

# Caracterización de la diversidad genética y de la dinámica evolutiva de influenzavirus humanos mexicanos.

Hugo I. Contreras Treviño<sup>a</sup>, Daniel E. Noyola<sup>b</sup>, Christian A. García-Sepúlveda<sup>a</sup>

<sup>a</sup> Laboratorio de Genómica Viral y Humana, Facultad de Medicina, Universidad Autónoma de San Luis Potosí. Avenida Venustiano Carranza #2405, Colonia Filtros Las Lomas, CP 78210, San Luis Potosí, México.

<sup>b</sup> Laboratorio de Virología, Departamento de Microbiología, Facultad de Medicina, Universidad Autónoma de San Luis Potosí. Avenida Venustiano Carranza #2405, Colonia Filtros Las Lomas, CP 78210, San Luis Potosí, México.

## Correspondencia:

Christian A. García-Sepúlveda, MD PhD  
Laboratorio de Genómica Viral y Humana  
Facultad de Medicina, UASLP  
Avenida Venustiano Carranza #2405  
Colonia Filtros Las Lomas, CP 78210  
San Luis Potosí, México  
Work +52 (444) 826-2345 extensión 568  
Mobile +52 (444) 1135-860  
ca.garcia.s@gmail.com

## Introducción

La influenza es una infección aguda de vías respiratorias causada por los virus de la influenza, la cual causa a nivel mundial entre 250 000 a 500 000 muertes anuales (50). En México se reportaron 7 200 muertes y 66 000 individuos que recibieron atención hospitalaria por complicaciones respiratorias en los años 2001 y 2002, respectivamente, entre individuos mayores a los 65 años de edad (26). Los virus de la influenza que generalmente infectan a los humanos corresponden a los géneros A, B y C; pertenecientes a la familia *Orthomyxoviridae*. El virus de la influenza A es el responsable de epidemias anuales durante el invierno lo cual ocasiona un incremento importante de ausentismo laboral así como en la demanda de consultas médicas (45). Así mismo, el influenzavirus A tiene la capacidad de ocasionar pandemias globales periódicas con un gran impacto económico, social y sanitario (27).

El virus de la influenza A se clasifica en subtipos en base a la antigenicidad de las glicoproteínas de superficie hemaglutinina (HA) y neuraminidasa (NA). Actualmente, 16 subtipos de HA (1-16) y 9 de NA (1-9) son reconocidos (18). Los influenzavirus infectan a una gran variedad de animales incluyendo a humanos, cerdos, caballos, mamíferos marinos y aves. La mayoría de los subtipos virales se replican en las células intestinales de aves acuáticas sin causar enfermedad evidente (62).

Los virus de la influenza poseen un genoma fragmentado de RNA en sentido negativo (-ssRNA) y una envoltura que posee los principales determinantes antigénicos (HA y NA). El virión es pleomórfico y tiene entre 80 a 120 nm de diámetro. La nucleocápside contiene al genoma viral distribuido en 8 segmentos genómicos cuyos tamaños oscilan entre 0.9 a 2.3 kb (44). Cada segmento codifica para una proteína, excepto los fragmentos 2(PB1), 7(M) y 8(NS), los cuales codifican para dos proteínas diferentes cada uno. Las subunidades de la polimerasa (PB2, PB1 y PA) son codificadas por los segmentos 1-3 mientras que las proteínas de la superficie (HA, NA y M2) son codificadas por los segmentos 4, 6 y 7, respectivamente. El segmento 5 codifica para la nucleoproteína (NP), la cual tiene la función de estabilizar el RNA genómico viral formando complejos ribonucleoproteicos (RNP) junto con las subunidades de la polimerasa. Los RNP's se encuentran unidos a la proteína de matriz (M1) también codificada por el segmento 7. Por último, las proteínas no estructurales NS1 y NS2 codificadas en el segmento 8 intervienen en la inhibición de la respuesta inmune innata y en la exportación del RNA viral al núcleo, respectivamente (5).

