya que el único dinucleótido sobrerrepresentado, considerado dentro de estos codones es GpA, el cual en el caso de la glicina es usado en el codón GGA, preferentemente sobre el codón que presenta el

dinucleótido subrepresentado GGC, sin embargo en el caso de la arginina existe una preferencia por usar el codón CGU sobre los codones AGA, AGG.

Finalmente en el caso del linaje determinado como CAII, se observa una situación similar a los linajes de Norteamérica, donde existe una predilección por codones que tienen dinucleótidos sobre representados para sintetizar los aminoácidos estudiados AGA, GGA, GCA, UCA, evitando el uso con dinucleótidos subrepresentados como CpG.

Por otro lado, al analizar los codones con valores RSCU mayores a 1 (figura 11), de estos el 67 % presentan en la posición 3 del codón una A o U, lo cual coincide con la composición de la secuencia que es A- 32.9 %, U -24.9%, G - 21.89%, C -20.24 %).

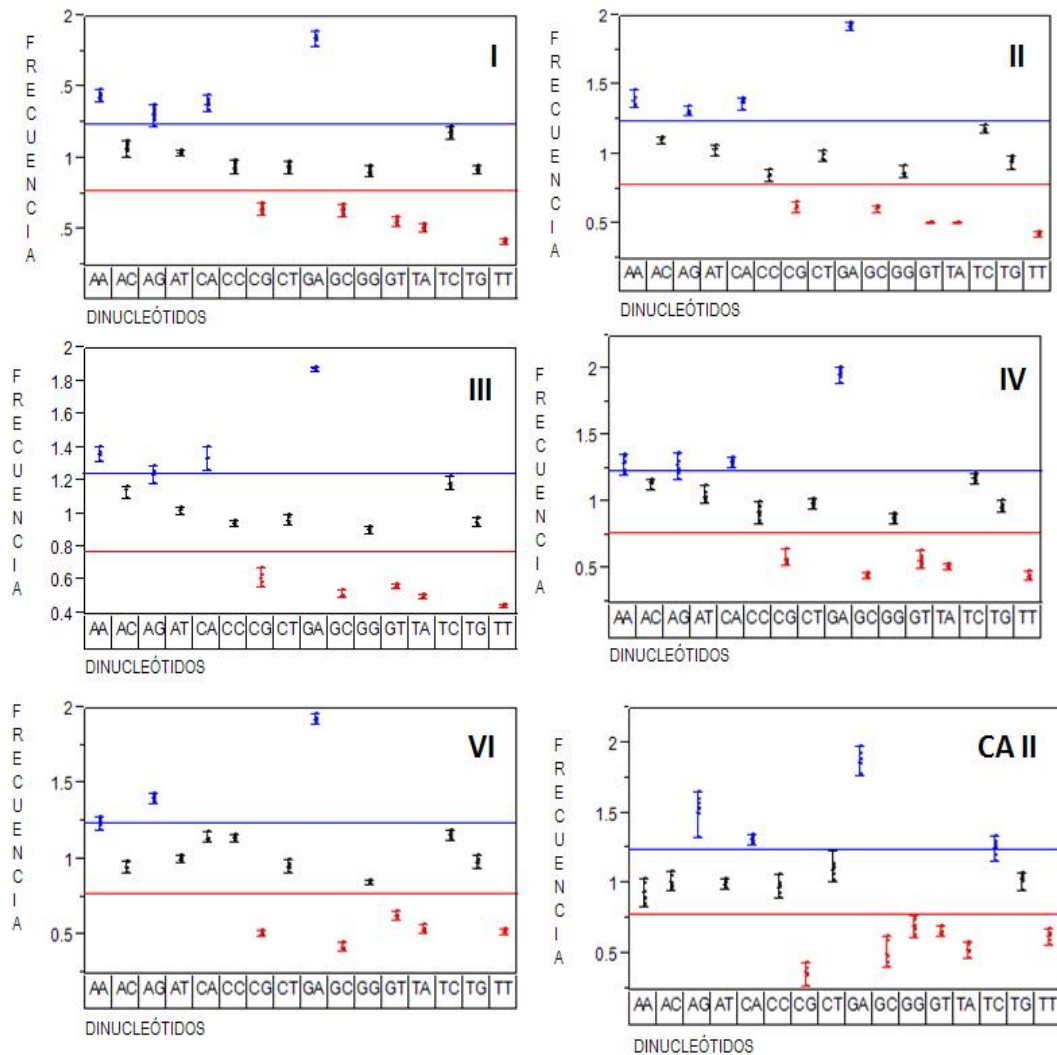


Figura 10. Comparación de las frecuencias de dinucleótidos entre los diferentes linajes del VEV NJ estudiados en este trabajo. Las frecuencias de dinucleótidos fueron calculadas para cada linaje de manera individual a partir de secuencias parciales del gen de la fosfoproteína del VEV NJ y posteriormente estos valores fueron graficados. Las líneas azul (1.25) y roja (0.78), representan los puntos de corte para indicar que un dinucleótido está sobrerrepresentado (color azul) o subrepresentado (color rojo) respectivamente. Los dinucleótidos en color negro presentan frecuencias con valores esperados de acuerdo a la composición nucleotídica de la secuencia.

I		II		III		IV		VI		CA II	
AGA	3.85	AGA	4	AGG	3	AGG	2.68	UCU	2.01	UCA	3.1
GGA	2.28	GGA	2.67	UCA	2.06	UCA	2.08	CAU	2	AGA	3.05
UCA	2.01	UUU	2	UAU	2	AGA	2.04	CGU	2	GCA	2.63
CAU	2	UAU	2	CAU	2	GGA	2.03	AAU	1.71	GGA	2.46
AGG	2	CAU	2	GGA	2	UUU	2	AGU	1.71	CCU	2.45
UUU	1.96	AGG	2	UUG	1.85	GCA	2	AUU	1.63	UUG	2.16
UAU	1.96	UCA	1.96	GCA	1.75	CAU	2	CCU	1.57	AAU	1.92
UUG	1.68	UUG	1.9	CGA	1.5	UAU	1.9	UUA	1.48	AUU	1.78
CCA	1.53	CCA	1.78	AGA	1.5	UUG	1.83	UCA	1.44	CAU	1.73
UGC	1.5	ACU	1.67	ACU	1.45	ACA	1.51	GCC	1.44	ACU	1.7
ACA	1.47	UCU	1.58	ACA	1.45	ACU	1.47	GUA	1.43	UUA	1.51
ACU	1.44	UGC	1.5	AAA	1.38	GUC	1.43	CAA	1.43	GUU	1.44
CAA	1.43	CAA	1.43	UUU	1.33	AAA	1.31	CGA	1.43	UUU	1.4
AAA	1.43	AAA	1.42	GUC	1.33	GAA	1.29	UUG	1.39	UAU	1.36
GUC	1.35	GUU	1.33	CCU	1.3	UCU	1.28	CUC	1.39	AAG	1.36
GUU	1.32	GUC	1.33	CCA	1.3	CGA	1.28	GAA	1.38	GAU	1.26
GAA	1.29	ACA	1.33	GAU	1.3	CCU	1.27	GGA	1.38	GUC	1.21
UCU	1.24	GCU	1.33	CAA	1.29	GAU	1.16	GCU	1.28	UCU	1.18
UCC	1.24	GAA	1.28	GAA	1.29	GUU	1.15	GCG	1.28	GUA	1.17
AAU	1.22	AAU	1.23	UCU	1.25	CCC	1.14	UAU	1.28	CGG	1.09
CCU	1.21	GAU	1.21	GUU	1.22	GCU	1.1	UGU	1.28	AGG	1.09
GAU	1.16	GCA	1.17	UGU	1.13	AUU	1.08	GAU	1.27	GAG	1.07
GGU	1.15	AUC	1.07	ACG	1.09	GGU	1.08	GGU	1.27	CAA	1.06
AUC	1.13	CCU	1.04	AAC	1.09	CAG	1.07	AAG	1.2	UGC	1.03
AUU	1.11	UCC	1	AUU	1	AAC	1.06	AGG	1.2	ACC	1.01
ACG	1.08	ACG	1	AUC	1	UCC	1.05	UCC	1.11	GCU	0.98
GCU	1.06	AUU	0.96	AUA	1	CUA	1.02	GUU	1.1	UGU	0.97
GCA	1.06	AUA	0.96	UCC	1	UGU	1	ACC	1.1	CAG	0.94
GCG	0.98	GUA	0.89	GCU	1	UGC	1	ACG	1.1	GAA	0.93
GCC	0.9	UUA	0.88	GCC	1	UUA	0.98	UUU	1.06	GGG	0.92
GUA	0.88	CUU	0.88	UUA	0.92	AUC	0.98	ACA	1.05	AUA	0.86
UUA	0.87	CUA	0.88	CUU	0.92	GUA	0.97	AUA	0.97	CCA	0.86
CUC	0.87	GCG	0.83	CUA	0.92	AUA	0.95	UUC	0.94	CUC	0.76
CUA	0.87	GAC	0.79	CUG	0.92	ACG	0.94	CCC	0.81	CGA	0.76
CUG	0.87	AAC	0.77	AAU	0.91	AAU	0.94	CCA	0.81	GAC	0.74
CUU	0.85	CCC	0.74	GUA	0.89	CAA	0.93	CCG	0.81	CUU	0.7
CCG	0.84	CUC	0.73	UGC	0.88	GCC	0.9	AAA	0.8	ACG	0.69
GAC	0.84	CUG	0.73	AGU	0.75	CCA	0.86	AGA	0.8	CUG	0.65
AAC	0.78	GAG	0.72	CAG	0.71	CUU	0.85	GUG	0.76	UAC	0.64
AUA	0.76	UCG	0.71	GAG	0.71	GAC	0.84	ACU	0.75	AAA	0.64
AGU	0.73	GCC	0.67	CCC	0.7	CCG	0.73	GAC	0.73	GGU	0.62
GAG	0.71	GGU	0.67	CCG	0.7	AGU	0.73	GGC	0.73	UCC	0.61
CAG	0.57	GGC	0.67	GAC	0.7	CUC	0.72	CUG	0.72	UUC	0.6
AAG	0.57	AAG	0.58	UCG	0.69	GAG	0.71	UAC	0.72	ACA	0.6
GGC	0.55	CAG	0.57	UUC	0.67	AAG	0.69	UGC	0.72	AGU	0.47
UGU	0.5	UGU	0.5	GGU	0.67	GGC	0.61	GUC	0.71	AUC	0.37
UCG	0.49	AGU	0.5	GGC	0.67	CUG	0.6	CUA	0.67	AGC	0.37
GUG	0.45	GUG	0.44	GGG	0.67	UCG	0.58	GAG	0.63	CCG	0.36
CCC	0.42	CCG	0.44	AAG	0.62	GUG	0.46	GGG	0.63	CCC	0.32
AGC	0.28	AGC	0.25	GUG	0.56	AGC	0.28	UCG	0.57	UCG	0.27
CGC	0.07	UUC	0	CUC	0.46	GGG	0.27	CAG	0.57	CAC	0.27
CGG	0.07	ACC	0	GCG	0.25	UAC	0.1	CGG	0.57	CUA	0.22
UUC	0.04	UAC	0	AGC	0.25	ACC	0.08	AUC	0.41	GCC	0.2
UAC	0.04	CAC	0	ACC	0	UUC	0	CUU	0.36	GCG	0.2
GGG	0.02	CGU	0	UAC	0	GCG	0	AAC	0.29	GUG	0.18
ACC	0	CGC	0	CAC	0	CAC	0	AGC	0.29	AAC	0.08
CAC	0	CGA	0	CGU	0	CGU	0	GCA	0	CGU	0
CGU	0	CGG	0	CGC	0	CGC	0	CAC	0	CGC	0
CGA	0	GGG	0	CGG	0	CGG	0	CGC	0	GGC	0

