

EPIDEMIOLOGÍA, VIROLOGÍA Y GENÓMICA DE LA NUEVA VARIANTE DEL VIRUS DE INFLUENZA A (H1N1)

Elizabeth E. Godoy-Lozano,¹ Christian A. García-Sepúlveda,² Daniel E. Noyola¹

¹ Laboratorio de Virología, Departamento de Microbiología, y

² Laboratorio de Genómica Viral y Humana,

Facultad de Medicina, Universidad Autónoma de San Luis Potosí,

Título corto: Caracterización del virus de influenza A (H1N1) 2009

Autor responsable de la correspondencia: Daniel E. Noyola, MD. Laboratorio de Virología, Departamento de Microbiología, Facultad de Medicina, Universidad Autónoma de San Luis Potosí. Avenida Venustiano Carranza 2405, San Luis Potosí, S.L.P., México, 78210. Tel y fax: (52) 444-826-2346, ext. 534. E-mail: dnoyola@uaslp.mx

Palabras Clave: Virus de influenza A (H1N1), pandemia, caracterización genómica, enfermedades infecciosas emergentes.

RESUMEN

Las infecciones respiratorias agudas son causa importante de morbilidad y mortalidad en todo el mundo siendo los niños menores de cinco años los más afectados. Entre los agentes que causan estas infecciones se encuentran bacterias y virus, siendo estos últimos su principal causa. El virus de influenza A es de los virus de mayor importancia debido a que causa enfermedades graves en tres a cinco millones de personas y entre 250,000 y 500,000 muertes por año a nivel mundial. La reciente pandemia causada por el virus de influenza A (H1N1) 2009 representó un grave problema de salud pública en México y el mundo. Es por tal motivo que nuestro objetivo principal fue definir las características epidemiológicas, virológicas y genómicas de cepas de la nueva variante pandémica del virus de influenza A (H1N1). Nuestro estudio nos permitió inferir la fecha de aparición del ancestro común más reciente de la cepa alrededor del 8 de julio del 2008. Adicionalmente encontramos que la tasa de mutación global es mucho menor que la reportada anteriormente para esta cepa. Entre las 10 cepas estudiadas no encontramos la presencia de mutaciones asociadas a resistencia a oseltamivir. No obstante, nuestro análisis permitió identificar la existencia de eventos recombinatoriales entre cepas de influenza con distintos orígenes geográficos e incluso hospederos. Lo cual brinda más información sobre la evolución y el origen de esta cepa pandémica.

INTRODUCCIÓN

Las infecciones respiratorias agudas (IRAS) son una de las principales causas de morbilidad y mortalidad en todo el mundo. Para niños menores de cinco años son la segunda causa de muerte en el mundo y en México son la tercera causa de mortalidad (Rudan et al., 2004; SSA, 2009). En México la tasa de mortalidad por IRAS en el 2005 fue de 18.8 por 100,000 habitantes incluyendo aquellas causadas por influenza y neumonía. Los niños menores de 5 años son los más afectados por estos padecimientos; en el año 2005 la tasa de mortalidad por IRAS a nivel nacional fue de 33.5 por 100,000 habitantes (SSA, 2008).

Entre los agentes que causan IRAS se incluye, entre otros, a las bacterias y a los virus. Estos últimos son la principal causa de estas infecciones ya que tienen la capacidad de infectar y de replicarse en el epitelio respiratorio con mucha facilidad. Entre los virus que causan IRAS más frecuentes se encuentran los virus parainfluenza 1-4, rinovirus, virus sincicial respiratorio y virus de la influenza A y B (Bellos et al., 2010; Maffey, 2008). Los virus de influenza se consideran los de mayor relevancia para la salud pública puesto que la Organización Mundial de la Salud (OMS) ha declarado que durante las epidemias anuales de influenza éstos afectan entre el 5 y el 15% de la población; además, hay de tres a cinco millones de casos graves y entre 250,000 y 500,000 muertes por año a nivel mundial (Taubenberger et al., 2010; Thompson et al., 2003).

Los influenzavirus pertenecen a la familia Orthomyxoviridae que se ubican en el Grupo V según la clasificación de Baltimore (Baltimore, 1971). Su genoma se encuentra compuesto por ocho segmentos de RNA monocatenario de sentido negativo. Cada segmento del genoma codifica para una proteína específica, excepto los segmentos 2(PB1), 7(MP) y 8(NS) que poseen marcos de lectura alternativos para codificar cada uno a dos proteínas

estructural y funcionalmente distintas (Viruses, 2006; Brown, 2000; Conenello, 2007; Swedish, 2010).

Una de las características principales de los virus de influenza es que tienen una gran diversidad debido a la gran cantidad de hospederos en los que circula el virus. Los virus de influenza A se clasifican de acuerdo a las proteínas de superficie,, en 16 subtipos según la hemaglutinina (HA) y 9 subtipos de acuerdo a la neuraminidasa (NA). Los subtipos que afectan al humano son los H1, H2, H3, N1 y N2 (Nicholson et al., 2003). Por otro lado existe variabilidad genética entre los mismos subtipos, generada por dos mecanismos: drift y shift antigénico. La deriva o “drift” antigénico es un cambio menor en los antígenos de superficie que resulta de la acumulación de mutaciones puntuales. Estos cambios contribuyen a la aparición de nuevas cepas responsables de epidemias anuales estacionales de influenza. El drift antigénico resulta en variantes virales para las cuales hay protección incompleta en la población debido a su exposición previa a virus similares. En cambio el desplazamiento antigénico o “shift” antigénico es una modificación mayor en uno o ambos antígenos de superficie (HA o NA) que ocurre a intervalos variables. Estos cambios se deben a un re-arreglo de fragmentos genómicos entre diferentes virus de influenza tanto humanos como de otras especies. Un desplazamiento antigénico puede ocasionar una pandemia si el virus se transmite en forma eficiente de persona a persona (Al Hajjar et al., 2010; Arias et al., 2009). El re-arreglo de genes ocurre comunmente en los cerdos, puesto que a diferencia de otros hospederos, los cerdos tienen la capacidad de hospedar a virus de influenza aviar, porcina y humana. Esta susceptibilidad a la infección por distintas cepas de influenza se debe a la presencia del receptor de ácido siálico tanto en posición $\alpha 2,3$ (empleado por el virus aviar) como en posición $\alpha 2,6$ (empleado por el virus humano) (Salomon et al., 2009).

En el mes de abril del 2009 se identificó una nueva variante del virus de influenza como causa de infección respiratoria en el humano; la nueva variante del virus de influenza se extendió por todo el mundo en muy poco tiempo (Boelle et al., 2009; Fraser et al., 2009). Este nuevo virus, a diferencia del virus de influenza estacional, causó infecciones atípicas y severas en adultos jóvenes. Debido a que el virus se transmitió fácilmente de persona a persona, de su rápida diseminación de un país a otro de la misma región y a que ulteriormente circuló en todo el mundo, en junio del 2009 la OMS elevó el nivel de alerta pandémica de influenza de fase 5 a 6, designándola como la primera pandemia de influenza del siglo XXI. A la fecha este nuevo virus se ha propagado en más de 214 países, causando al menos 18,366 muertes confirmadas (WHO, 2009). Cabe mencionar que México fue el primer país en donde se registró la diseminación extendida de este nuevo virus, siendo San Luis Potosí uno de los estados en donde el virus fue identificado inicialmente. Debido a esto nos hemos propuesto definir las características epidemiológicas, virológicas y genómicas de cepas de la nueva variante del virus de influenza A (H1N1) responsable de la epidemia de infecciones respiratorias ocurrida en la ciudad de San Luis Potosí.

MATERIALES Y MÉTODOS

Muestras. Las muestras fueron tomadas de pacientes con sospecha de infección por virus de influenza obtenidas a través del sistema de vigilancia de infecciones respiratorias del Hospital Central “Dr. Ignacio Morones Prieto” y del Laboratorio de Virología de la UASLP en la ciudad de San Luis Potosí (n=706). Estas muestras corresponden a secreciones respiratorias del tracto respiratorio (exudado faríngeo, exudado nasofaríngeo y lavado nasofaríngeo) de pacientes con infección respiratoria aguda. Una vez obtenidas, las muestras fueron transportadas al laboratorio de Virología de la UASLP, a 4°C en solución

salina isotónica o en medio de transporte viral; al llegar al laboratorio se les agregó solución de antibióticos y fueron almacenadas a -70°C hasta su procesamiento.

Extracción de RNA viral. El RNA viral fue extraído empleando reactivos comerciales (High Pure Viral RNA Kit; Roche Diagnostics, Mannheim, Germany) de acuerdo a las especificaciones del fabricante y almacenado posteriormente a -80°C .

Síntesis de cDNA. El RNA extraído se sometió a transcripción inversa (RT) empleando 2 μL de RNA, 2.25 μM del oligonucleótido UniFlu-RT (5'-AGG-AAA-AGC-AGG-3') y 50 UI totales de la transcriptasa inversa M-MuLV a un volumen final de 10 μL sometiéndola a un ciclo de 60 minutos a 38°C (Hoffmann et al., 2001). El cDNA obtenido se almacenó a -80°C .