Las epidemias anuales causadas por estos virus son generadas por la emergencia de virus con nuevas variantes antigénicas como resultado de la acumulación de mutaciones puntuales, fenómeno al que se le denomina *drift* antigénico. Este acumulo de mutaciones se debe a la baja fidelidad de la RNA polimerasa viral la cual posee una tasa de mutación de  $7.3 \times 10^{-5}$  mutaciones/nucleótido replicado (12). Los virus de RNA rápidamente acumulan mutaciones que les permiten escapar a la respuesta inmunológica. Este hecho ha sido exhaustivamente estudiado para el virus de la inmunodeficiencia humana tipo 1 (HIV-1). Este virus presenta gran divergencia en las regiones hipervariables aun dentro de una misma célula infectada (37). La tasa de mutación de los influenzavirus ( $7 \times 10^{-3}$  sustituciones por sitio, s/s) es aproximadamente 15 veces superior que la del ser humano ( $5 \times 10^{-4}$  s/s) pero inferior a la del virus de la inmunodeficiencia humana tipo 1 ( $8 \times 10^{-2}$  s/s) (3). Las grandes pandemias ocurren cuando surge una variante viral cuya antigenicidad es tan diferente que no es reconocida por la memoria del sistema inmune. Estos cambios surgen del reacomodo de segmentos genómicos virales resultado de la co-infección de una célula hospedera por dos subtipos virales diferentes. A este fenómeno se le denomina *shift* antigénico (54).

Las pandemias de influenza se han caracterizado por una elevada morbi-mortalidad y han destacado a lo largo de la historia para incluir a la de 1918 (50 millones de muertes), la de 1957 (2 millones de muertes), la de 1968 (0.5 millones de muertes) y la más reciente del 2009 con una mortalidad estimada en el estado de San Luis Potosí de 19.71 muertes/100,000 habitantes y fue similar a la mortalidad ocasionada por los virus pre-pandémicos de baja virulencia (5, 52). En la aparición de los virus pandémicos de 1957 y 1968, *el shift* antigénico tuvo un rol importante al ser responsable la adquisición de nuevos segmentos genómicos para la subunidad PB1 de la polimerasa, hemaglutinina y neuraminidasa de hospederos aviares (31, 53). Por otro lado varios estudios han resaltado la homología que existió entre el genoma del virus de la influenza pandémica de 1918 con un precursor viral de hospedero aviar introducido a los humanos poco años antes de la pandemia (58, 59). Los subtipos H1N1 y H3N2 han circulado a lo largo de la historia registrada de las pandemias de influenza con distinta prevalencia estacional cada año. Durante la primavera del 2009 surgió una nueva variante del virus A(H1N1) resultado del triple rearrreglo genómico (52).

Los estudios de caracterización genómica han demostrado ser una herramienta muy valiosa para el control y entendimiento de los influenzavirus y de las enfermedades que ocasionan. La descripción de la diversidad genética, tasa de mutación y de la existencia de eventos de rearrreglo/recombinación genética nos brindan información sobre la manera en que evolucionan los virus, como surgen nuevas variantes antigénicas, así como del tipo de fuerzas selectivas que guían la dirección y rapidez de su evolución (57). Por otro lado, la caracterización de las sustituciones proteicas brinda información útil respecto a la antigenicidad viral, resistencia a medicamentos antivirales, especificidad (tropismo) para un hospedero o sobre la eficacia de los métodos diagnósticos comúnmente empleados. Por último e importantemente, la caracterización genómica permite definir el tipo de cepas de influenzavirus que se encuentran circulando en una determinada región y con ello guiar el desarrollo de estrategias de vacunación eficaces (9).

Para monitorear la emergencia de nuevas cepas, la organización mundial de la salud (OMS) mantiene un programa global de vigilancia, donde un panel de expertos se reúne dos veces por año para analizar los datos antigénicos, genéticos y epidemiológicos de las cepas virales circulantes. Este panel dirige y extiende recomendaciones respecto a la composición de la vacuna a aplicar durante la temporada invernal siguiente (51).