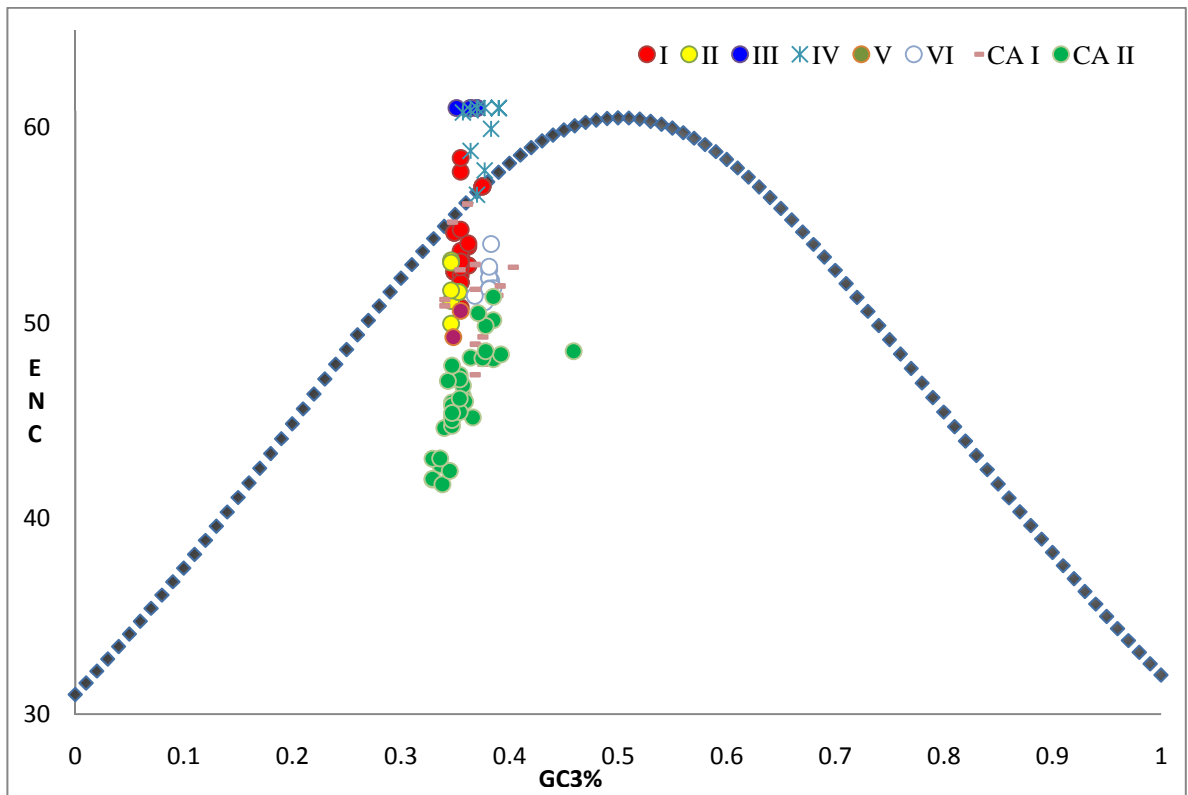
Figura 11. Lista ordenada de codones en los diferentes linajes del VEV NJ estudiados en este trabajo en relación a sus valores RSCU. Los promedios de los valores RSCU, fueron calculados para cada linaje de acuerdo a lo definido en la sección de material y métodos. Los codones de cada linaje fueron ordenados de mayor a menor de acuerdo a su valor RSCU. Cada linaje presenta 2 columnas, la izquierda que contiene los codones, los cuales están en azul y en rojo si presentan algún dinucleótido sobre o subrepresentado respectivamente, de acuerdo a lo representado en la figura 10. La columna derecha contiene los valores RSCU en amarillo >1 y en gris <1.

6.2.4 Numero efectivo de codones.

Los resultados del análisis del uso efectivo de codones mostraron que existe una diferencia por el uso de determinados codones entre los diferentes linajes del VEV NJ, obteniendo valores que van de 41.75 a 61. Estos valores en general tuvieron una distribución homogénea entre cada uno de los respectivos linajes, siendo los linajes de Norteamérica III y IV los que presentan los valores ENC mas altos y el linaje CAII de Centro-Suramérica los valores más bajos. (Tabla 1). Al comprobar si el GC3% estaba influenciando el valor ENC, se encontró que los valores de GC3% variaban entre 32.9 a 45.9%, y estos solo explicaban el valor ENC del 3% de las secuencias, el 97 % de las secuencias presentaron valores ENC los cuales no pueden ser explicados por el porcentaje de GC3, sugiriendo que un proceso selectivo podría estar influenciando el uso de codones en el VEV NJ. (Tabla 1 y gráfica 1).

Linaje	ENC	GC3%	ENC esperado
I	53.5	35.6	55.56
II	51.71	34.9	54.95
III	61	36.38	56.14
IV	59.88	37.42	57.22
VI	51.95	38.05	57.5
CAI	51.22	36.5	56.3
CAII	46.2	35.82	55.56

Tabla 1. Comparación de los valores del número efectivo de codones esperados versus calculados entre los diferentes linajes del VEV NJ. En relación a su porcentaje de CG3 %, se obtuvieron los valores esperados de ENC para cada uno de los linajes y estos fueron comparados con su valor ENC sin considerar su porcentaje de CG3, demostrando que la composición de CG3 no explica el uso de codones en el VEV NJ.



Gráfica 1. Gráfica representando el número efectivo de codones. El número efectivo de codones usado por cada linaje del VEV NJ fue graficado contra su composición de GC3. La curva representa el valor esperado de ENC, si este estuviera influenciado únicamente por su contenido de GC3. En general es admitido que si el uso de codones de un gen en particular está relacionado con el contenido de GC3, este se verá reflejado en la gráfica sobre o ligeramente bajo la línea de la curva esperada, por otro lado si un gen está sujeto a selección por traducción, este caerá por debajo de la curva esperada

6.3 Epidemiología molecular del VEV NJ en México.

Con base en las características filogenéticas, de uso de codones y geográficas, es posible sustentar en este periodo de estudio (2005-2009) la presencia 3 zonas de circulación viral. La primera localizada aproximadamente entre $17^{\circ}49'28.51''\text{N}$ $95^{\circ}48'54.05''\text{O}$ y $20^{\circ}57'52.45''\text{N}$ $97^{\circ}24'16.19''\text{O}$, la cual abarca una zona de circulación de 350 kilómetros aproximadamente, que corresponde a una zona ecológica clasificada como tropical húmeda (anexo 3) (50), la cual comprende el estado de Veracruz en los municipios de Alvarado, Boca del Rio, Playa Vicente, Tierra Blanca, Tuxpan, Cosautlán de Carvajal. En esta zona se mantuvieron circulando entre 2005 y 2009 los virus que fueron clasificados dentro del linaje II (anexo 2), a una altura a nivel del mar entre 7 y 64 metros, con excepción del virus encontrado en Cosautlán de Carbajal que se aisló a 1250 metros sobre el nivel del mar.

La segunda zona se encontró localizada $15^{\circ}25'58.68\text{N}$ $92^{\circ}53'52.22\text{O}$ y $18^{\circ}11'07.94\text{N}$ y $93^{\circ}02'48.15\text{O}$, abarcando una zona de circulación de aproximadamente 325 kilómetros y comprende zonas ecológicas clasificadas como tropical húmeda y tropical subhúmeda, la cual comprende los estados de Tabasco (Huimanguillo, Jonuta, Emiliano Zapata, Jalpa), Chiapas (Tecpan, Tapalapa, Juarez, Pichulalco, Coapilla, Villa Flores, Mapastepec) y Oaxaca (Reforma de Pineda). En estas zonas los virus clasificados en los linajes III y IV se mantuvieron circulando entre 2005 y 2008, a una altura sobre el nivel del mar entre 5 y 29 metros en los estados de Tabasco y Oaxaca y entre 450 y 1645 metros sobre el nivel del mar en el estado de Chiapas.

La tercera zona localizada entre $16^{\circ}25'37.48\text{N}$ $97^{\circ}58'35.71''\text{O}$ y $28^{\circ}38'07.00''\text{N}$ $106^{\circ}05'20.00''\text{O}$, es una zona de circulación de aproximadamente 2000 kilómetros en la que se encuentran principalmente las zonas ecológicas tropical húmeda y templada subhúmeda en los estados de Aguascalientes, Chihuahua, Colima, Estado de México, Guanajuato, Guerrero, Jalisco, Michoacán, Morelos, Nayarit, Oaxaca, Puebla, Querétaro, Sinaloa y Sonora. En esta zona circuló el linaje I entre 2006 y 2009. Este linaje presenta un uso de codones similar a los veracruz pertenecientes al linaje numero II, los cuales circulan en el estado de Veracruz, lo cual puede indicar el posible origen genético de este linaje, considerando que linajes que circulan en Veracruz relacionados al linaje linaje II

(NJ00VCB1, NJ00VCB2 y NJ00VCB4) han estado circulando en ese estado por lo menos 6 años antes de la aparición del linaje I (ver figura 6).