Detección de influenza A (H1N1) pandémica. La técnica utilizada para la detección de influenza A (H1N1) es una reacción en cadena de la polimerasa (PCR) anidada que hace uso de oligonucleótidos previamente descritos (Gomez-Gomez et al., 2009). Para la primera PCR, se tomaron 2 μL de cDNA previamente sintetizado como se mostró anteriormente, 1.6 μM de los oligonucleótidos SwNS-F 5'-ATG-GAC-TCC-AAC-ACC-3' y SwNS-R 5'-TTA-AAT-AAG-CTG-AAA-CGA-G-3', 0.05 UI totales de Taq (Vivantis Technologies Sdn Bhd Selangor D., Malaysia), 200 μM de dNTP's y 1.5 mM de MgCl_2 , llevados a un volumen final de 10 μL . Se utilizó el siguiente programa de termociclaje: una desnaturalización inicial de 95°C por 5 minutos, seguidos de 30 ciclos de 95°C por 20 segundos, 56°C por 30 segundos y 72°C por 90 segundos, seguidos de una extensión final de 72°C por 5 minutos. El producto de la primera PCR se diluyó 1:1 empleando 1 μL de esta dilución para la 2da PCR utilizando los siguientes oligonucleótidos: Sw1-F 5'-CTT-GAA-AGA-GGA-ATC-GAG-CG-3' y Sw1-R 5'-GTC-TCC-CAT-TCT-CAT-CAC-AGT-3' a una concentración de 800 nM; el resto de los componentes de la mezcla de reacción

fueron los mismos antes mencionados a excepción de que la concentración del MgCl₂ fue de 1 mM. La 2nda PCR genera un amplicón de 429 bp. Las condiciones de termociclaje fueron las mismas que en la primera PCR a excepción de que se aumentó el número de ciclos a 40. Los productos de amplificación de la segunda PCR fueron sometidos a electroforesis en gel de agarosa al 1.5% teñido con bromuro de etidio a 4.5 volts/cm. Posteriormente fueron observados bajo luz UV por medio del fotodocumentador y las imágenes almacenadas digitalmente (Biodoc-it, UVP, Cambridge, UK).

Diseño de oligonucleótidos para la caracterización genómica. Se realizó el alineamiento de las secuencias de la nueva variante del virus de influenza A (H1N1) depositadas en la base de datos “Influenza Virus Resource” del National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>) utilizando el programa Clustal W versión 1.83 (<http://www.ebi.ac.uk/Tools/clustalw2/>) para cada uno de los 8 segmentos del genoma viral. Posteriormente se realizó el reformato a unanimidad por medio de una aplicación en línea localizada en <https://www.alleles.org/internal/common/reformat.html> para resaltar las zonas conservadas de los alineamientos y diseñar los oligonucleótidos. Se analizaron las características de las secuencias de los oligonucleótidos deseadas incluyendo su especificidad, la presencia de hairpins, homodímeros y heterodímeros por medio de la herramienta en línea OligoAnalyzer de Integrated DNA Technologies (<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>). Para todos los segmentos se diseñaron oligonucleótidos sentido y antisentido que flanquean al gen completo (Tabla 1). Para los segmentos HA, PA, PB1 y PB2 se diseñaron 4 pares de oligonucleótidos con la finalidad de obtener secuencias cortas (800 bp) de buena calidad. Así mismo, para los segmentos NS y NP se utilizaron 2 y 3 pares, respectivamente, de los oligonucleótidos

diseñados y publicados por los Centers for Disease Control and Prevention (CDC) (Tabla 2).

Amplificación de fragmentos genómicos completos. Se utilizaron 2 μL de cDNA sintetizado como se menciona previamente, 1.6 μM de cada oligonucleótido (Tabla 1), 200 μM de dNTP's, 1.5 mM MgCl_2 y 0.5 UI totales de una mezcla 9:1 de Taq:Pfu (Vivantis Technologies Sdn Bhd Selangor D., Malasia) a un volumen final de 12.5 μL . Se utilizó el siguiente programa de termociclaje: desnaturalización inicial de 95°C por 5 minutos, seguida de 35 ciclos de 95°C por 20 segundos, la temperatura de alineamiento correspondiente a cada segmento por 30 segundos y 72°C por 90 segundos, seguidos de una extensión final de 72°C por 5 minutos. El producto de PCR fue almacenado a -4°C.

Amplificación de fragmentos subgenómicos. Se realizó una dilución 1:10 de la PCR de los fragmentos genómicos completos, a excepción del segmento de HA para el cual la dilución fue de 1:16, y del fragmento 2 de PA para el cual fue de 1:5. Posteriormente se realizó una segunda PCR con 1 μL del producto de PCR diluido (a excepción del fragmento 4 de PA que se colocó 1 μL directo de la primera PCR), 200 μM de dNTP's, 0.05 UI totales de la mezcla Taq:Pfu relación 9:1 (Vivantis Technologies Sdn Bhd Selangor D., Malasia) a la concentración de MgCl_2 correspondiente para cada fragmento. El programa de termociclaje para los fragmentos que corresponden a HA, PA, PB1 y PB2 inicia con una desnaturalización inicial de 95°C por 5 minutos seguido de 35 ciclos de 95°C por 30 segundos, la temperatura de alineamiento correspondiente para cada fragmento (Tabla 1) por 30 segundos y 72°C por 90 segundos, finalizando con un tiempo de extensión de 72°C por 5 minutos. El programa de termociclaje así como las condiciones de la PCR para los fragmentos NP y NS es el descrito por los CDC (CDC, 2009). Para confirmar que se produjo la amplificación de los fragmentos de interés, los productos de PCR se corrieron en

un gel de agarosa al 1.5% teñido con Bromuro de Etidio por electroforesis a 150 voltios por 40 minutos, observados bajo luz UV por medio del fotodocumentador y las imágenes fueron almacenadas digitalmente (Biodoc-it, UVP, Cambridge, UK).

Secuenciación de DNA. Los productos de PCR se enviaron al Laboratorio Nacional de Genómica para la Biodiversidad del CINVESTAV-Irapuato para su secuenciación. Los electroferogramas obtenidos fueron evaluados y editados con la aplicación 4peaks versión 1.7 (<http://mekentosj.com/science/4peaks/>). Para validar los polimorfismos encontrados, las secuencias obtenidas fueron comparadas con las secuencias depositadas en la base de datos “Influenza Virus Resource” de NCBI.

Análisis bioinformático. Se efectuó un análisis de las secuencias de influenza generadas localmente (n=17) junto con las secuencias de virus de influenza mexicanos previamente depositadas en la base de datos “Influenza Virus Resource” de NCBI (n=159). El análisis de homología global (de longitud completa) de cada segmento genómico completo se realizó por medio de la aplicación en línea Basic Local Alignment Search Tool (www.ebi.ac.uk/Tools/blast/). El análisis de las substituciones no sinónimas con respecto a las substituciones sinónimas (dN/dS) para cada una de las proteínas codificadas por el virus se realizó empleando el algoritmo “Codon-based Z-test of selection” y “Position-wise codon based selection estimation (HyPhy)” del paquete de análisis MEGA 5 (Kumar et al., 2008). El análisis de la homología regional así como el análisis de las secuencias de influenza A (H1N1) pandémica mexicanas se realizó con la aplicación SimPlot versión 3.5.1 (Lole et al., 1999). El análisis de Bootscan utilizó un tamaño de ventana de 400 bases en los segmentos 1(PB2), 2(PB1), 3(PA), 6(NA) y de 200 bases en los segmentos 4(HA), 5(NP), 7(MP) y 8(NS); con pasos de 20 bases, 100 repeticiones, con el modelo Kimura (2-parameter), T/t:2.0, Neighbor-Joining. El modelo evolutivo basado en el análisis de la

distribución de los polimorfismos se basó en el algoritmo FindModel disponible en línea en la base de datos de HIV del Laboratorio Nacional de Los Álamos (www.hiv.lanl.gov) utilizando el set de modelos reducidos. El análisis coalescente de inferencia Bayesiana en donde se calculó la tasa de mutación global de cada proteína codificada por el virus, la incidencia de mutaciones por posición codónica, al igual que los árboles filogenéticos y el cálculo del ancestro común más reciente (MRCA por sus siglas en inglés) se realizó empleando las aplicaciones Bayesian Evolutionary Analysis Utility (BEAUti) v1.5, Bayesian Evolutionary Analysis by Sampling Trees (BEAST) v1.5.4, Tracer v1.5, TreeAnnotator v1.5.4 y FigTree v1.3.1 (Drummond et al., 2007). El tamaño de la cadena utilizado para el cálculo del MRCA fue de 50 millones para las proteínas PB2, PB1, PA y NA; 40 millones para HA y NP; 20 millones para M1, M2, NS1 y NS2. Las secuencias que se utilizaron para los análisis de dN/dS, tasa de mutación global de cada proteína codificada por el virus, de la incidencia de mutaciones por posición codónica, al igual que los árboles filogenéticos y el cálculo del MRCA fueron las secuencias nucleotídicas mexicanas de tamaño completo que se encuentran depositadas en la base de datos “Influenza Virus Resource” de NCBI en conjunto con las secuencias generadas en este estudio. Las secuencias consenso para el análisis de Simplot se generaron en la base de datos “Influenza Virus Resource” de NCBI con secuencias de tamaño completo de cada uno de los segmentos de virus aislados en todo el mundo y de diversos orígenes (humano, aviar, porcino) depositadas en esta misma base de datos.