El análisis bioinformático de la información genética ha permitido describir los patrones de la dinámica evolutiva que siguen los influenzavirus, hecho que se ha visto facilitado por el desarrollo de las grandes bases de datos de información genética. Los repositorios de información genética de los virus de la influenza como NCBI Influenza Virus Resource y el Global Initiative on Sharing All Influenza Data *GISAID* (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html> y [www.gisaid.org](http://www.gisaid.org)) han demostrado jugar un papel crucial en la respuesta mundial al surgimiento de nuevos patógenos y variantes del virus (44). El objetivo de este trabajo fue el de estudiar la extensión de la diversidad genómica de influenzavirus estacionales pre-pandémicos y estudiar su relevancia clínica y epidemiológica.

## **Materiales y métodos.**

### **Bioespecímenes y población de estudio.**

Muestras de secreciones del tracto respiratorio correspondientes a diversos programas de vigilancia epidemiológica a nivel estatal (n = 839) y de la ciudad de San Luis Potosí (n = 644) fueron empleadas para este estudio. Las muestras biológicas fueron compiladas durante la temporada invernal 2007-2008 al igual que durante los primeros meses de la pandemia de la influenza A(H1N1) 2009. Las muestras fueron almacenadas a -80°C en medio de transporte suplementado con antibióticos, buffer y albúmina bovina hasta que fueron requeridas para el proceso de extracción de RNA.

### **Extracción, evaluación y almacenamiento de RNA viral.**

El RNA viral se aisló empleando el kit comercial *High Pure Viral RNA* (Roche Diagnostics, Mannheim, Alemania) de acuerdo a las especificaciones del fabricante. La extracción de los ácidos nucleicos virales se realizó en un gabinete de seguridad biológica clase II dentro de un laboratorio nivel de bioseguridad II conforme a las recomendaciones establecidas por los CDC's para el trabajo con muestras respiratorias

extemporáneas del virus de la influenza y de acuerdo a prácticas y disciplina correspondiente a un nivel de bioseguridad 3 ([http://www.cdc.gov/h1n1flu/guidelines\\_labworkers.htm](http://www.cdc.gov/h1n1flu/guidelines_labworkers.htm)). La calidad del RNA extraído fue evaluada por espectrofotometría. El RNA viral extraído permaneció en congelación a -80 °C hasta ser empleado para RT-PCR.

#### **Síntesis de cDNA.**

La síntesis de primer cadena (cDNA) del RNA viral extraído de las muestras positivas para virus de la influenza A al igual que del las muestras destinadas al tamizaje de la variante pandémica de la influenza A(H1N1) 2009 se realizó con un paso de RT separado de la PCR. Dicha reacción se llevó a cabo a 38 °C por 60 minutos empleando el oligonucleótido genérico UniFlu-RT (5'-AGG-AAA-AGC-AGG-3') a la concentración final de 2.25 µM, 2 µL de RNA y 50 IU totales de la transcriptasa inversa M-MuLV (Vivantis Technologies Sdn Bhd Selangor D.E., Malasia) a un volumen final de 10 µL de acuerdo a un protocolo previamente descrito (48).

#### **Detección molecular/tipificación del virus de la influenza A y B.**

Las muestras fueron tamizadas para la detección/tipificación de virus de la influenza A y B de acuerdo a un protocolo de un solo paso de RT-PCR previamente estandarizado (1, 7).