Soportando esta hipótesis se encuentra el hecho de que el primer reporte de circulación de este linaje fue en el mes de octubre del 2006 en el estado de Puebla, estado que colinda con el estado de Veracruz, posteriormente apareció en Guerrero, Morelos, Michoacán. Para 2007 el linaje circulo en Guerrero, Michoacán, Oaxaca, Estado de México, durante el 2008 en Aguascalientes, Colima, Guanajuato, Guerrero, Jalisco, Michoacán, Morelos, Nayarit, Oaxaca, Querétaro. Finalmente para el 2009 este linaje se encontraba en la parte norte de México en los estados de Chihuahua, Sinaloa y Sonora, para completar un total de 240 casos confirmados por este linaje entre 2006 y el 2009, que comparado con 16 casos reportados durante el 2005 antes de que iniciara su circulación este linaje, representa un aumento importante en la incidencia de esta enfermedad en México (figura 12 A).

Finalmente, durante este periodo de estudio, se detectó la presencia del linaje VI un linaje , procedente de Centroamérica, el cual fue detectado durante 2008 y 2009 en los estados de Tabasco, Chiapas, Oaxaca, Campeche y Veracruz. Considerando su relación genética (figura 6) con virus de Guatemala, se puede proponer que este linaje ha logrado diseminarse 682 km en 2 años hasta Veracruz (figura 12 B). Además es importante mencionar que para 2010 este linaje fue el responsable de la mayor parte de los casos positivos al VEV NJ en México y se diseminó hacia el norte, pues fue aislado en los estados de Querétaro y Guanajuato es decir a 1288 kilómetros de su lugar de origen. En Guanajuato durante el 2010 se aisló a partir de un hato de carbas, lo cual contrasta con lo que sucede con los linajes de origen Norteamericano, donde se reportan infecciones principalmente en bovinos y equinos (Figura 12 B).

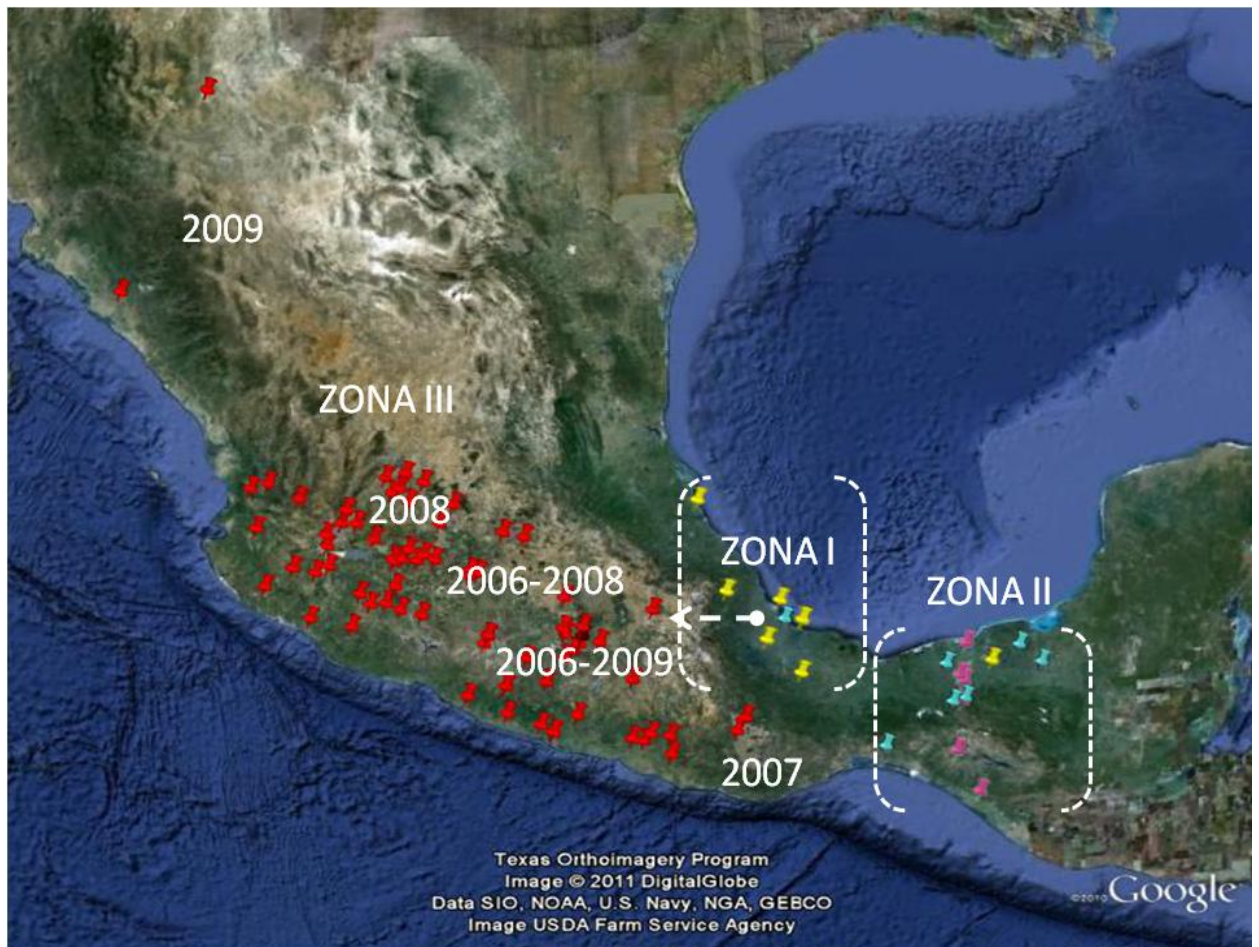


Figura 12(A). Representación grafica de la distribución geográfica del VEV NJ en México. Utilizando los datos referentes a estados y municipios donde fueron aisladas cada una de las secuencias, estos fueron localizadas en el mapa de la republica mexicana utilizando el software Google maps (<http://www.google.com/intl/es/earth/index.html>), los colores para cada linaje fueron designados de acuerdo con la reconstrucción filogenética presentada en la figura 5, correspondiendo el linaje I en rojo, el linaje II en amarillo, el linaje III en azul y el linaje IV en rosa. Las zonas I y II fueron divididas de acuerdo con su ubicación geográfica. La flecha en la zona I indica el posible origen del linaje I el cual circulo en la zona denominada como numero III. Esta relación fue sustentada por la característica de uso de codones similares que presentan los linajes I y II.

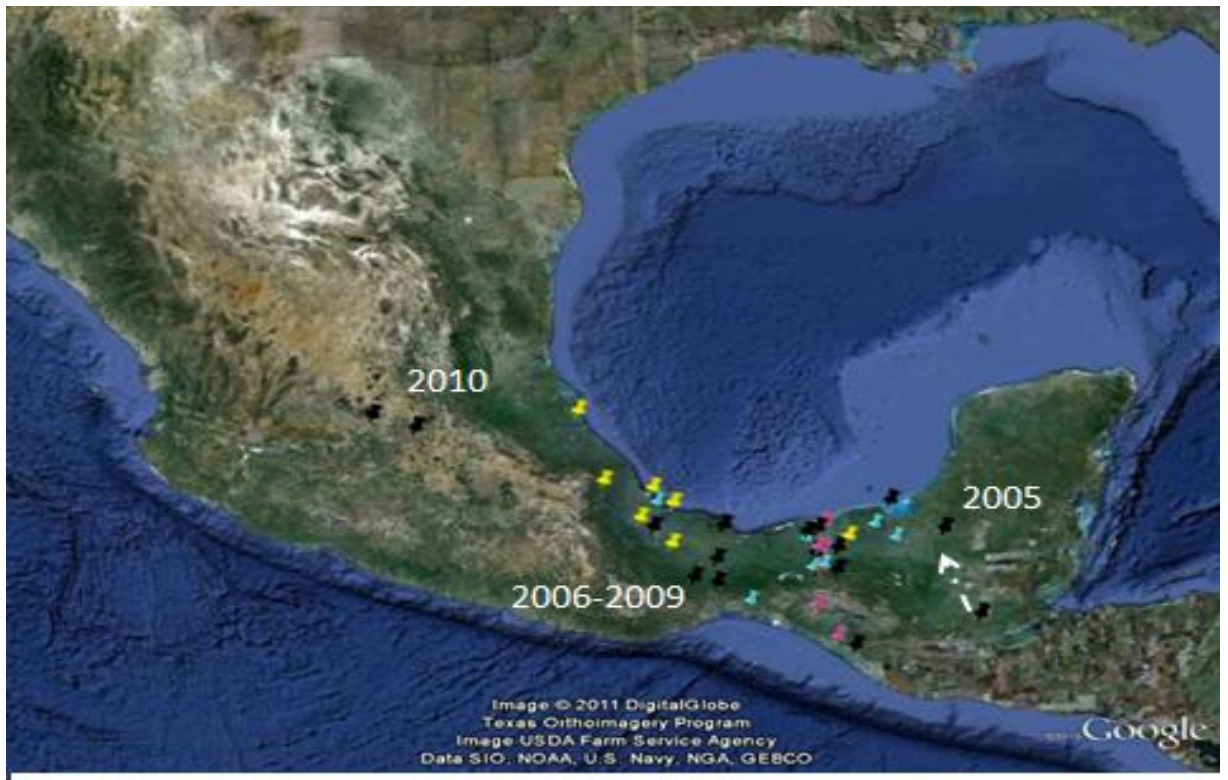


Figura 12 (B). Representación grafica de la distribución geográfica del VEV NJ en México

Se presenta la distribución geográfica del linaje VI (negro) el cual tiene relación genética con linajes circulantes en Centroamérica. Este linaje se mantuvo circulando entre el 2005 y el 2009 junto con los linajes II (amarillo), III (azul) y IV (rosa) en la zona sur de México. Sin embargo para el 2010 este linaje se fue aislado en la zona centro norte de México, en este año no hubo reportes de circulación del linaje I (Rojo), lo que podría representar un potencial brote en la zona norte de México en años próximos.