RESULTADOS

Durante el periodo comprendido entre el 1º de marzo de 2009 y el 9 de abril de 2010 se evaluaron 706 muestras para la detección de influenza pandémica A(H1N1) 2009. Las

muestras fueron referidas a través del sistema de vigilancia de virus respiratorios del Hospital Central “Dr. Ignacio Morones Prieto” y de la Facultad de Medicina de la UASLP. Se detectó al virus de influenza pandémica en 133 (18.8%) de las 706 muestras analizadas. En la Figura 1 se muestra el número de muestras evaluadas y el número de casos de influenza detectados en cada una de las semanas epidemiológicas del periodo incluido en este análisis. Se seleccionaron 17 muestras positivas para el virus de influenza distribuidas a lo largo de todo el año de estudio para realizar el análisis genómico. Se generó información de secuencia nucleotídica de alta resolución para 11 aislados clínicos y para los segmentos 1(PB1), 2(PB2), 3(PA) y 5(NP); 10 para el segmento 6(NA); 9 para los segmentos 4(HA) y 7(MP); y 7 para el segmento 8(NS), todas de secuencia genómica completa. De la información generada se obtuvieron 2 genomas completos, 7 parciales [de los cuales solo faltó el segmento 8(NS)] y 8 parciales (de los cuales faltaba más de un segmento). Se obtuvieron amplicones de buena resolución tanto para los segmentos que fueron amplificados completamente en una sola reacción [segmento 6(NA) y 7(MP)], como para aquellos que fueron amplificados en diversos fragmentos [segmentos 1(PB1), 2(PB2), 3(PA), 5(NP), 4(HA) y 8(NS)]. El porcentaje de éxito de la amplificación genómica completa fue aproximadamente del 87.5%. Cuando se realizó el empalmamiento de fragmentos subgenómicos para la construcción del contig correspondiente a la secuencia completa no se observaron posiciones ambiguas en ninguna de las secuencias. Las secuencias obtenidas mostraron un 99% de similitud entre sí para cada uno de los segmentos analizados. Así mismo el análisis de homología basado en BLAST nucleotídico reveló una similitud del 99% en relación a secuencias del virus de influenza A (H1N1) pandémica de todo el mundo aisladas durante el año 2009.

El análisis de la relación dN/dS para cada una de las proteínas codificadas reveló un predominio de sustituciones sinónimas. Este hallazgo se verificó al encontrar valores estadísticamente significativos en apoyo a la hipótesis de selección purificante ($dN < dS$) para la mayoría de los genes empleando el test Global Z (proteína PB2 $p=0.001$, PB1 $p=0.00003$, PA $p=0.004$, HA $p=0.0004$, NP $p=0.0001$, NA $p=0.04$ y M1 $p=0.008$). Para las proteínas M2, NS1 y NS2 no se obtuvieron valores significativos.

El análisis de similitud realizado a cada segmento reveló una similitud superior al 99% entre las secuencias locales. Este análisis se utilizó para comparar las secuencias locales contra aquellas derivadas de diferentes hospederos. Con el análisis se corroboró la diversidad del origen de los distintos segmentos génicos al demostrar mayor similitud con secuencias de influenza porcina para los segmentos 1(PB2), 4(HA), 5(NP), y 8(NS), mientras que los segmentos 3(PA), 6(NA) y 7(MP) demostraron mayor similitud con secuencias de influenza aviar y el segmento 2(PB1) demostró mayor similitud con secuencias de influenza humana.

Adicionalmente para los segmentos 4(HA) y 6(NA) se compararon las secuencias obtenidas con aquellas reportadas para los diversos subtipos virales. Tomando en cuenta los resultados previos, este análisis se limitó a las secuencias de influenza porcina y aviar para los genes HA y NA, respectivamente. Para el segmento 4(HA) encontramos mayor similitud con secuencias del subtipo H1 (91.8%) que con las secuencias de influenza porcina del subtipo H2 (66.6%), H5 (64.6%), H9 (35.6%), H4 (28.2%) y H3 (18.8%). Por otro lado, para el segmento 6(NA) encontramos mayor similitud con el subtipo N1 (84.8%) que con secuencias de influenza aviar del subtipo N4 (60.0%), N3 (36.3%), N5 (36.3%), N9 (36.3%), N8 (23.2%), N2 (20.6%), N7 (17.4%) y N6 (13.3%).

Para analizar el origen geográfico de las secuencias de H1 del segmento 4(HA) y N1 del segmento 6(NA) se utilizaron secuencias consenso de influenza porcina y aviares por regiones, respectivamente. Para el análisis de los segmentos 1(PB2), 5(NP), y 8(NS) se utilizaron secuencias consenso por regiones de origen porcino; para los segmentos 3(PA) y 7(MP) se emplearon secuencias consenso de influenza aviar. En el caso del segmento 2(PB1), el análisis se efectuó con el uso de secuencias de influenza de origen humano. En la Tabla 3 se muestra el porcentaje de similitud entre las secuencias incluidas en nuestro estudio y las secuencias consenso de influenza porcina, aviar y humana para cada uno de los segmentos génicos del virus; así como la similitud por región geográfica. Una vez que identificamos las secuencias por región geográfica con mayor similitud con las secuencias locales, se analizaron con la estrategia Bootscan de ventana desplazante para evaluar la existencia de eventos de recombinación (Figura 2). Este análisis reveló para el segmento 1(PB2) la existencia de un evento de recombinación entre las bases 876-1124 en el cual se identificó una mayor similitud entre las secuencias de influenza A(H1N1) 2009 con secuencias porcinas provenientes de Asia, mientras que el resto de la secuencia mostró mayor similitud con secuencias de Norte América. Para el segmento 2(PB1) se observaron vestigios de recombinación en las bases 568 a 726 y de las bases 1704 a 1819 entre secuencias de Asia y Norte América así como de Asia y Oceanía respectivamente; demostrando que entre estas secuencias existió una recombinación, pero debido a la acumulación de mutaciones estos segmentos se han ido alejando paulatinamente a través del tiempo. Al analizar el segmento 3(PA) se observó la presencia de un evento de recombinación en las primeras bases en donde la secuencia de la base 1 a la 277 era más similar a secuencias originarias de Europa, mientras que entre las bases 1014 a 1170 se encontró otro evento de recombinación en donde la secuencia era más similar a secuencias

de Oceanía, el resto de la secuencia era más similar a secuencias de Norte América. Para el segmento 4(HA) se observó la presencia de un evento de recombinación en donde las primeras 300 bases eran más similares a secuencias de Europa, mientras que el resto de la secuencia mostró mayor similitud con cepas norteamericanas. El segmento 5(NP) solo manifestó vestigios de eventos de recombinación entre las bases 431 a 484 y 836 a 967 entre secuencias norteamericanas y europeas. Para el segmento 6(NA) se encontró mayor similitud en las primeras 253 bases con secuencias de Oceanía y se observó un evento de recombinación entre las bases 1330 a 1349 donde la secuencia era más similar a secuencias de Sudamérica, así como vestigios de eventos de recombinación entre las bases 704 a 820 con mayor similitud a secuencias de Oceanía y para las bases 1199 a 1246 con secuencias de Sudamérica. Al analizar el segmento 7(MP) se encontró un evento de recombinación entre las bases 196 a 225 en donde esta parte de la secuencia era más similar a secuencias de Oceanía y en las últimas 177 bases la similitud era mayor con secuencias de Europa, el resto de la secuencia tuvo mayor similitud con secuencias de origen asiático. Finalmente el segmento 8(NS) no reveló ningún evento de recombinación.

Los resultados al determinar el mejor modelo evolutivo de acuerdo con el nivel de Akaike Information Criteria (AIC) indican que para las proteínas PB2, PB1, PA, HA, NA, M1 y M2 fue Hasegawa-Kishino-Yano (HKY) (PB2 AIC=7742.22 PB1 AIC=7012.52, PA AIC=7012.52, HA AIC=6635.83, NA AIC= 4919.87, M1 AIC=2379.96 y M2 AIC=948.39), para las proteínas NP y NS2 fue Tamura-Nei (TrN) (NP AIC=4901.85, NS2 AIC=1158.08) y para la proteína NS1 fue General Time Reversible (GTR) (NS1 AIC = 2134.81).

El cálculo del MRCA indicó que la proteína con el MRCA más antiguo fue M1 y la proteína de más reciente aparición fue NS2. Las tasas de mutación global encontradas

concuerdan con estos resultados, pues la tasa de mutación global más baja es la de la proteína de menor tiempo de evolución (NS2), en cambio la tasa de mutación global más alta es la de la proteína con mayor tiempo de evolución (M1). Los resultados del análisis del MRCA y la tasa de mutación global de cada proteína se muestran en la Tabla 4. Los datos obtenidos del análisis de la incidencia de mutaciones por posición codónica de todas las proteínas a excepción de NS1 nos indicaron que la tercera posición tiene una tasa de mutación superior a la de la primera y segunda posición, lo que es indicativo del mecanismo de conservación de las proteínas. La proteína NS1 mostró la misma tasa de mutación en cada una de las posiciones codónicas, lo que es consistente con la hipótesis de acumulación neutra de mutaciones.

Los árboles filogenéticos generados fueron árboles de cladas de máxima credibilidad, en donde la raíz del árbol es el MRCA y las distancias evolutivas del eje horizontal se indican en días. Como podemos observar las secuencias locales se encuentran ubicadas en la parte superior de los árboles puesto que son los aislados más recientes. En la Figura 3 observamos los árboles correspondientes a las secuencias nucleotídicas que codifican para las proteínas de mayor longitud virus, en donde se puede visualizar que el gen que codifica para la proteína PA tiene mayor tiempo de evolución (aproximadamente dos meses), en relación con PB2, PB1 y NP. En la figura 4 se observan los árboles correspondientes a las secuencias nucleotídicas que codifican para los antígenos de superficie, NA y HA. En la Figura 5 se observan los árboles correspondientes a las secuencias nucleotídicas que codifican para las proteínas de menor longitud del virus, las proteínas que se leen con marcos de lectura alternativos; puede observarse que las proteínas M2, NS1 y NS2 tienen un menor tiempo de evolución en comparación con la proteína M1 (aproximadamente ocho meses). Los árboles nos muestran que la mayoría de las proteínas, a excepción de M1,

tuvieron al momento del brote inicial una gran explosión de diversidad genética, misma que con el paso del tiempo ha ido evolucionando hacia la conservación de proteínas. En cambio la proteína M1 aun permanece en la fase de expansión.