#### **Detección de la variante pandémica del virus de la influenza A(H1N1) 2009.**

Oligonucleótidos dirigidos contra el segmento 8(NS) de la variante pandémica de la influenza A(H1N1) 2009 fueron diseñados en base a las primeras secuencias nucleotídicas depositadas en la base de datos EpiFlu (<http://plataform.gisaid.org>) para su uso en la técnica de detección de éste virus que fue desarrollada localmente. La técnica de detección de la variante pandémica del virus de la influenza A(H1N1) 2009 utilizó un acercamiento RT-PCR anidado empleando para la PCR externa los oligos de caracterización genómica del segmento 8(NS) diseñados por nosotros previamente (Sw-NS-F 5'-ATG-GACTCC-AAC-ACC-3' y Sw-NS-R 5'-TTA-AAT-AAG-CTG-AAA-CGA-G-3'). Para la segunda PCR se emplearon los oligonucleótidos (Sw1-F 5'-CTT-GAA-AGA-GGA-ATC-GAG-CG-3' y Sw1-R 5'-GTC-TCC-CAT-TCT-CAT-CAC-AGT-3') los cuales generaban un amplicón de 409 bp en presencia del citado virus. La primer PCR utilizó los oligonucleótidos a una concentración final de 1.6 µM, 0.5 IU de Taq DNA polimerasa y MgCl<sub>2</sub> 1.5 mM. Las condiciones de amplificación para la PCR externa constan de un paso inicial de desnaturalización a 95 °C por 5 min, 30 ciclos a 95 °C por 20 s, 56 °C por 30 s, y 72°C por 90 s, seguidos de un paso final de 2 min a 72 °C. Para la PCR interna se tomó 1 µL de una dilución 1:1 del producto de la primer PCR empleando MgCl<sub>2</sub> 1.0 mM y 800 nM de los oligonucleótidos internos. Las condiciones de amplificación para la PCR interna fueron similares a las de la externa pero empleando 40 ciclos y un tiempo de extensión de 30 s (21). Los amplicones se visualizaron y documentaron tras electroforesis en geles de agarosa al 1.5% a 4.7 V/cm en presencia de bromuro de etidio.

### **Subtipificación del virus de la influenza tipo A.**

Oligonucleótidos específicos para los segmentos 4(HA) y 6(NA) de subtipos patogénicos humanos (HA1, HA3, HA5, NA1 y NA2) fueron diseñados para su uso en la técnica de detección y subtipificación molecular a partir de una base de datos de 416 secuencias nucleotídicas compiladas de *Influenza Virus Resource Database* NCBI (2, 10). Las secuencias fueron alineadas empleando la aplicación Muscle (15) y reformateados para unanimidad empleando una herramienta de bioinformática brindada por el Anthony Nolan Research Institute de Londres, Inglaterra (<https://www.alleles.org/internal/common/reformat.html>). Las características y la secuencia de los oligonucleótidos empleados para la técnica multiplex de subtipificación molecular de influenza virus A se muestran en la tabla 1. La PCR empleada para tal fin utilizó 5 µL de cDNA y concentraciones finales de MgCl<sub>2</sub> 3.0 mM, oligonucleótidos 400 nM y 0.5 IU de Taq (Vivantis Technologies Sdn Bhd Selangor D.E., Malaysia). Las condiciones de PCR constaron de un paso inicial de desnaturalización a 95 °C por 5 min, 40 ciclos de 30 s a 95 °C, 30 s y 2 min a 72 °C, seguidos de un paso final de extensión a 72 °C por 5 min.