7. Discusión.

7.1 Detección de la selección.

Mucho del conocimiento sobre el proceso de evolución molecular se deriva del análisis del tipo de mutaciones acumuladas a través del tiempo. Debido a que el código genético es degenerado, algunas provocan un cambio en el aminoácido codificado mientras que otras no tendrán este efecto. Tradicionalmente se considera que una mutación sinónima o silenciosa no altera la función de la proteína codificada, por lo tanto estas mutaciones se consideran neutras. Se puede inferir que tipo de fuerzas evolutivas están operando sobre una población, al comparar la relación del número de mutaciones no-sinónimas (dN) por sitio no sinónimo con el número de mutaciones sinónimas por sitio sinónimo (dS). En ausencia de selección, cualquier mutación tiene la misma posibilidad de fijación por lo que el valor ω (dN/dS) debe ser cercano a 1. Cuando opera la selección negativa o purificadora, la tasa de fijación de las mutaciones sinónimas es superior a la tasa de fijación de las mutaciones no-sinónimas, esto debido a su efecto deletéreo en la adecuación de la proteína ($dS > dN$). Por otro lado un incremento en la tasa de fijación de mutaciones no-sinónimas ($dS < dN$), es considerado como evidencia de selección positiva, en sitios donde las mutaciones representan un aumento en la adecuación de la proteína correspondiente (51).

Sin embargo, evidencia que indica que las mutaciones silenciosas pueden no ser siempre neutras. Por ejemplo, mutaciones silenciosas pueden tener efecto en el control de la expresión de las proteínas, splicing, disponibilidad de RNAs de transferencia y resistencia a los RNAs de interferencia, este efecto sobre la función que tienen las mutaciones silenciosas, pueden contribuir a que estas mutaciones no se comporten de manera neutral (52, 53,54,55,56).

En este estudio observamos que una de las principales fuerzas evolutivas que está actuando sobre la evolución del virus de estomatitis vesicular, es la selección negativa. Este mismo patrón evolutivo ha sido reportado en otros virus RNA transmitidos por artrópodos (57), sugiriendo que el requerimiento en los arbovirus para replicarse en mamíferos e invertebrados como parte de su ciclo infeccioso, impone una mayor restricción selectiva que en los virus que sólo infectan especies filogenéticamente similares (58).

Una posible explicación es que cambios a nivel de proteico, que podrían ser ventajosos para establecer la infección en el mamífero, podrían resultar adversos para el invertebrado o viceversa (59).

Otra posible explicación a este fenómeno, es dada por Arrigo y colaboradores en un estudio sobre encefalitis equina del Este, donde encuentran, de manera similar a los resultados obtenidos en este trabajo, que los virus de encefalitis equina Venezolana y del Este, tienen patrones de divergencia genética, caracterizados por la evolución de múltiples subtipos y linajes altamente conservados filogenéticamente que se agrupan en clados geográficos, sugiriendo a la selección negativa como una de las principales fuerzas evolutivas actuando en este virus, como consecuencia a su adaptación y estabilidad en determinados ambientes ecológicos (60).

En nuestro estudio, aunque todos los linajes presentaron niveles globales de dN/dS indicativos de selección Negativa, en el caso de los linajes I y VI que son linajes que surgieron de zonas endémicas en México y Centroamérica respectivamente, estos presentan niveles dN/dS mucho más cercanos a la neutralidad, posiblemente como cambio en el tamaño de la población.

En el 2004 Novella y colaboradores en un estudio llevado a cabo en estomatitis vesicular Indiana, obtuvieron la secuencia completa de 10 poblaciones virales que evolucionaron alternando pases en líneas celulares derivadas de mamíferos e insectos, encontraron como resultado del proceso que , un número significativo de las mutaciones fijadas eran silenciosas, concluyendo que las mutaciones silenciosas pueden contribuir significativamente en la adaptación de los virus RNA y las frecuencias relativas de sustituciones dN/dS podrían no ser útiles para determinar las fuerzas evolutivas que están actuando sobre estas poblaciones (51).

7.2 Uso de codones

Para determinar el efecto que las mutaciones silenciosas pueden tener sobre la evolución del virus de EV, exploramos en este trabajo el efecto que estas tenían sobre la preferencia en el uso de codones, los resultados sugieren que el uso de codones puede ser filogenéticamente informativo en el virus de EV NJ, proponiendo que en este virus existe una selección positiva que está moldeando el uso de codones entre los linajes circulantes en

América, como resultados de las necesidades de replicación del virus en diferentes organismos para mantenerse en diferentes nichos ecológico a través del continente americano. Para soportar esta hipótesis analizamos las secuencias utilizadas por Rodríguez y colaboradores en 1996, donde ellos encontraron una relación filogenética entre VEV NJ que comparten una misma zona ecológica y concluyeron que factores ecológicos más que temporales dominan la evolución del virus de estomatitis vesicular (29). En nuestra investigación utilizando los cálculos de uso relativo de codones sinónimos encontramos la misma topología obtenida por Rodríguez y colaboradores, y proponemos que existe una relación entre el uso de codones de un VEV NJ y el nicho ecológico donde este es mantenido.

Los 2 modelos principales que han sido propuestos para explicar la preferencia en el uso de codones son el relacionado con la traducción (modelo selectivo) o con las mutaciones (modelo neutral). En el modelo relacionado con la traducción, se postula que existe una coadaptación de uso de codones sinónimos y la abundancia de RNAs de transferencia para optimizar la eficiencia durante la traducción. Se espera una correlación entre el uso de codones y la expresión de un gen (61). En el virus Epstein- Barr el estado latente de los genes está relacionado con la desoptimización del uso de codones para reducir la competencia por la traducción con las células del hospedador (62). En el virus del papiloma, el uso de codones se ve optimizado durante la replicación en cierto tipo de células (63). Ha sido posible realizar la atenuación del virus de la polio, reduciendo la traducción viral de los genes, construyendo un virus con un uso de codones sinónimos desfavorables (64).

En un estudio donde se analizó el uso de codones en 100 bacteriófagos como parte de su adaptación a diferentes bacterias, se encontró que los genomas de los bacteriófagos se encuentran bajo una presión selectiva a nivel de codones impuestas por las preferencias de uso de codones de los respectivos hospedadores, ya que esta selección otorga una ventaja para la síntesis viral al mejorar la eficiencia en la traducción (65).

El modelo mutacional, postula que la composición genética puede influenciar el uso de codones de una manera no selectiva y esto ha sido determinado en diferentes especies (66). En particular la influencia en el uso de codones en una secuencia reducida en el

dinucleótido CpG puede ser el resultado del efecto de respuestas antivirales por parte de la célula y no una selección directa del uso de codones (66). Es importante hacer notar que ambos modelos no son mutuamente excluyentes (61).

En este sentido, los resultados de este trabajo sugieren que la composición genómica del VEV NJ tiene un efecto en el uso de codones sinónimos, habiendo una predisposición por usar codones con terminación en A o T, los cuales están en mayor proporción en la composición nucleotídica del virus. Por otro lado, encontramos evidencia de selección a favor de la traducción, al graficar los valores ENC de las secuencias contra los GC3%. En general es admitido que si el uso de codones de un gen en particular está relacionado con el contenido de GC3%, éste se verá reflejado en la gráfica sobre o ligeramente bajo la línea de la curva esperada, por otro lado si un gen está sujeto a selección por traducción, este caerá por debajo de la curva esperada (67). La mayor parte de los linajes, presentó valores que indican selección a nivel de uso de codones. Sin embargo no se puede determinar cual modelo contribuye más a la preferencia por el uso de codones en el VEV NJ.

7.3 Frecuencias de dinucleótidos.

El termino arbovirus no es una designación taxonómica, se usa para referirse a virus de RNA, cuyo ciclo de vida comprende la replicación en hospedadores vertebrados y vectores hematófagos (68). El VEV es un parásito intracelular obligado, el cual depende de la maquinaria celular para llevar a cabo su ciclo de replicación, por lo que la interacción virus-hospedador, lleva consigo un proceso de coevolución, donde el proceso evolutivo del virus está ligado a la biología del hospedador. Esta interacción biológica entre el virus y el hospedador, es un delicado balance de acciones y contracciones entre el sistema inmune del hospedador y los mecanismos de escape del virus (21).

Se sabe que los genomas de los vertebrados tienen una subrepresentación de los dinucleótidos CpG, la hipótesis actual para explicar esto, está relacionada con el modelo de metilación-desaminación de las citocinas (69). Este modelo dice que los genomas de los vertebrados, están sujetos a la metilación en contexto de los dinucleótidos CpG, con la finalidad de preservar información epigenética de una región determinada en el genoma, propensa a la transcripción o el silenciamiento (70). Las citocinas metiladas son propensas a mutar a timinas a través de un proceso de desaminación espontanea, generando el

dinucleótido TpG, produciendo un apareamiento incorrecto con la cadena opuesta de DNA. Este suceso causa una mutación en la cadena opuesta si la replicación de lleva acabo sin reparación de este error, produciendo la aparición del dinucleótido CpA (71). Los organismos que no metílan DNA como los insectos, no presentan una subrepresentación de CpG, y un exceso de CpA (72).

Recientemente, surgió una hipótesis para explicar la subrepresentación del dinucleotido CpG inducida por el hospedador en los genomas de los virus de RNA que infectan vertebrados, la cual se basa en el reconocimiento del dinucleótido CpG por el sistema inmune del hospedador. Se sabe que el dinucleotido CpG no metilado, puede inducir la respuesta inmune en los mamíferos (73).

El patrón de CpG es reconocido por el receptor de reconocimiento de patrón intracelular (PPR) Toll-like 9 (TLR9). Esto siguiere que los genomas de los virus de RNA que infectan vertebrados, son seleccionados para imitar las cualidades moleculares de los RNAs mensajeros del hospedador con la finalidad de evadir la respuesta inmune del hospedador por un mecanismo inmunológico aún desconocido (72)

Otro dinucleótido que se encuentra subrepresentado en los genomas de vertebrados e insectos es el TpA, para lo cual se han propuesto 2 teorías. La primera se refiere a que el dinucleótido TpA se encuentra presente en 2 de 3 codones de paro, por lo que la subrepresentacion de TpA en secuencias codificantes, reduce la posibilidad de mutaciones sin sentido, es decir mutaciones que pudieran limitar la traducción de un gen al encontrar una señal de paro.