Al analizar las secuencias del segmento 6(NA) no se encontró la presencia de mutaciones en la posición 274 ni en la 119, que confieren resistencia a oseltamivir. Tampoco se encontró la mutación en la posición 151 que confiere resistencia a zanamivir.

DISCUSIÓN

El estado de San Luis Potosí es el tercer estado de la república Mexicana con el mayor número de casos confirmados por influenza A (H1N1) pandémica. Del total de casos confirmados, San Luis Potosí contribuyó con el 6% de los infectados de todo el país (SSA, 2010). Los datos obtenidos en este estudio solo representan una pequeña parte de las infecciones por influenza A (H1N1) pandémica durante el periodo del 1° de marzo de 2009 y el 9 de abril de 2010 en México. Sin embargo, nuestros resultados muestran el desarrollo de las dos primeras oleadas de influenza pandémica en nuestra región. Asimismo, el muestreo continuo nos permitió contar con muestras representativas del primer año de la pandemia.

Uno de los objetivos que nos planteamos fue el de visualizar el origen geográfico de nuestras secuencias; esto debido a que México (principalmente el estado de San Luis Potosí) fue uno de los primeros países en reportar la diseminación del virus. Existen varios estudios previos que realizaron un análisis con ventanas temporales más cortas. El origen de los segmentos de influenza A (H1N1) pandémica según varios autores, proviene principalmente de influenza porcina tanto de Norte América como de Asia y Europa (Babakir-Mina et al., 2009; Chang et al., 2009; Dawood et al., 2009; Garten et al., 2009).

Estos datos corroboran con lo obtenido en los segmentos 1(PB2), 4(HA), 5(NP), y 8(NS). Smith et al., 2009b realizaron un estudio más a fondo sobre la caracterización genómica temprana del virus de influenza A(H1N1) pandémico en donde refieren que el origen de los segmentos 1(PB2), 3(PA), 6(NA) y 7(MP) proviene de influenza aviar; de los segmentos 4(HA), 5(NP) y 8(NS) proviene de influenza porcina; así como el segmento 2(PB1) tiene su origen en cepas de influenza humana. Los datos que obtuvimos en este estudio corroboran lo reportado en los estudios antes mencionados para todos los segmentos excepto el segmento 1(PB2) que refirió mayor similitud con secuencias porcinas.

De manera interesante estudios ulteriores han pretendido inferir las fechas del MRCA debido a que este dato estima la fecha más probable en que el ancestro putativo circuló, en este caso, del cual derivan los virus que causaron el brote de la pandemia. Un estudio previo reportó que las fechas oscilan entre el 27 de marzo del 2008 y el 24 de octubre del 2008 (Smith et al., 2009b). Así como en otro estudio determinaron para el segmento 6(NA) que la fecha de aparición del MRCA fue el 17 de agosto del 2008 (Goni et al., 2009). Se observa una clara diferencia entre las fechas reportadas por estos autores y las nuestras. Esto puede deberse a que nuestro estudio incluye secuencias que abarcan desde el 30 de marzo del 2009 hasta el 31 de marzo del 2010; en cambio los estudios previos tienen ventanas temporales mucho más cortas. Puesto que las mutaciones ocurren al azar, la estimación del tiempo del MRCA no es un número exacto, sino más bien una distribución de probabilidad, a medida que más información se compara, la estimación se vuelve más refinada, es por eso que las fechas obtenidas del MRCA en nuestro estudio son mucho más confiables que las reportadas anteriormente.

Es de gran importancia mencionar como ha evolucionado el virus a lo largo del año en que ha circulado por todo el mundo. Un marcador claro de evolución es la tasa de mutación

global, en las que podemos observar que las tasas de mutación obtenidas son menores a las reportadas previamente (Goni et al., 2009; Smith et al., 2009a; Smith et al., 2009b). Esto se puede deber a que en un inicio hay un surgimiento de distintas variantes del virus, lo cual permite explorar diferentes combinaciones de cepas, para que finalmente preponderé la más resistente y que posiblemente sea la que encontremos en etapas tardías. Adicionalmente, los resultados obtenidos de los árboles filogenéticos de cladas de máxima credibilidad, apoyan la noción de que la mayoría de las proteínas tienden a la conservación, después de haber pasado por un periodo de expansión, que muy probablemente es cuando los estudios previos realizaron su análisis.

Es de alta importancia mencionar el uso de los antivirales, el impacto de estos sobre la presión selectiva del virus y la manera en que la incrementan. Para personas de alto riesgo el acceso a estos medicamentos constituye un beneficio significativo; sin embargo el uso indiscriminado de los mismos a largo plazo puede resultar en el desarrollo de mutaciones que hacen al virus resistente a estos antivirales. Las mutaciones asociadas a fármacos antivirales más importantes que han sido reportadas, se localizan en la posición 119 y 274 en el segmento 6(NA), que confieren resistencia a oseltamivir y en la posición 151 que confiere resistencia a zanamivir. Harvala et al., 2010 recientemente reportaron cepas de influenza A(H1N1) pandémica resistentes a oseltamivir. Además la OMS reportó el 11 de agosto del 2010, la existencia de 120 cepas virales resistentes a oseltamivir (WHO, 2010). Ninguna de estas mutaciones fue encontrada en las secuencias generadas en este estudio para el segmento 6(NA). Esto indica, aunque de forma limitada dado el pequeño número de muestras estudiadas, la ausencia de resistencia a los inhibidores de NA entre las cepas más comúnmente aisladas en nuestra región, más no constituye evidencia de la inexistencia de

las mismas. Lo que nos refiere un punto importante para continuar con esta línea de investigación e indagar más en el tema.

Una de las principales características de los influenzavirus A es que las cepas circulantes en una temporada sean iguales o muy parecidas entre sí. Puesto que influenza A (H1N1) pandémica es un virus de linaje relativamente nuevo, el comportamiento observado ha sido el esperado. Es decir, el virus tiende a la conservación después de venir de una serie de cambios que provocaron que el virus saltara de hospedero a hospedero. Esto se ratifica con los datos obtenidos de la relación dN/dS , en donde para la mayoría de las proteínas dio una significancia estadística para la teoría de selección purificante. Otra probable explicación es que solo hemos evaluado un año en el que el virus ha circulado.

Se sabe que las mutaciones puntuales acumuladas debido a la presión selectiva que se encuentran en los influenzavirus son la principal causa de diversidad de estos virus (Khiabanián, Trifonov, and Rabadan, 2009; Rambaut et al., 2008). Sin embargo He et al., 2009 recientemente reportaron para el segmento 1(PB2) de influenza H9N2 de pollos una recombinación. En donde las primeras 1066 bases eran más similares a cepas del subtipo H9N2 de patos y el resto de la secuencia era más similar al subtipo H9N2 de pollos. Ellos proponen que la recombinación puede jugar un papel importante en la evolución del virus y en la generación de una diversidad mayor en influenzavirus aviar. Dentro de la literatura no existen reportes para México o el mundo de eventos de recombinación entre cepas de distintos hospederos o de cepas de influenzavirus humanos. Es por eso que los resultados del análisis de similitud son de extrema relevancia ya que permiten establecer bases y evidencias de la recombinación de cepas entre diferentes hospederos de diversas regiones del mundo. Este mecanismo nos da una clara evidencia de que puede jugar un papel

relevante en la generación de nuevas variantes de influenza así como de la evolución del virus mismo.

En resumen nuestros resultados sugieren que el MRCA para cada una de las proteínas que codifica el virus de influenza A (H1N1) pandémica es más reciente que el reportado por estudios previos en donde el análisis se realizó cuando comenzaba a circular el virus. La tasa de mutación global encontrada es menor a lo anteriormente descrito debido a que el genoma del virus va encaminado a mantener conservadas las estructuras de la proteínas así como la funcionalidad de las mismas. Entre las cepas estudiadas no encontramos la presencia de mutaciones que confieran resistencia a oseltamivir. La recombinación de los segmentos de influenzavirus con origen de diversas regiones y hospederos es una posible manera en la que se pueda explicar la gran diversidad que existe en este virus y no solo el re-arreglo génico que se da por el intercambio de segmentos génicos completos entre cepas de diferente origen. La vigilancia epidemiológica a nivel genético es importante puesto que la aparición de cepas que tengan las mutaciones que resultan en resistencia a los antivirales puede conducir a la expansión de dichas cepas durante la siguiente temporada en la que el virus circule.

Es muy interesante poder comparar estudios ulteriores en donde el virus comenzaba a circular contra estudios en donde el virus se encuentra entrando en un periodo post-pandémico. Pues esto demuestra el proceso evolutivo en donde un virus nuevo se ve envuelto, tratando de perpetuarse. En donde al principio de la pandemia se originó un número elevado de experimentos biológicos destinados a generar mucha diversidad con la posibilidad de que alguna de las cepas lograra colonizar a un hospedero y convertirse en una cepa biológicamente exitosa.

AGRADECIMIENTOS

Extendemos nuestro agradecimiento al personal del Laboratorio de Genómica Viral y Humana de la UASLP, en especial a la QFB Sandra E. Guerra-Palomares y al QFB Hugo I. Contreras-Treviño así como al personal del Laboratorio de Virología del departamento de Microbiología de la UASLP en especial a la QBF Lidia Barrios-Compeán y al Dr. Andreu Comas-García por su valiosa contribución al proyecto.