### **Caracterización genómica del virus de la influenza tipo A, subtipo H1N1 y H3N2.**

Oligonucleótidos dirigidos contra los segmentos genómicos individuales fueron diseñados para su empleo en las técnicas de caracterización genómica de los aislados virales a partir de alineamientos de secuencias nucleotídicas compiladas a partir de *Influenza Virus Resource Database* NCBI (2, 10). Los alineamientos correspondientes a cada segmento genómico de ambos subtipos virales (H1N1 y H3N2) incluyeron a 613 secuencias para PB2, 595 secuencias para PB1, 591 secuencias para PA, 659 secuencias para HA, 574 secuencias para NP, 569 secuencias para NA, 479 secuencias para M y 456 secuencias para NS. Los alineamientos fueron generados con la aplicación ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2>) y posteriormente reformateados a unanimidad. Los oligonucleótidos fueron diseñados para ser específicos y evitar hibridación cruzada con otras regiones en la secuencia u otros segmentos genómicos. Los ocho segmentos genómicos de cinco aislados virales A(H1N1) y de cinco A(H3N2) fueron amplificados y secuenciados empleando los oligonucleótidos descritos en la tabla 2. La estrategia de caracterización genómica empleó una primera PCR (externa) para amplificar el segmento completo seguida de la amplificación de uno o más fragmentos subgenómicos a través de una segunda PCR (interna). La amplificación de los segmentos PB2, PB1, PA, NP, NA, M y NS se realizó a 1.5 mM de MgCl<sub>2</sub>, 1.6 µM de cada oligonucleótido y 0.5 IU totales de una mezcla Taq:Pfu (9:1). La amplificación del segmento codificante para HA se realizó a 1.0 mM de MgCl<sub>2</sub>. Las condiciones de amplificación para los segmentos PB2, PB1, PA y NP constó de un paso de desnaturalización inicial a 95 °C por 5 min, 40 ciclos a 95 °C por 20 s, 58 °C por 30 s y 72 °C por 3 min, seguidos de un paso final a 72 °C por 5 min. Las condiciones de amplificación para los segmentos HA y NA fueron similares pero una temperatura de hibridación de 60 °C y un tiempo de extensión en cada ciclo de 90 s. Para los segmentos NS y M el tiempo de extensión de cada ciclo se redujo a 30 s. Con motivo de permitir la amplificación de secuencias de alta calidad para los segmentos genómicos largos (PB2, PB1, PA, HA, NP y NA), se diseñaron y emplearon oligonucleótidos internos que permitieran generar

amplicones correspondientes a fragmentos subgenómicos de entre 740 y 990 bp (ver tabla 2). Para ello se emplearon reacciones individuales de 12.5  $\mu$ L de volumen final con 2  $\mu$ L de cDNA, MgCl<sub>2</sub> 1.5 mM, 800 nM de cada oligonucleótido y 0.5 IU de la mezcla Taq:Pfu (9:1). Las condiciones de amplificación para todos los segmentos genómicos (excepto PA-1) constaron de un paso inicial de desnaturalización a 95 °C por 5 min, 40 ciclos a 95 °C por 20 s, 54 °C por 30 s y 72 °C por 30 s, seguidos de un paso de extensión final a 72 °C por 5 min. La amplificación del fragmento PA-1 empleó una temperatura de hibridización de 57 °C. Aquellos aislados que no permitieron generar amplicones de calidad adecuada para secuenciación empleando esta estrategia fueron amplificados utilizando un acercamiento anidado para la amplificación de los fragmentos subgenómicos. En este caso se utilizó 1  $\mu$ L del producto de la primera PCR (segmento completo) diluido 1:8 con dH<sub>2</sub>O para amplificar por PCR el fragmento de interés en un programa tipo touch-down (20 ciclos empleando 54 °C para la hibridización seguido de 20 ciclos adicionales con una temperatura de hibridización de 50 °C). Una vez logrados los amplicones, el producto de PCR sin purificar ni clonar fue referido al Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO) de Irapuato, Guanajuato para secuenciación. Los cromatogramas crudos de secuenciación fueron analizados localmente empleando la aplicación 4peaks versión 1.7.2 (<http://mekentosj.com/science/4peaks/>) para ordenadores Macintosh o FinchTV versión 1.4 (<http://www.geospiza.com/Products/finchtv.shtml>) para ordenadores IBM-compatibles. Las secuencias correspondientes a cada fragmento subgenómico fueron exportadas en formato FASTA y los contigs correspondientes al empalmamiento de las secuencias subgenómicas armadas en un editor de textos para cada muestra. Las secuencias completas correspondientes a cada segmento genómico fueron entonces alineadas empleando ClustalW2 y reformateadas a unanimidad o usadas directamente como archivos FASTA para su análisis ulterior.