La otra teoría atribuye la ausencia de TpA en secuencias codificantes, ya que este dinucleotido se encuentra presente en regiones no codificantes controlando la transcripción como en la caja TATA, por lo tanto minimizando la presencia de este dinucleótido en regiones codificantes, se previene la unión de factores de transcripción en sitios falsos de iniciación (74).

En este trabajo observamos que al comparar el patrón de dinucleótidos entre los diferentes linajes existen diferencias entre los linajes circulantes en Norteamérica, Centroamérica y Centro-Suramérica, probablemente como consecuencia a la presión selectiva por la adaptación en los diferentes nichos ecológicos donde estos circulan.

7.4 Epidemiología Molecular del VEV NJ en México.

Como resultado de los análisis evolutivos realizados en este estudio, estamos proponiendo que el uso de codones es una característica filogenéticamente informativa, la cual puede ser usada para distinguir nichos endémicos de circulación viral así como para explicar el origen endémico que le dio origen a cepas capaces de diseminarse fuera de estas zonas como es el caso del linaje I descrito en este trabajo. La metodología, usada en este trabajo con el virus de estomatitis vesicular, fue usada en el virus de influenza para identificar el origen del hospedador como un método alternativo para entender la evolución de los virus de influenza en el surgimiento de nuevos virus emergentes. En este estudio los autores concluyen que el análisis de la preferencia por el uso de codones, da una nueva perspectiva para examinar la evolución viral. Los estudios de uso de codones, mantienen información sobre estructuras básicas de las secuencias codificantes y en conjunto con los estudios filogenéticos pueden dar una perspectiva más amplia de la evolución de los virus de RNA (75).

En este trabajo, fue posible distinguir 2 nichos ecológicos de circulación viral en México, uno localizado en ciertas áreas del estado de Veracruz, y otro localizado en áreas de los estados de Tabasco, Chiapas y Oaxaca. Entre ellos se distinguen características de uso de codones diferentes, las cuales podrían explicar la adaptación a diferentes nichos ecológicos ya que en el estudio fue posible correlacionar estas zonas de uso de codones con localización geográfica, altitud sobre nivel del mar y zonas ecológicas. Sin embargo y debido a que la única información que se tenía sobre el lugar donde fueron colectados las muestras epiteliales que contenían estos virus, eran estado y municipio y no las coordenadas exactas de latitud y longitud, no fue posible describir de manera más detallada las zonas ecológicas donde circulaban cada uno de estos linajes, con la finalidad de entender con mayor detalle el proceso de adaptación de estos linajes a estas zonas ecológicas y la relación con sus características evolutivas.

Por otro lado se ha descrito que linajes de las zonas endémicas de México son responsables de afectar la parte norte del país, así como causar brotes en la parte sur de los Estados Unidos (25). Sin embargo, se desconoce el origen exacto de estos linajes, nuestros resultados indican que el linaje I descrito en este trabajo, el cual se diseminó de sur a norte

entre el 2006 y 2009 y que adicionalmente en el 2009 muy probablemente fue responsable del brote en los estados de Texas y Nuevo México en los Estados Unidos, durante el mismo año (http://www.aphis.usda.gov/vs/nahss/equine/vsv/vsv2009_final.htm), tuvo su origen de las zonas endémicas en el estado de Veracruz (linaje II), de manera interesante una característica nunca antes descrita, es que a pesar de 4 años de circulación por la zona Centro- Norte de México este linaje parece no modificar su característica intrínseca de uso de codones, lo cual hace posible conocer su origen a pesar de haber escapado de una zona endémica años atrás. Este hecho es similar a lo encontrado en el linaje VI proveniente de Centroamérica, en ambos linajes fue posible encontrar evidencia de selección positiva, aunque este estudio se llevó a cabo únicamente utilizando la región hipervariable de la fosfoproteína del VEV NJ, es posible que otros genes como el de la glicoproteína tengan mayor evidencia de selección positiva y esto pueda explicar en parte su capacidad para diseminarse fuera de su zona endémica de origen. En el caso del virus de encefalitis equina venezolana, se ha demostrado que las cepas endémicas difieren de las aisladas en eventos epidémicos por su nula o baja capacidad para producir viremia en equinos experimentalmente infectados. En encefalitis equina venezolana al igual que en estomatitis vesicular, los estudios filogenéticos concluyen que los linajes endémicos son los precursores de los linajes epidémicos, y que estos no evolucionan de manera independiente. Sin embargo, en encefalitis equina venezolana, se sabe que una mutación en el aminoácido 213 de la glicoproteína es el principal determinante para la producción de viremia en los equinos. (76).

8. Conclusiones.

I.- Este es el primer trabajo realizado en México, explicando la epidemiología molecular de este virus en un periodo determinado, utilizando las características evolutivas del virus como una herramienta para este fin.

II.- Existe una selección positiva de mutaciones silenciosas en el virus de estomatitis vesicular, la cual está moldeando el uso de codones entre los diferentes linajes circulantes en América. Por esto es necesario considerar los estudios de selección positiva midiendo la relación de mutaciones no silenciosas (dN) y silenciosas (dS), no es necesariamente adecuado sea la mejor manera para llevar a cabo análisis en poblaciones naturales del virus de estomatitis vesicular.

III. En condiciones naturales, el patrón de evolución del VEV sugiere una evolución Negativa, con linajes discretos asociados a zonas geográficas específicas, y una ausencia de reloj molecular, donde genotipos encontrados en una determinada zona geográfica 40 años después se mantienen sin cambios considerables. Utilizando los resultados obtenidos en este trabajo, proponemos que la fijación de mutaciones silenciosas permite optimizar el uso de codones, mejorando así en los diferentes linajes del VEV NJ su eficiencia traduccional, como parte de su adaptación a zonas ecológicas definidas y una vez que estos han alcanzado su máximo nivel de adaptación, son mantenidos genéticamente estables por largos periodos de tiempo.

IV.- El uso de codones se correlaciona con la información filogenética en el virus de estomatitis vesicular y esta característica permite predecir nichos ecológicos de circulación viral y el origen de linajes capaces de circular fuera de las zonas endémicas.

9.- Perspectivas.

I.- Analizar el genoma completo del virus de estomatitis vesicular, utilizando virus de estomatitis vesicular New Jersey, representativos de las diferentes regiones endémicas en América, con la finalidad de comprender mejor las características evolutivas de cada gen, y comprobar si los resultados presentados en este trabajo utilizando la región hipervariable del gen de la fosfoproteína son representativos de todo el genoma del virus.

II.- Realizar estudios de co-evolución utilizando secuencias completas del virus de estomatitis vesicular New Jersey y comparar las características evolutivas de estas con las de RNAs mensajeros de los diferentes hospedadores vertebrados e invertebrados los cuales son parte del ciclo infeccioso del virus, con la finalidad de entender el efecto de estos sobre la evolución del virus.

III. Utilizando el análisis de codones sinónimos, establecer el origen de linajes epidémicos que han sido previamente reportados circulando en el norte de México y en los Estados Unidos, con la finalidad de comparar biológica y molecularmente estos linajes con los que les dieron origen en las zonas endémicas y así entender de una mejor manera los mecanismos biológicos que dan origen a la aparición de nuevos brotes epidémicos.

10. Bibliografía.

1. Wagner R.R and Rose J.K.(1996) *Rhabdoviridae*, the viruses and their replication. In fields Virology, 3er edn, PP.1121-1133. Edited by B.N fields , D.M Knipe. Philadelphia, PA: Lippincott-Raven.
- 2.-Goodger W.J., Thurmond M., Hehay J., Mitchell J and Smith P. (1985). Economic Impact of an epizootic of bovine vesicular stomatitis in California. Journal of the American Veterinary Medical Association. **186**, 370-373.
- 3.- Letchworth G.J., Rodriguez L and Barrera J.(1999) Vesicular stomatitis. The veterinary Journal.**157** 239-260.
- 4.- Wunner W.H., Calisher C.H., Dietzgen R.G., Jackson A.O., Kitajima E.W., Lafon M., Leong J.C., Nichol S., Peters D., Smith J.S and Walker P.J.(1995) Rhabdoviridae. Archives of Virology **140** (suppl 10) 275-288.
- 5.- Hanson R.P., Estupian J., Castaneda J.(1968). Vesicular stomatitis in the Americas. Bull. Off. Int. Epizoot. 70, 37-47.
- 6.-Hanson R.P. (1981) Vesicular stomatitis. In: Gibbs, E.P (Ed), Virus diseases of food animals. Academic Press, New York, pp. 517-539.
- 7.- Allende R., Germano P.M. (1993). Comparison of virus neutralization and enzyme-linked immunosorent assay for the identification of antibodies against vesicular stomatitis (Indiana 3) virus. Rev, Sci. Tech. **12**, 849-855.
- 8.-Rodriguez L.L., Bunch T.A., Fraire M., Llewellyn Z.N., (2000). Re –emergence of vesicular stomatitis in the western United States is associated with distinct viral genetic lineages. Virology, **271** 171-181.
- 9.-Rodriguez L.L.(2002). Emergence and re-emergence of vesicular stomatitis in the United States. Virus Research **85**, 211-219.