Para la realización de este trabajo se contó con la aportación del Fondo Mixto de Fomento a la Investigación Científica y Tecnológica CONACYT-Gobierno del Estado de San Luis Potosí (Proyecto FMSLP-2008-C01-86384).

REFERENCIAS

- Al Hajjar, S., and McIntosh, K. (2010). The first influenza pandemic of the 21st century. *Ann Saudi Med* **30**(1), 1-10.
- Arias, C. F., Escalera-Zamudio, M., Soto-Del Rio Mde, L., Cobian-Guemes, A. G., Isa, P., and Lopez, S. (2009). Molecular anatomy of 2009 influenza virus A (H1N1). *Arch Med Res* **40**(8), 643-54.
- Babakir-Mina, M., Dimonte, S., Perno, C. F., and Ciotti, M. (2009). Origin of the 2009 Mexico influenza virus: a comparative phylogenetic analysis of the principal external antigens and matrix protein. *Arch Virol* **154**(8), 1349-52.
- Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriol Rev* **35**(3), 235-41.
- Bellos, A., Mulholland, K., O'Brien, K. L., Qazi, S. A., Gayer, M., and Checchi, F. (2010). The burden of acute respiratory infections in crisis-affected populations: a systematic review. *Confl Health* **4**(1), 3.
- Boelle, P. Y., Bernillon, P., and Desenclos, J. C. (2009). A preliminary estimation of the reproduction ratio for new influenza A(H1N1) from the outbreak in Mexico, March-April 2009. *Euro Surveill* **14**(19).
- Brown, E. G. (2000). Influenza virus genetics. *Biomed Pharmacother* **54**(4), 196-209.
- Centers of Disease Control and Prevention (2009). Sequencing primers and protocol. CDC. 12 mayo 2009.
- Chang, L. Y., Shih, S. R., Shao, P. L., Huang, D. T., and Huang, L. M. (2009). Novel swine-origin influenza virus A (H1N1): the first pandemic of the 21st century. *J Formos Med Assoc* **108**(7), 526-32.
- Conenello, G. M., and Palese, P. (2007). Influenza A virus PB1-F2: a small protein with a big punch. *Cell Host Microbe* **2**(4), 207-9.
- Dawood, F. S., Jain, S., Finelli, L., Shaw, M. W., Lindstrom, S., Garten, R. J., Gubareva, L. V., Xu, X., Bridges, C. B., and Uyeki, T. M. (2009). Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med* **360**(25), 2605-15.
- Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.
- Fraser, C., Donnelly, C. A., Cauchemez, S., Hanage, W. P., Van Kerkhove, M. D., Hollingsworth, T. D., Griffin, J., Baggaley, R. F., Jenkins, H. E., Lyons, E. J., Jombart, T., Hinsley, W. R., Grassly, N. C., Balloux, F., Ghani, A. C., Ferguson, N. M., Rambaut, A., Pybus, O. G., Lopez-Gatell, H., Alpuche-Aranda, C. M., Chapela, I. B., Zavala, E. P., Guevara, D. M., Checchi, F., Garcia, E., Hugonnet, S., and Roth, C. (2009). Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* **324**(5934), 1557-61.
- García-Sepulveda, C. A. (2009). Oligonucleótidos para la amplificación genómica completa de influenza A (H1N1) pandémica, San Luis Potosí, México.
- Garten, R. J., Davis, C. T., Russell, C. A., Shu, B., Lindstrom, S., Balish, A., Sessions, W. M., Xu, X., Skepner, E., Deyde, V., Okomo-Adhiambo, M., Gubareva, L., Barnes, J., Smith, C. B., Emery, S. L., Hillman, M. J., Rivaller, P., Smagala, J., de Graaf, M., Burke, D. F., Fouchier, R. A., Pappas, C., Alpuche-Aranda, C. M., Lopez-Gatell, H., Olivera, H., Lopez, I., Myers, C. A., Faix, D., Blair, P. J., Yu, C., Keene, K. M., Dotson, P. D., Jr., Boxrud, D., Sambol, A. R., Abid, S. H., St George, K., Bannerman, T., Moore, A. L., Stringer, D. J., Blevins, P., Demmler-Harrison, G. J., Ginsberg, M., Kriner, P., Waterman, S., Smole, S., Guevara, H. F., Belongia, E. A.,

- Clark, P. A., Beatrice, S. T., Donis, R., Katz, J., Finelli, L., Bridges, C. B., Shaw, M., Jernigan, D. B., Uyeki, T. M., Smith, D. J., Klimov, A. I., and Cox, N. J. (2009). Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* **325**(5937), 197-201.
- Godoy-Lozano, E. E. (2009). Oligonucleótidos para la secuenciación de influenza A (H1N1) pandémica, San Luis Potosí, México.
- Gomez-Gomez, A., Magana-Aquino, M., Garcia-Sepulveda, C., Ochoa-Perez, U. R., Falcon-Escobedo, R., Comas-Garcia, A., Aranda-Romo, S., Contreras-Trevino, H. I., Jimenez-Rico, P. V., Banda-Barbosa, M. A., Dominguez-Paulin, F., Bernal-Blanco, J. M., Perez-Gonzalez, L. F., and Noyola, D. E. (2009). Severe pneumonia associated with pandemic (H1N1) 2009 outbreak, San Luis Potosi, Mexico. *Emerg Infect Dis* **16**(1), 27-34.
- Goni, N., Fajardo, A., Moratorio, G., Colina, R., and Cristina, J. (2009). Modeling gene sequences over time in 2009 H1N1 influenza A virus populations. *Virol J* **6**, 215.
- Harvala, H., Gunson, R., Simmonds, P., Hardie, A., Bennett, S., Scott, F., Roddie, H., McKnight, J., Walsh, T., Rowney, D., Clark, A., Bremner, J., Aitken, C., and Templeton, K. (2010). The emergence of oseltamivir-resistant pandemic influenza A(H1N1) 2009 virus amongst hospitalised immunocompromised patients in Scotland, November-December, 2009. *Euro Surveill* **15**(14).
- He, C. Q., Xie, Z. X., Han, G. Z., Dong, J. B., Wang, D., Liu, J. B., Ma, L. Y., Tang, X. F., Liu, X. P., Pang, Y. S., and Li, G. R. (2009). Homologous recombination as an evolutionary force in the avian influenza A virus. *Mol Biol Evol* **26**(1), 177-87.
- Hoffmann, E., Stech, J., Guan, Y., Webster, R. G., and Perez, D. R. (2001). Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* **146**(12), 2275-89.
- Khiabani, H., Trifonov, V., and Rabadan, R. (2009). Reassortment patterns in Swine influenza viruses. *PLoS Curr Influenza*, RRN1008.
- Kumar, S., Nei, M., Dudley, J., and Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* **9**(4), 299-306.
- Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., and Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**(1), 152-60.
- Maffey, A. F. (2008). [New viruses associated with respiratory infections in children]. *Arch Argent Pediatr* **106**(4), 341-50.
- Nicholson, K. G., Wood, J. M., and Zambon, M. (2003). Influenza. *Lancet* **362**(9397), 1733-45.
- Rambaut, A., Pybus, O. G., Nelson, M. I., Viboud, C., Taubenberger, J. K., and Holmes, E. C. (2008). The genomic and epidemiological dynamics of human influenza A virus. *Nature* **453**(7195), 615-9.
- Rudan, I., Tomaskovic, L., Boschi-Pinto, C., and Campbell, H. (2004). Global estimate of the incidence of clinical pneumonia among children under five years of age. *Bull World Health Organ* **82**(12), 895-903.
- Salomon, R., and Webster, R. G. (2009). The influenza virus enigma. *Cell* **136**(3), 402-10.

- Smith, G. J., Bahl, J., Vijaykrishna, D., Zhang, J., Poon, L. L., Chen, H., Webster, R. G., Peiris, J. S., and Guan, Y. (2009a). Dating the emergence of pandemic influenza viruses. *Proc Natl Acad Sci U S A* **106**(28), 11709-12.
- Smith, G. J., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., Ma, S. K., Cheung, C. L., Raghwani, J., Bhatt, S., Peiris, J. S., Guan, Y., and Rambaut, A. (2009b). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**(7250), 1122-5.
- SSA (2008). Boletín epidemiológico. Semana 8 vols. Secretaria de Salud, San Luis Potosí.
- Secretaría de Salud (2009). Rendición de Cuentas en Salud 2008. SSA. 2009.
- Secretaria de Salud (2010). Situación actual de la epidemia. SSA. 19 Julio 2010.
- Swedish, K. A., Conenello, G., and Factor, S. H. (2010). First Season of 2009 H1N1 Influenza. *Mt Sinai J Med* **77**(1), 103-13.
- Taubenberger, J. K., and Morens, D. M. (2010). Influenza: the once and future pandemic. *Public Health Rep* **125 Suppl 3**, 16-26.
- Thompson, W. W., Shay, D. K., Weintraub, E., Brammer, L., Cox, N., Anderson, L. J., and Fukuda, K. (2003). Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA* **289**(2), 179-86.
- Viruses, I. C. o. T. o. V. I. o. (2006). Orthomyxoviridae
- World Health Organization (2009). Pandemic (H1N1) 2009 weekly update. WHO. 11. 30 julio 2010.
- World Healt Organization (2010). Weekly update on oseltamivir resistance to pandemic influenza A (H1N1) 2009 viruses. WHO. 11 august 2010.

TABLAS

Tabla 1. Oligonucleótidos de segmentos completos utilizados para la amplificación genómica completa de cada uno de los segmentos de influenza A (H1N1) pandémica.

Segmento	Nombre	Secuencia §	Tamaño amplicón (bp)	Tm (°C)	Referencia
1 (PB2)	SwPB2-F	5'-ATG-GAG-AGA-ATA-AAA-GAA-C-3'	2279	58	(Garcia-Sepulveda, 2009)
	SwPB2-R	5'-TAA-TTG-ATG-GCC-ATC-C-3'			
2 (PB1)	SwPB1-F	5'-ATG-GAT-GTC-AAT-CCG-A-3'	2273	56	
	SwPB1-R	5'-TAT-TTT-TGC-CGT-CTG-AG-3'			
3 (PA)	SwPA-F	5'-ATG-GAA-GAC-TTT-GTG-C-3'	2151	59	
	SwPA-R	5'-CTA-CTT-CAG-TGC-AT-GTG-3'			
4 (HA)	SwHA-F	5'-ATG-AAG-GCA-ATA-CTA-GTA-G-3'	1696	56	
	SwHA-R	5'-CATATTCTACACTGTAGAGAC-3'			
5 (NP)	SwNP-F	5'-ATG-GCG-TCT-CAA-GG-3'	1496		
	SwNP-R	5'-TCA-ACT-GTC-ATA-CTC-CTC-3'			
6 (NA)	SwNA-F	5'-ATG-AAT-CCA-AAC-CAA-AAG-3'	1409		
	SwNA-R	5'-TTA-CTT-GTC-AAT-GGT-AAA-TG-3'			
7 (MP)	SwMP-F	5'-TAACCGAGGTCGAAA-3'	970		
	SwMP-R	5'-TAC-TCT-AGC-TCT-ATG-TTG-A-3'			
8 (NS)	SwNS-F	5'-ATG-GAC-TCC-AAC-ACC-3'	890		
	SwNS-R	5'-TTA-AAT-AAG-CTG-AAA-CGA-G-3'			

§ Nota: Las secuencias del oligonucleótido antisentido indicadas son el complemento-reverso de las secuencias presentes en los alineamientos

Tabla 2. Oligonucleótidos para secuenciación genómica utilizados para la caracterización genómica por fragmentos.

Segmento	Nombre	Secuencia §	Tamaño amplicón (bp)	Tm (°C)	MgCl ₂ (mM)	Referencia
1 (PB2)	SwPB2-F1	5'-AAA-GAA-CTG-AGA-GAT-CTA-3'	504	50	2.5	(Godoy-Lozano, 2009)
	SwPB2-R1	5'-CCA-CTT-CAT-TTG-GGA-A-3'				
	SwPB2-F2	5'-AGG-TTG-AAA-CAT-GGT-ACC-3'	740	58	2.0	
	SwPB2-R2	5'-GCT-CTT-CTC-CCA-ACC-A-3'				
	SwPB2-F3	5'-GCT-AAC-GGG-CAA-CCT-3'	889	58	2.0	
	SwPB2-R3	5'-CAC-ATT-CAC-AGTC-AAT-GAG-G-3'				
	SwPB2-F4	5'-CCT-AAG-GCA-ACC-AGA-AGC-3'	491	58	2.0	
	SwPB2-R4	5'-TGG-CTG-TCA-GTA-AGT-ATG-CTA-G-3'				
2 (PB1)	SwPB1-F1	5'-ATG-GAT-GTC-AAT-CCG-ACT-C-3'	616	58	1.0	(Godoy-Lozano, 2009)
	SwPB1-R1	5'-CTA-TTG-TTC-TTT-GCG-TGA-CC-3'				
	SwPB1-F2	5'-AGG-AAG-GCT-AAT-AGA-TTT-CTT-A-3'	639	58	1.0	
	SwPB1-R2	5'-AGC-ATT-TCT-GCT-GGT-AT-3'				
	SwPB1-F3	5'-GAG-TGG-TTC-AGA-AAC-ATC-3'	738	58	1.0	
	SwPB1-R3	5'-TAA-CTC-AAA-TGA-TCT-TCT-CGT-3'				
	SwPB1-F4	5'-CAG-ATG-GCT-CTT-CAA-TTG-T-3'	639	58	1.0	
	SwPB1-R4	5'-TTT-TTG-CCG-TCT-GAG-TTC-3'				
3 (PA)	SwPA-F1	5'-GGA-AGA-CTT-TGT-GCG-AC-3'	703	59	1.5	(Godoy-Lozano, 2009)
	SwPA-R1	5'-TCC-ATC-TAC-ATA-GGC-TCT-AAA-3'				
	SwPA-F2	5'-CAG-TAG-GAG-TCT-ATG-GGA-T-3'	607	59	1.5	
	SwPA-R2	5'-TTT-GCA-GTC-ATC-AAA-GTC-TA-3'				
	SwPA-F3	5'-TTG-GAA-GCA-GGT-GCT-3'	700	59	1.5	
	SwPA-R3	5'-GGT-TCC-ATT-GGT-TCT-CAC-3'				
	SwPA-F4	5'-TGA-TGT-GGT-GAA-CTT-TGT-AAG-3'	600	59	1.5	
	SwPA-R4	5'-AGT-GCA-TGT-GTG-AGG-A-3'				
4 (HA)	SwHA-F1	5'-CCG-CAA-ATG-CAG-ACA-CAT-3'	510	54	1.5	(Godoy-Lozano, 2009)
	SwHA-R1	5'-TTA-ATG-TAG-GAT-TTG-CTG-A-3'				
	SwHA-F2	5'-GGC-CCA-ATC-ATG-ACT-CGA-3'	501	54	1.5	
	SwHA-R2	5'-AGG-CTG-GTG-TTT-ATA-GCA-CC-3'				
	SwHA-F3	5'-CCG-AGA-TAT-GCA-TTC-GC-3'	636	54	1.5	
	SwHA-R3	5'-CGT-TTC-CAA-TTT-CCT-TGG-C-3'				
	SwHA-F4	5'-TTG-ATG-ATG-GTT-TCC-T-3'	387	54	1.5	
	SwHA-R4	5'-TTA-GAG-CAC-ATC-CAG-AAA-3'				
5 (NP)	NPCDC-F1	5'-TGT-AAA-ACG-ACG-GCC-AGT-CAG-GGT-AGA-TAA-TCA-CTC-AC-3'	592	50	2.0	(CDC, 2009)
	NPCDC-R1	5'-CAG-GAA-ACA-GCT-ATG-ACC-AGA-GCA-CAT-YCT-GGG-ATC-CAT-3'	566		3.0	
	NPCDC-F3	5'-TGT-AAA-ACG-ACG-GCC-AGT-TGG-CAT-TCH-AAT-TTR-AAT-GAT-3'			732	
	NPCDC-R3	5'-CAG-GAA-ACA-GCT-ATG-ACC-CTG-RCT-CTT-GTG-TGC-DGG-3'				
	NPCDC-F5	5'-TGT-AAA-ACG-ACG-GCC-AGT-TTC-TGA-GRG-GRT-CAG-TTG-CTC-3'	492		2.0	
	NPCDC-R5	5'-CAG-GAA-ACA-GCT-ATG-ACC-AGT-AGA-AAC-AAG-GGT-ATT-TTT-C-3'				
8 (NS)	NSCDC-F1	5'-TGT-AAA-ACG-ACG-GCC-AGT-AGC-AAA-AGC-AGG-GTG-ACA-AAG-ACA-3'	492	50	2.0	(CDC, 2009)
	NSCDC-R1	5'-CAG-GAA-ACA-GCT-ATG-ACC-TCG-GTG-AAA-GCC-CTT-A-3'				
	NSCDC-F2	5'-TGT-AAA-ACG-ACG-GCC-AGT-TGA-GGC-WYT-TAA-AAT-GAC-CA-3'	682		3.0	
	NSCDC-R2	5'-CAG-GAA-ACA-GCT-ATG-ACC-AGT-AGA-AAC-AAG-GGT-GTT-TTT-TAT-3'				

§ Nota: Las secuencias del oligonucleótido antisentido indicadas son el complemento-reverso de las secuencias presentes en los alineamientos

Tabla 3. Comparación de los porcentajes de similitud entre las secuencias nucleotídicas locales y las derivadas de otros hospederos y/o con otro origen geográfico .

Segmento	Hospedero	Similitud (%) §	Origen geográfico	Similitud (%) §
1(PB2)	Porcinos	92.2-87.6	Asia	92.6-87.0
			Norte América	88.0-78.3
			Europa	82.5-77.7
			Norte América	95.9-88.5
2(PB1)	Humanos	95.6-93.5	Oceania	94.2-87.4
			Asia	94.1-85.1
			Europa	93.2-85.6
			Sudamérica	92.9-82.8
			África	92.5-80.1
			Oceania	94.5-89.3
3(PA)	Aves	94.2-88.0	Norte América	94.0-90.4
			Asia	92.8-86.8
			Europa	92.8-86.5
			África	89.3-83.5
			Sudamérica	88.0-76.8
			Norte América	92.2-89.6
4(HA)	Porcinos	84.2-92.7	Asia	85.5-83.7
			Europa	76.3-62.1
			Norte América	95.5-92.6
5(NP)	Porcinos	92.6-96.1	Asia	89.0-87.5
			Europa	86.9-73.9
			Asia	88.7-80.1
6(NA)	Aves	70.7-44.1	Europa	88.0-79.1
			Oceania	87.4-81.9
			Norte América	86.6-73.2
			Sudamérica	85.0-72.4
			África	84.8-78.7
7(MP)	Aves	92.4-88.9	Asia	93.6-89.8
			Oceania	92.2-89.0
			Europa	91.4-89.3
			Norte América	90.6-87.2
			África	89.1-84.7
			Sudamérica	88.7-84.1
8(NS)	Porcinos	94.2-91.6	Norte América	94.5-91.8
			Asia	91.9-87.5
			Europa	84.8-72.6

§ Valor máximo y mínimo de similitud porcentual obtenido con SimPlot.