- 10.-Tesh R.B., Peralta P.H., Johnson K.M. (1969). Ecological studies of vesicular stomatitis virus. Prevalence of infection among animals and humans living in an area of endemic VSV activity. *Am J Epidemiol.* **90**, 225-268.
- 11.-Sudia W.D., Fields B.N., Calisher C.H. (1967). The isolation of vesicular stomatitis virus (Indiana Strain) and other viruses from mosquitoes in New Mexico, 1965. *Am, J Epidemiol* **86**, 598-602.
- 12.- Walton T.E., Webb P.A., Kramer W.L., Smith G.C., Davis T., Holbrook F.R., Moore C.G., Schiefer T.J., Jones R.H., Janney G.C. (1987). Epizootics vesicular stomatitis in colorado, 1982, epidemiologic and entomologic studies. *Am J Trop. Med Hyg*, **36**, 166-176.
- 13.- Mead D.G., Mare C.J., Ramberg F.B (1999). Bite transmission of vesicular stomatitis virus (New Jersey serotype) to laboratory mice by *Simulium vittatum* (Diptera: Simuliidae). *J Med. Entomol.***36**, 410-413.
- 14.- Jimenez A.E., Jimenez C., Castro L., Rodriguez L.L. (1996). Serological survey of small mammals in a vesicular stomatitis virus enzootic area. *J. Wildl. Dis.* **32**, 274-279.
- 15.-Beaty B.J., Borucki M., Farfan J., White D. (1996). Arbovirus-vector interactions: determinants of arbovirus evolution. In Saiuzzo, J.F., Dodet B. (Eds). *Factors in the Emergence of Arbovirus Diseases, Emerging diseases.* Elsvier, Paris, pp 23-35.
- 16.-Howerth E.W., Stallknecht D.E., Dorminy M., Pisell T., Clarke G.R. (1997). Experimental vesicular stomatitis in swine; effects of route inoculation and steroid treatment. *J. Vet Diagn. Invest.* **9**, 136-142.
- 17.-Cornish T.E., Stallknecht D.E., Brown C.C., Seal B.S., Howerth E.W. (2001). Pathogenesis of experimental vesicular stomatitis virus (New Jersey serotype) infection in the deer mouse (*peromyscus maniculatus*). *Vet Pathol.* **38**, 396-406.

- 18.-Mead D.G., Ramberg F.B., Besselsen D.G., Mare C.J. (2000). Transmission of vesicular stomatitis virus from infected to noninfected black flies co-feeding on nonviremic deer mice. *Science* **287**, 485-487.
- 19.- Drake J.W., Holland J.J. (1999) Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. USA* **96**, 13910-13913.
- 20.-Holland J.J., Domingo E., De la Torre J.C., and Steinhauer D.A. (1990). Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *Journal of virology*, **64**, 3960-3962.
- 21.- Novella I.S., Clarke D.K., Quer J., Duarte E.A., Lee C.H., Weaver S.C., Elena S.F., Moya A., Domingo E., and Holland J.J. (1995). Extreme fitness differences in mammalian and insect hosts after continuous replication of vesicular stomatitis virus in sandfly cells. *Journal of Virology*. **69**, 6805-6809.
- 22.-Novella I.S., Quer J., Domingo E., Holland J.J. (1999). Exponential fitness gains of RNA populations are limited by bottleneck effects. *J Virol*. **73**, 1668-1671.
- 23.-Nichol S.T., Rowe J.E., Fitch W.M. (1993). Punctuated equilibrium and positive Darwinian evolution in vesicular stomatitis virus. *Proc.Natl.Acad Sci. USA* **90**,10424-10428.
- 24.- Bisel P.A., Rowe J.E., Fitch W.M and Nichol S.T. (1990) Phosphoprotein and nucleocapsid protein evolution of vesicular stomatitis virus New Jersey. *J Virol* **64**, 2498-2504.
- 25.-Rainwater-Lovett K., Pausek S.J., Kelley W.N., Rodriguez L.L.(2007). Molecular epidemiology of vesicular stomatitis New Jersey virus from 2004-2005 US outbreak indicates a common origin with Mexican strains. *Journal of General Virology*. **88**, 2042-2051.

- 26.-Rodriguez L.L., Vernon S., Morales A.I., Letchworth G.J.(1990). Serological monitoring of vesicular stomatitis New Jersey virus in enzootic regions of Costa Rica. *Am. J. Trop.Med. Hyg.* **42**, 272-281.
- 27.-Vanleeuwen J.A., Rodriguez L.L., Waltner-Toews D. (1995). Cow, farm and ecologic risk factors of clinical vesicular stomatitis on Costa Rica dairy farm. *Am. J. Trop. Med. Hyg.* **53**,342-350.
- 28.-Vernon S.D., Rodriguez L.L., Letchworth G.J. (1990). Vesicular stomatitis New Jersey virus glycoprotein gene sequence and neutralizing epitope stability in an enzootic focus. *Virology* **177**, 209-215.
- 29.- Rodriguez L.L., Fitch W.M., Nichol S.T. (1996) Ecological factors rather than temporal factors dominate the evolution of vesicular stomatitis virus. *Proc. Natl. Acad. Sci.* **93**,13030-13035.
- 30.- Thompson J.D., Higgins D.G., Gibson T.J. (1994) Clustal W; improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position –specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680.
- 31.-Hall T.A. (1999) BioEdit: a user-frendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucl. Acids. Symp. Ser.***41**:95-98.
- 32- Chomczynski P., Sacchi N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-choloform extraction. *Anal Biochem.***1**:159-161.
- 33- Tamura K, Dudley J, Nei M, Kumar S. (2007).MEGA 4: Molecular Evolutionary genetics analysis (MEGA), software 4.0. *Molecular Biology and Evolution.***24**:1596-1599.
- 34.-Saitou M and Nei M (1987), The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution.***4**:406-425.
- 35.-Tamura K and Kumar S (2002). Evolutionary distance estimation under heterogenous substitution pattern among lineages *Molecular Biology and Evolution.* **19**: 1727-1736.

- 36.- Felsenstein J. (1985). Confidence limits on phylogenies; an approach using the bootstrap. *Evolution*. **39**:783-791.
- 37.- Kosakovsky S., Frost S. (2005). No so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*. **22**: 1208-1222.
- 38.- Kosakovsky S, Scheffler K, Gravenor M, Poon A, Frost S. (2009). Evolutionary fingerprint of genes. *Molecular Biology and Evolution*. **27**: 520-536.
- 39.- Kosakovsky S., Frost S.(2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*.**21**:2531-2533.
- 40.-Sharp P.M., Tuohy TM., Mosurski KR. (1986). Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res*. **14**:5125-5143.
- 41.-Zhou T, Gu W, Ma J, Sun X, Lu Z: Analysis of synonymous codon usage in H5N1 virus and other influenza A viruses. *Biosystems*. **81**:77-86.
- 42.-Shields D.C and Sharp P.M. (1989) Evidence that mutation patterns vary among *Drosophila* transposable elements. *Journal of Molecular Biology*. **207**: 843-846.
- 43.- Grantham R., Gautier C., Gouy M., Jacobzone M and Mercier R. (1981). Codon catalog usage is genome strategy modulated for gene expressivity. *Nucleic Acids Research*. **9**: 43-74.
44. Peden J.(2005). Correspondence Analysis of Codon Usage. (<http://codonw.sourceforge.net/>)
- 45.- Karlin S., Doerfler W, Cardon LR. (1994) Why is CpG suppressed in the genomes of virtually all small eukaryotic viruses but not in those of large eukaryotic viruses? *J Virol* **68**: 2889-2897.
- 46.- Powell J.R., Moriyama E.N. (1997). Evolution of codon usage bias in *Drosophila*. *Proc. Nat. Acad. Sci. Usa* **94**: 7784-7790

- 47.- Rima B.K., Neil V.M. (1997). Dinucleotide and stop frequencies in single-stranded RNA viruses. *J of Gen Virol.* **78**: 2859-2870.
- 48.-Jernigan R.W., Baran RH. (2002). Pervasive properties of the genomic signature. *BMC Genomics.***3**:23.
- 49.-Mellor P.S. (2000) Replication of arboviruses in insect vectors. *J. Comp Pathol* **123**: 231-247
- 50.- Toledo V.M., Ordoñez M.J. (2009) Zonas Ecologicas de México. Diagnóstico de los escenarios de la biodiversidad en México. Centro de ecología de la UNAM.
- 51.- Novella I.S., Zárate S., Metzgar D., Ebendick-Corpus B.E. (2004) Positive Selection of Synonymous mutations in vesicular stomatitis virus. *J. Mol. Biol.* **342**:1415-1421.
- 52.-Day D.A., Truite M.F. (1998). Post-transcriptional gene regulatory mechanisms in eukaryotes: an over view. *J. Endocrinol.* **157**: 361-371.
- 53.- Saier M. (1995). Differential codon usage-a safe guard against inappropriate expression of specialized genes. *FEBS letters.***362**:1-4.
- 54.-Lynn D., Singer G., Hickey D.(2002). Synonymous codon usage is subject to selection in thermophilic bacteria. *Nucl. Acids Res.* **30**:4272-4277.
- 55.- Forsdke D.(2002). Selective pressures that increase synonymous mutations in *Plasmodium falciparum*. *Trends Parasitol.* **18**:411-417.
- 56.- Bull J.J., Jacobson A., Badgett M.R., Molineux I.J. (1998). Viral escape from antisense RNA. *Mol. Microbiol.* **28**:835-846.
- 57.- Jenkins G.M., Rambaut A., Pybus O.G., Holmes E.C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J.Mol. Evol.* **54**:152-161.
- 58.-Weaver S.C., Brault A.C., Kang W., Holland J.J. (1999) Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. *J. Virol.* **73**:4361-4326.

- 59.-Woelk C.H., Holmes E.C. (2002). Reduced positive selection in vector-born RNA viruses. *Mol. Biol. Evol.* **19**:2333-2336.
- 60.- Arraigo N.C., Adams P.A., Weaver S.C. (2010). Evolutionary patterns of eastern equine encephalitis virus in north versus south America suggest ecological differences and taxonomic revision. *J. Virol.* **84**:1014-1025.
- 61.- Duret L.(2002). Evolution of synonymous codon usage in metazoans. *Current opinion in Genetics & Development.* **12**:640-649.
- 62.-Karlin S., Blaisdell B.E., Schachtel G.A. (1990). Contrasts in codon usage of latent versus productive genes of Epstein-Barr virus. Data and hypothesis. *J. Virol.* **64**:4264-4273.
- 63.- Zhou J., Liu W.J., Peng S.W., Sun X.Y., Fraizer I. (1999). Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. *J. Virol.* **73**:4972-4982.
- 64.-Coleman J.R., Papamichail D., Skiena S., Fitcher B., Wimmer E., Mueller S. (2008). *Science* **320**: 1784-1787.
- 65.- Bahir I., Fromer M., Prat Y., Linial M. (2009). Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. *Molecular Systems Biology* **5**:311.
- 66.-Kliman R.M., Bernal C.A. (2005) Unusual usage of AGG and TTG codons in humans and their viruses. *Gene.* **352**: 92-99.
67. - Jiang P., Sun X., Lu Z (2007). Analysis of synonymous codon usage in *Aeropyrum pernix* K1 and other Crenarchaeota microorganisms. *J. Genet. Genomics* **34**:275-284.
68. - Mellor P.S (2002). Replication of arboviruses in insect vectors. *J. Comp Pathol.* **123**:231-247.

69. - Bird A.P (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* **8**: 1499-1504.
- 70.-Bird A. (2007) Perceptions of epigenetics. *Nature* **477**:396-398.
71. - Kow YW. (2002). Repair of determinants bases in DNA. *Free Radic Biol Med* **33**:886-893.
- 72.- Lobo F.P., Mota B.E.F., Pena S.D.J., Azevedo V., Macedo A.M. (2009). Virus-host coevolution: Common Patterns of Nucleotide motif usage in Flaviviridae and their hosts. *Plos ONE* **4**:e6282. Doi:10.1371/journal. Pone 0006282.
- 73.- Greenbaum B.D., Levine A.J., Bhanot G., Rabadan R. (2008). Patterns of evolution and host gene mimicry in influenza and other RNA viruses. *PloS Pathog.* **4**: e1000079.
74. - Karlin S., Mrazek J. (1997). Compositional differences within and between eukaryotic genomes. *Proc Natl Acad Sci USA* **94**: 10227-10232.
75. - Wong H.M., Smith D.K, Rabadan R., Peiris M., Poon L.M. (2010). Codon usage bias and the evolution of influenza A viruses. Codon usage biases of influenza virus. *BMC Evolutionary Biology* **10**: 253.
- 76.- Weaver S.C., Barrett D.T. (2004). Transmission cycles, host range, evolution and emergence of arboviral disease. *Nature Reviews.* **2**:789-801.

Anexo 1.

Muestras.

ID= Identificación, MSNM= Metros sobre el nivel del mar,

AC= año de colección.

ID	CIUDAD	MUNICIPIO	MSNM	AC
10	GUERRERO	COYUCA DE CATALAN	210	2007
13	OAXACA	MIXTEPEC	1641	2007
14	ESTADO DE MÉXICO	TLAPALTITLAN	2661	2007
15	CHIAPAS	PICHULALCO	147	2007
16	TABASCO	JALPA	7	2007
17	ESTADO DE MÉXICO	TLAPALTITLAN	2661	2007
19	CHIAPAS	JUAREZ	109	2007
20	JALISCO	ZAPOTLANEJO	1523	2008
21	JALISCO	SAN JUAN DE LOS LAGOS	1741	2008
22	AGUASCALIENTES	AGUASCALIENTES	1873	2008
23	JALISCO	EL SALTO	871	2008
24	JALISCO	TAMAZULA	1373	2008
25	MICHOACAN	CHAVINDA	1571	2008
26	CAMPECHE	CARMEN	0	2008
27	MICHOACAN	ECUANDUREO	1581	2008
28	JALISCO	JALOSTOTITLÁN	1739	2008
29	JALISCO	SAN JULIAN	2064	2008
30	JALISCO	IXTLAHUACAN DEL RIO	1633	2008
31	JALISCO	JILOTLÁN DE LOS DOLORES	853	2008
32	OAXACA	SAN JUAN MAZATLAN	422	2008
33	JALISCO	TUXPAN	1140	2008
34	MICHOACAN	SAHUAYO	1535	2008
35	TABASCO	HUIMANGUILLO	40	2008
36	TABASCO	TACOTALPA	20	2008
37	CAMPECHE	CANDELARIA	40	2008
38	JALISCO	SAN GABRIEL	1540	2008
39	JALISCO	EL LIMÓN	891	2008
40	JALISCO	ZACOALCO DE TORRES	1362	2008
41	COLIMA	TECOMAN	41	2008
42	MICHOACAN	TOMATLAN	321	2008
43	MICHOACAN	HUETAMO	303	2008
44	GUERRERO	CHILPANCINGO DE LOS BRAVO	1104	2008
45	MICHOACAN	APATZINGAN	331	2008

46	JALISCO	LA HUERTA	297	2008
47	GUERRERO	COCULA	622	2008
48	MICHOACAN	TLAZAZALCA	1819	2008
49	CHIAPAS	SUNUAPA	172	2008
50	VERACRUZ	TLALIXCOYAN	10	2008
51	MICHOACAN	TANGANCICUARO	574	2008
52	GUANAJUATO	SILAO	1780	2008
53	MICHOACAN	CHARO	1903	2008
54	CHIAPAS	COAPILLA	1645	2008
55	CHIAPAS	LA CONCORDIA	545	2008
56	MICHOACAN	TARIMBARO	1870	2008
57	CHIAPAS	PUEBLO NUEVO	32	2008
58	VERACRUZ	COSAUTLÁN	1250	2008
59	CHIAPAS	TECPATAN	450	2008
60	MICHOACAN	TANGAMANDAPIO	1682	2008
61	GUERRERO	JUAN R. ESCUDERO	218	2008
62	TABASCO	CUNDUACÁN	15	2008
63	CHIAPAS	COPAINALÁ	441	2008
64	CHIAPAS	JITOTOL	1693	2008
65	MICHOACAN	MORELIA	1916	2008
67	MICHOACAN	ZAMORA	1565	2008
68	TABASCO	H. CARDENAS	27	2008
69	JALISCO	OCOTLÁN	1526	2008
70	JALISCO	TLAQUEPAQUE	1596	2008
72	MICHOACAN	COALCOMAN DE VASQUEZ	1027	2008
73	MICHOACAN	PANINDICUARO	1823	2008
75	JALISCO	MASCOTA	1272	2008
76	JALISCO	TEOCALTICHE	1736	2008
77	MICHOACAN	PENJAMILLO	1782	2008
80	JALISCO	HOSTOTIPAQUILLO	1208	2008
81	JALISCO	ENCARNACIÓN DE DÍAZ	1828	2008
82	JALISCO	VILLA OBREGÓN	534	2008
84	GUANAJUATO	CUERÁMARO	1725	2008
86	JALISCO	ACATLÁN DE JUÁREZ	1372	2008
89	GUERRERO	PILCAYA	1615	2008
91	GUANAJUATO	APASEO EL ALTO	1849	2008
92	MICHOACAN	PERIBÁN	1649	2008
93	MICHOACAN	TEPALCATEPEC	383	2008
94	NAYARIT	SANTA MARÍA DEL ORO	1154	2008
95	QUERÉTARO	HUIMILPAN	2303	2008

101	CHIAPAS	TAPALAPA	1713	2007
102	CHIAPAS	COAPILLA	1645	2007
103	GUERRERO	AZOYU	400	2007
104	GUERRERO	JUAN R ESCUDERO	218	2007
105	GUERRERO	PETATLAN	39	2007
106	GUERRERO	TLACOACHISTLAHUACA	463	2007
107	GUERRERO	ZIRANDARO DE LOS CHAVES	232	2007
108	OAXACA	LOS LLANOS DE AMUZGOS	1627	2007
109	OAXACA	PINOTEPA DE DON LUIS	501	2007
111	OAXACA	STA MARIA ZACATEPEC	388	2007
114	NAYARIT	COMPOSTELA	848	2008
116	CHIAPAS	VILLA FLORES	558	2006
117	GUERRERO	BUENA VISTA DE CUELLAR	1270	2006
118	GUERRERO	IGUALA DE LA INDEPENDENCIA	749	2006
119	GUERRERO	TETIPAC	1747	2006
122	MORELOS	MIACATLÁN	1026	2006
123	PUEBLA			2006
124	TABASCO	HUIMANGUILLO	36	2006
125	VERACRUZ	ALVARADO	10	2006
126	VERACRUZ	BOCA DEL RIO	5	2006
127	VERACRUZ	PLAYA VICENTE	51	2006
128	VERACRUZ	TIERRA BLANCA	64	2006
129	GUERRERO	ACAPULCO DE JUAREZ	84	2007
130	GUERRERO	COYUCA DE BENITEZ	18	2007
131	GUERRERO	COYUCA DE CATALAN	1742	2007
132	GUERRERO	HELIODORO CASTILLO	1533	2007
133	GUERRERO	OMETEPEC	320	2007
134	GUERRERO	TECPAN	119	2007
135	GUERRERO	TLAPEHUALA	670	2007
136	JALISCO	LAGOS DE MORENO	1882	2008
137	TABASCO	HUIMANGUILLO	36	2006
138	VERACRUZ	TUXPAM	21	2006
140	CHIAPAS	MAPASTEPEC	44	2005
141	TABASCO	HUIMANGUILLO	36	2005
142	TABASCO	HUIMANGUILLO	36	2005
143	CHIAPAS	MAPASTEPEC	44	2005
144	TABASCO	EMILIANO ZAPATA	26	2005
145	CHIAPAS	PIJIAPAN	46	2005
146	TABASCO	JONUTA	5	2005
148	TABASCO	CENTRO	7	2005

149	OAXACA	REFORMA DE PINEDA	29	2005
150	CHIAPAS	TUMBALA	1450	2005
201	VERACRUZ	SAN PEDRO SOTEAPAN	430	2009
202	VERACRUZ	JESUS CARRANZA	37	2009
203	CHIHUAHUA	CHIHUAHUA	1430	2009
204	OAXACA	MATIAS ROMERO	198	2009
206	VERACRUZ	SAN PEDRO SOTEAPAN	430	2009
207	SINALOA	CULIACAN	32	2009
208	VERACRUZ	JESUS CARRANZA	37	2009
209	SINALOA	CULIACAN	32	2009
210	SINALOA	CULIACAN	32	2009
211	OAXACA	MATIAS ROMERO	198	2009
212	OAXACA	MATIAS ROMERO	198	2009
213	MICHOACAN	BUENA VISTA TOMATLAN	458	2009
214	VERACRUZ	COSAUTLÁN	1250	2009
215	VERACRUZ	JESUS CARRANZA	37	2009
217	MICHOACAN	BUENA VISTA TOMATLAN	455	2009
218	OAXACA	MATIAS ROMERO	198	2009
219	SINALOA	CULIACAN	32	2009
221	SINALOA	CULIACAN	32	2009
222	SONORA	CAJEMA	24	2009
223	TABASCO	TLACOTALPA	25	2009
227	SINALOA	CULIACAN	426	2009
300	TABASCO	CARDENAS	27	2009
301	VERACRUZ	TRES VALLES	50	2009
302	MORELOS	HUITZUCO DE LOS FIGEROA	1127	2009
303	GUERRERO	HUAMUXTITLAN	990	2009

Anexo 2. Información complementaria de la figura 5.