Tabla 4. Estimación del ancestro común más reciente (MRCA) para cada fragmento genómico de las cepas de la influenza pandémicas mexicanas basado en el análisis coalescente de inferencia Bayesiana.

Segmento	Proteína	Fecha estimada del MRCA	Tasa de mutación x 10 ⁻³ (s/s/a) §
1	PB2	11 ENE 2009 (6 OCT 2008- 18 MAR 2009)	6.32 (3.99-8.87)
2	PB1	16 DIC 2008 (3 SEP 2008- 5 MAR 2009)	5.31 (3.17-7.70)
3	PA	2 NOV 2008 (28 MAY 2008 - 28 FEB 09)	4.22 (2.24-6.33)
4	HA	30 ENE 2009 (19 NOV 2008 – 18 MAR 09)	17.6 (11.9-23.6)
5	NP	20 DIC 2008 (10 SEP 08 – 6 MAR 2009)	6.53 (3.83-9.28)
6	NA	19 OCT 2008 (13 JUN 2008 – 20 FEB 2009)	6.30 (3.13-10.0)
7	M1	8 JUL 2008 (27 MAR 2007 – 24 MAR 2009)	5.09 (1.22-9.62)
7	M2	16 MAR 2009 (20 MAR 2009 – 3 ABR 2009)	49.5 (19.0-87.3)
8	NS1	21 FEB 2009 (12 DIC 08 – 30 MAR 2009)	17.7 (6.12-33.1)
8	NS2	22 MAR 2009 (25 MAR 2009 – 30 MAR 2009)	58.5 (25.5-93.6)

Nota: Los resultados en paréntesis corresponden al IC 95%.

§ Substituciones nucleotídicas por sitio por año.

FIGURAS
Figura 1.

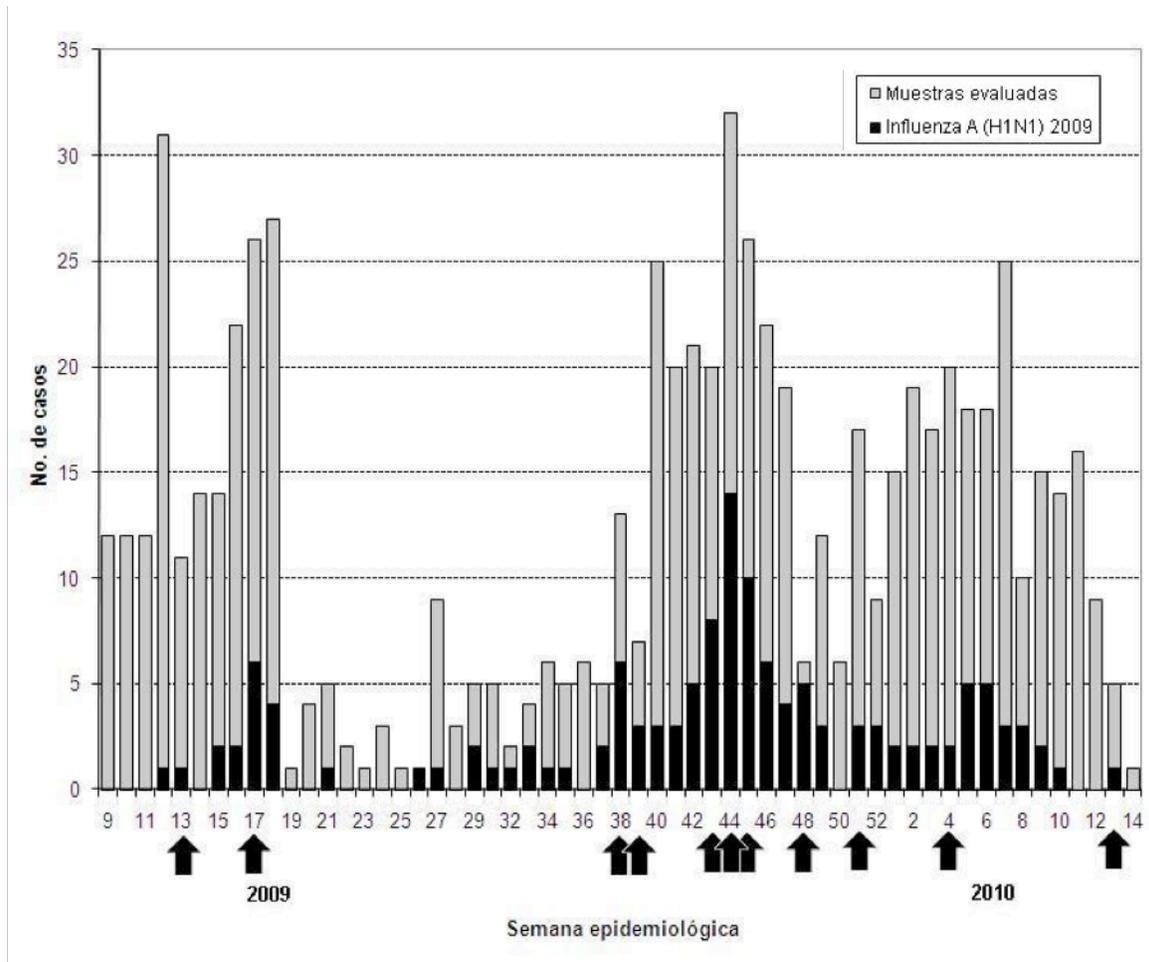


Figura 2.

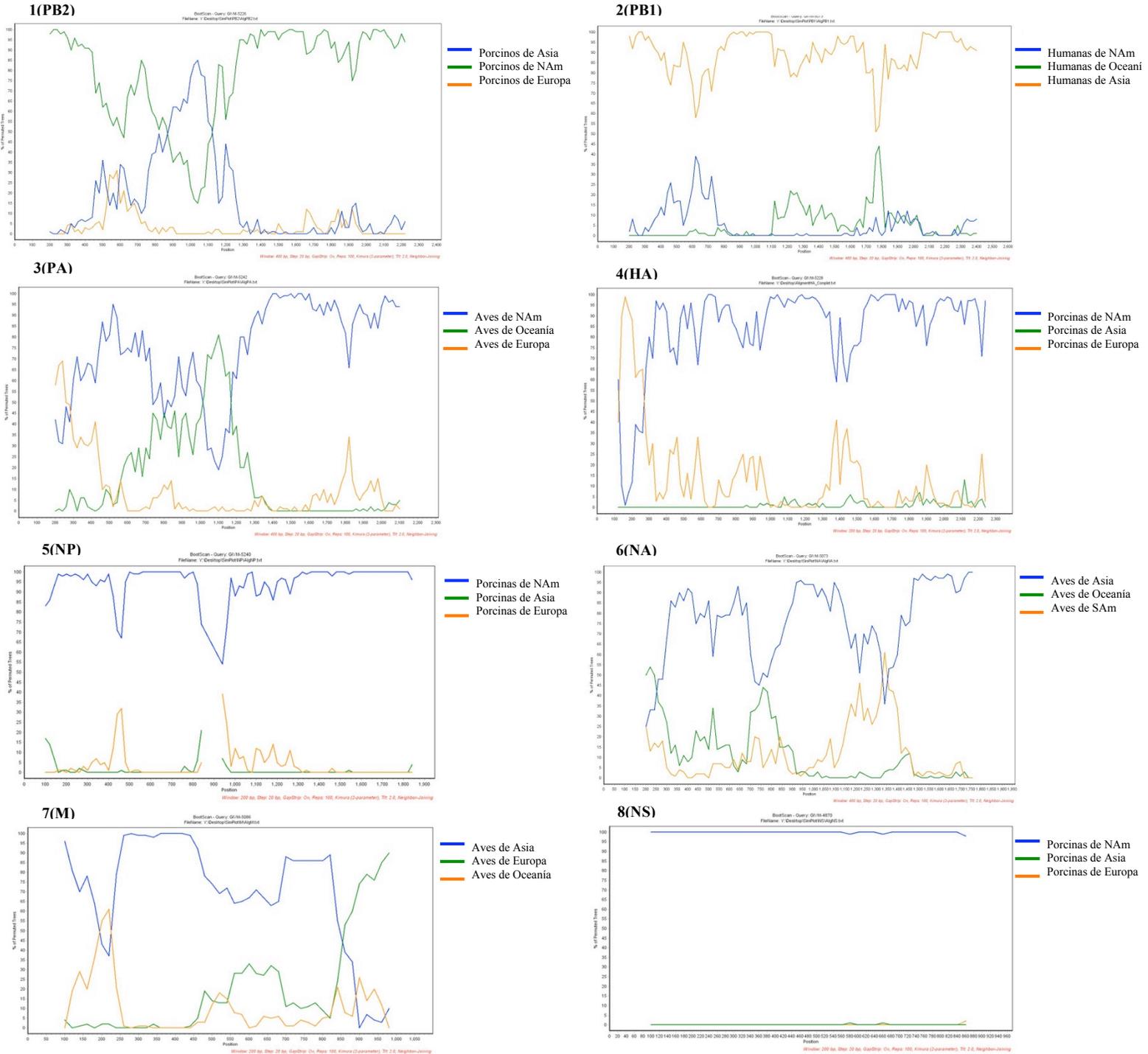


Figura 3.

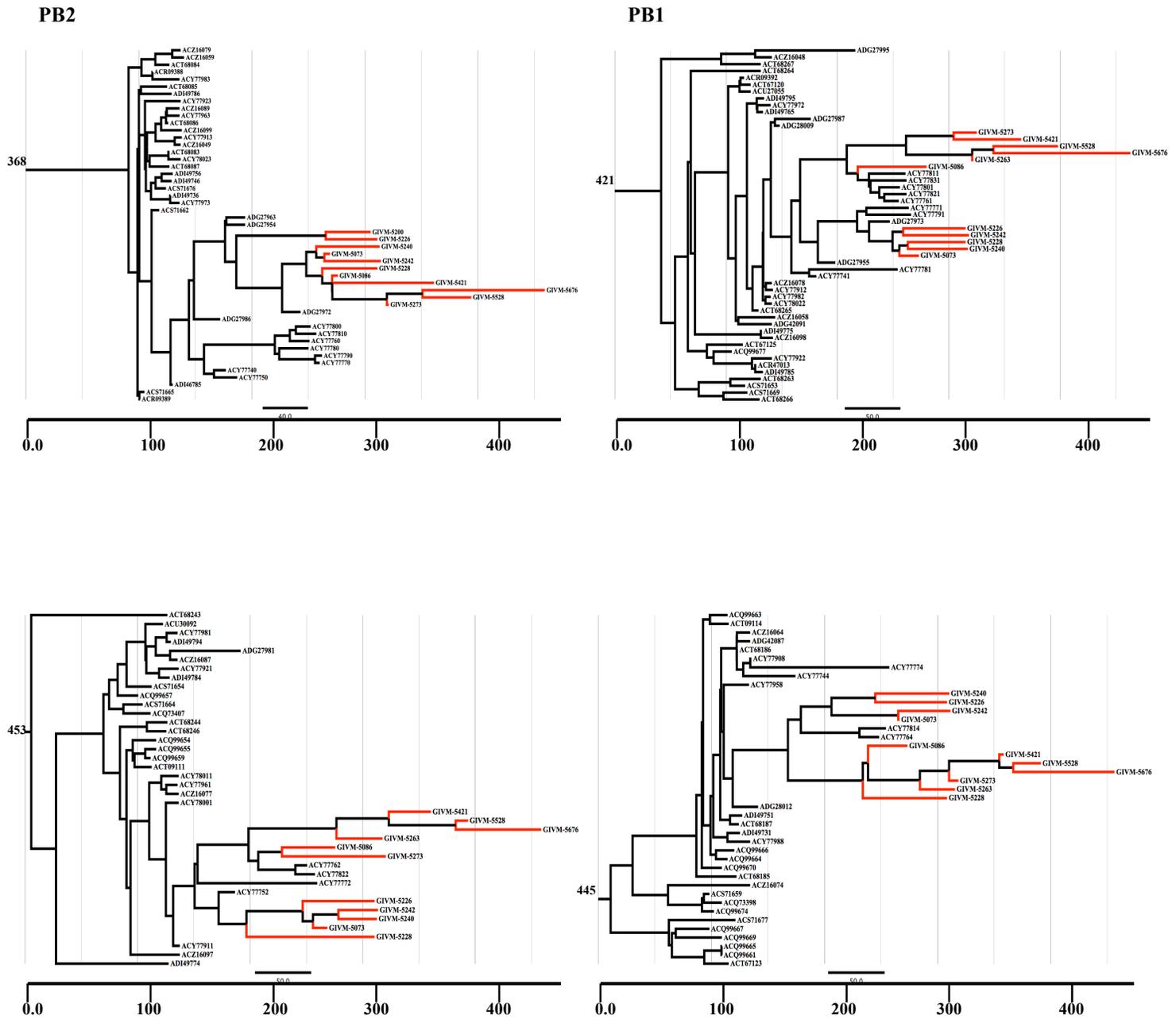
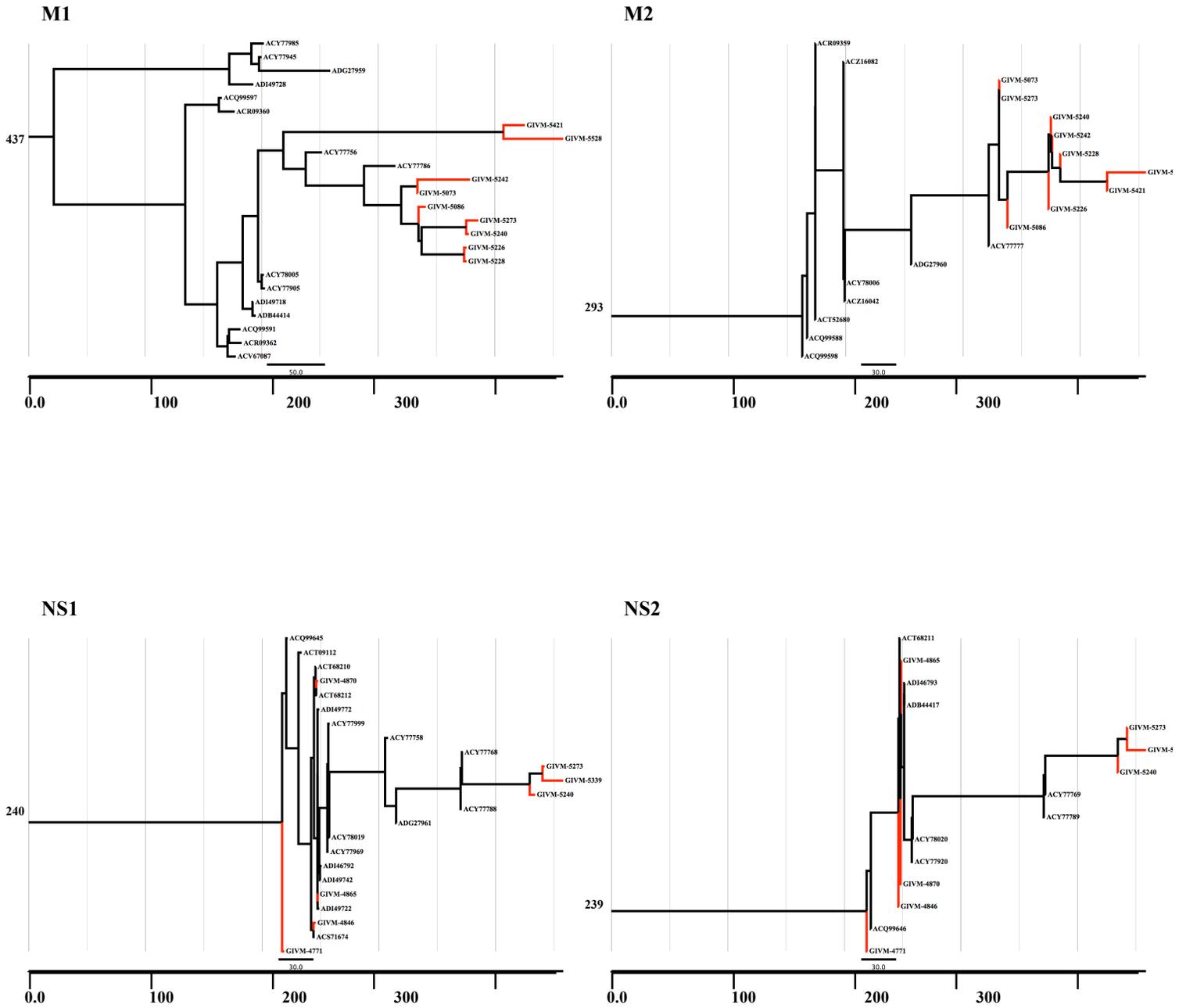


Figura 5.



PIE DE FIGURAS

Figura 1. Número semanal de muestras evaluadas para la detección de influenza pandémica A (H1N1) 2009; sistema de vigilancia de virus respiratorios del Hospital Central “Dr. Ignacio Morones Prieto” y Facultad de Medicina, UASLP entre el 1° de marzo de 2009 y el 9 de abril de 2010. Las flechas indican las semanas en que se obtuvieron las muestras incluidas en el análisis genómico. Se incluyó una muestra de cada una de las semanas señaladas a excepción de las semanas 17, 44 y 45 en las cuales se incluyeron 3, 4 y 2 muestras, respectivamente.

Figura 2. Análisis de recombinación por Bootscan en donde la posición nucleotídica se encuentra en el eje de las abscisas y el porcentaje de permutación en las ordenadas.

Figura 3. Árboles filogenéticos de cladas de máxima credibilidad para las proteínas PB2, PB1, PA y NP. Las distancias evolutivas se muestran en días y la raíz del árbol corresponde al MRCA. Las muestras analizadas en este estudio se muestran en rojo mientras que el resto se muestran en negro.

Figura 4. Árboles filogenéticos de cladas de máxima credibilidad de las proteínas HA y NA. Las distancias evolutivas se muestran en días y la raíz del árbol corresponde al MRCA. Las muestras analizadas en este estudio se muestran en rojo mientras que el resto se muestran en negro.

Figura 5. Árboles filogenéticos de cladas de máxima credibilidad de las proteínas M1, M2, NS1 y NS2. Las distancias evolutivas se muestran en días y la raíz del árbol corresponde al MRCA. Las muestras analizadas en este estudio se muestran en rojo mientras que el resto se muestran en negro.