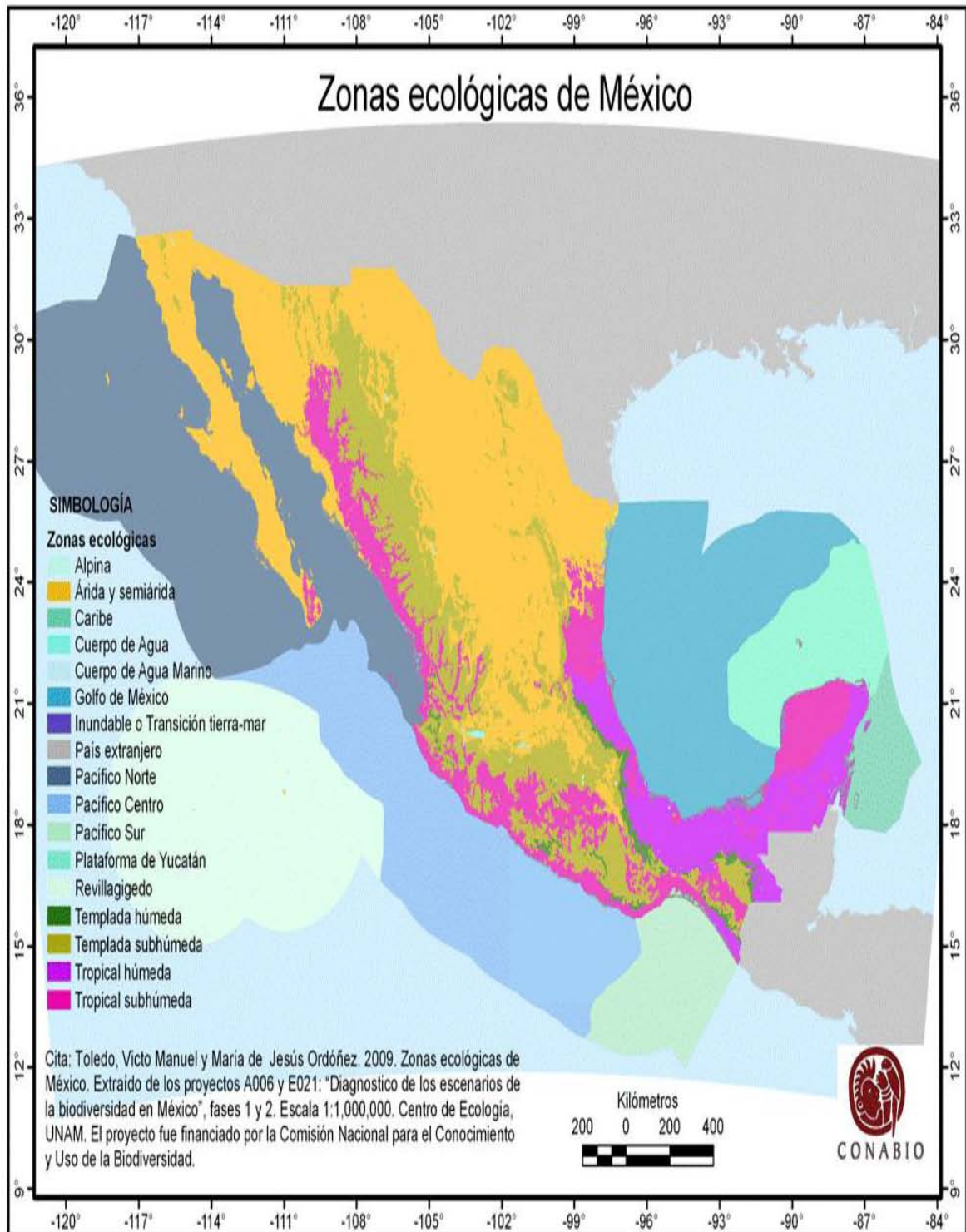
L=linaje, G= genotipo, DG= distribución geográfica, MSNM= metros sobre el nivel del mar, AC= año de colección, ID= identificación.

L	G	DG	MSNM	AC	ID
I	1	AGS, COL, EDO MEX, GTO, GRO JAL, MICH, OAX, SIN	32-2661	2006 AL 2009	14,17,20,22,23,24,27,28,30 40-43, 45,46,48,53,56,60,65 69,72,73,75,77,81,82,84,86 91,93,103,106,107-109,111 117, 131-133, 135, 209, 217, 219, 221, 227.
	2	GTO, JAL	1741-1882	2008	21,52,136
	3	JAL	891-1540	2008	38,39
	4	GRO	622-1615	2008	44,47,89
	5	MICH	574-1649	2008	51,92
	6	GRO	84-218	2007	104,129
	7	GRO,PUE	18-749	2006-2007	118,123,130,134
	8	MICH	1571	2008	25
	9	JAL	1739	2008	28
	10	JAL	1140	2008	33
	11	MICH	1565	2008	67
	12	QRO	2303	2008	95
	13	GRO	218	2008	61
	14	JAL	1596	2008	70
	15	JAL	1208	2008	80
	16	NAY	1154	2008	94
	17	JAL	1736	2008	76
	18	GRO	39	2007	105
	19	NAY	848	2008	114
	20	GRO	1747	2006	119
	21	MOR	1026	2006	122
	22	CHI	1430	2009	203
	23	SIN	32	2009	207
	24	SIN	32	2009	210
	25	MICH	458	2009	213
	26	SON	24	2009	222

II	27	VER, TAB	7--64	2005-2006	125.128,137,148
	28	VER	1250	2008	58
	29	VER	16	2006	126
	30	VER	51	2006	127
	31	VER	21	2006	138
	32	VER	1250	2009	214
III	33	TAB, VER CHP	10-450	2008	35,50,59
	34	CHP, TAB	5-1713	2005-2007	101-146
	35	TAB	26	2005	144
	36	OAX	29	2005	149
IV	37	CHP	109	2007	19
	38	CHP	147	2007	15
	39	TAB	7	2007	16
	40	CHP	1645	2007	102
	41	CHP	558	2006	116
	42	TAB	36	2006	124
	43	CHP	44	2005	140
	44	TAB	36	2005	141
	45	TAB	36	2005	142
	46	CHP	44	2005	143
V	47	CHP	1450	2005	150
	48	VER	1250	2009	231
VI	49	OAX,CAM,CHP,VER TAB	0-1693	2008-2009	218, 26,32,37,62,63,64,201,202, 204,206,208,212,216
	50	TAB	20	2008	36
	51	TAB	25	2009	223
	52	CHP	32	2008	57
	53	TAB	27	2008	68
	54	CHP	1645	2008	54
	55	OAX	198	2009	211
	56	CHP	172	2008	49

Anexo 3.

Zonas ecológicas en México.



Anexo 4.

Definiciones.

Evolutionary Fingerprint: Técnica bioinformática, para de inferir cuales sitios en una secuencia dada están siendo sujetos de selección positiva, negativa.

Frecuencia de dinucleótidos: La distribución de la frecuencia de dinucleótidos ha sido reconocida como un buen estimador para determinar presiones selectivas, así mismo se ha asociado con influenciar el uso de codones . Los valores obtenidos son usualmente diversos entre especies y muy invariantes a lo largo de un genoma individual, por estas razones el patrón de dinucleótidos de una secuencia es referido como su firma genética

Genotipo: El genotipo es la secuencia genética específica de un individuo en este caso un virus.

Hospedador: En Biología se llama hospedador, hospedero u hospedante, a aquel organismo que alberga a otro en su interior o lo porta sobre sí.

Linaje: En que tienen la misma línea de descendencia.

Linaje monofilético: un grupo es monofilético (del griego: de una rama) si todos los organismos incluidos en él han evolucionado a partir de una población ancestral común, y todos los descendientes de ese ancestro están incluidos en el grupo.

Mutación no sinónima- . Es aquella en la que se produce un cambio en un nucleótido de la cadena de ADN produciendo un cambio en el codon correspondiente que trae como consecuencia un cambio en la cadena polipeptídica.

Mutación sinónima: Es aquella en la que se produce un cambio en un nucleótido de la cadena de ADN produciendo un cambio en el codón, y que no cambia la secuencia de la proteína , debido a que el código genético es degenerado, es decir varios codones pueden codificar para el mismo aminoácido.

Nicho Ecológico- El concepto formal de nicho incluye a todos los factores bióticos y abióticos con los cuales el organismo se relaciona. Formalmente, el nicho ha sido descrito como un hipervolumen de n-dimensiones, donde cada dimensión corresponde a los factores antes descritos. De esta forma, el nicho involucra a todos los recursos presentes del

ambiente, las adaptaciones del organismo a estudiar y cómo se relacionan estos dos (nivel de adaptación, eficiencia de consumo, etc.)

Número efectivo de codones (ENC): Este método asigna a un gen un número entre 20 y 61, que indica cual es el grado en el que el total de codones sinónimos es usado. Un valor de 20 indica que solo un codón es usado para la síntesis de cada aminoácido, mientras que un valor de 61 indica que todos los codones sinónimos son usados de manera equitativa.

Reloj Molecular: La hipótesis del reloj molecular, propone que la fijación de las mutaciones ocurre a una velocidad constante a través del tiempo. Por lo tanto el número de sustituciones entre 2 especies, es equivalente al tiempo que ha pasado desde que ellos compartieron un ancestro en común.

Selección Negativa: Selección que elimina nuevas mutaciones de la población, debido a su efecto deletéreo, por lo tanto remueve cambios de una población y la mantiene estable.

Evolución Neutral: Son los cambios en los genes de una especie que son el resultado de mutaciones al azar las cuales no tienen ningún efecto en la adecuación de los individuos.

Selección Positiva: Promueve la fijación de una nueva mutación que aumenta la adecuación de los individuos.

Uso de codones: La frecuencia relativa a la cual los codones alternativos para sintetizar un aminoácido en particular son usados.