### **Análisis bioinformático.**

Los archivos independientes de cada muestra en formato FASTA fueron utilizados para realizar el análisis de homología global (de segmento completo) a través de la herramienta en línea *Basic Local Alignment Search Tool* (BLAST) disponible gratuitamente en el servidor del Instituto de Bioinformática Europeo ([www.ebi.ac.uk/Tools/blast/](http://www.ebi.ac.uk/Tools/blast/)). El análisis de sustituciones no sinónimas (dN) versus sustituciones sinónimas (dS) por sitio codónico para cada uno de los segmentos genómicos se realizó empleando tanto nuestras secuencias localmente generadas como aquellas secuencias únicas para la temporada invernal 2007-2008 generadas por otros laboratorios del mundo compiladas de la base de datos *The Flu Project* (2). Para cada subtipo viral por separado se tomaron secuencias genómicas completas, únicas (no duplicadas), del subtipo correspondiente, solo aislados derivados de hospederos humanos, de cualquier región del mundo generadas entre el 01 de octubre del 2007 y el 31 de marzo del 2008. El modelo evolutivo más apropiado para cada segmento genómico se calculó alimentando las secuencias compiladas a la aplicación en línea FindModel del Laboratorio Nacional de Los Alamos (<http://www.hiv.lanl.gov/>) empleando el set reducido de 12 modelos evolutivos. El análisis de la proporción de sustituciones (dN/dS) para cada una de las proteínas codificadas por cada segmento genómico de cada subtipo viral se realizó empleando el algoritmo *Codon-based Z-test of selection* y *Position-wise codon based selection estimation* (HyPhy) del paquete de análisis molecular

evolutivo MEGA 5 (34). El algoritmo HyPhy se corrió empleando un árbol *Neighbor-joining* de referencia con un acercamiento de *Maximum Likelihood* empleando el código genético estándar, y el modelo evolutivo indicado por la aplicación FindModel. La proporción dN/dS fue calculada y graficada automáticamente por HyPhy. El tipo de presión selectiva que actúa sobre cada segmento genómico del virus se calculó a partir de la proporción dN/dS tomando como reglas  $dN = dS$  correspondiente a selección neutra,  $dN > dS$  correspondiente a selección positiva (diversidad) y  $dN < dS$  selección correspondiente a selección purificante (conservación). Las sustituciones dN detectadas en las proteínas que actúan como antígenos de superficie (HA y NA) se mapearon hacia cristales tridimensionales de referencia generados por difracción de rayos X a una resolución de entre 2.5 y 3.0 Å. Para el mapeo de las sustituciones encontradas en la hemaglutinina de nuestras secuencias H1 y H3 se emplearon los cristales 2WRG y 1HGG, respectivamente. Para el mapeo de las sustituciones encontradas en la neuraminidasa se utilizó el cristal 2HTY para ambos subtipos. Para ello se utilizó la aplicación PyMol Molecular Graphics System versión 1.3 (Schrödinger, LLC) en una plataforma OS X 10.6.

El análisis filogenético de las secuencias de cada segmento genómico viral se realizó tanto con nuestras secuencias locales como con las secuencias consenso de las distintas temporadas invernales ocurriendo entre 1968 y 2008 para cada subtipo viral por separado (2). Este análisis empleó la aplicación SimPlot versión 3.5.1 en un acercamiento de ventana de 20 bp desplazante, con 100 repeticiones, utilizando al modelo Kimura de 2-parametros y arboles del tipo Neighbor-Joining (42). El análisis coalescente de inferencia Bayesiana se utilizó para calcular la tasa de mutación global de cada región codificante, para generar filogramas así como para estimar el tiempo del ancestro común más reciente (MRCA por sus siglas en inglés). Para ello se emplearon 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 (14). Para el análisis BEAST se tomó el promedio de los árboles de máxima credibilidad a partir del 90% posterior de los ensayos generacionales realizados (60 millones para todas las regiones codificantes de H3N2 al igual que para las proteínas PB2, PA, HA, M1, M2, NS1 y NS2 de H1N1, de 30 millones para la proteína PB1 de H1N1 y 100 millones para las proteínas NA y NP de H1N1). Las secuencias nucleotídicas consenso codificantes para las proteínas de cada segmento genómico viral fueron empleadas para calcular la tasa de mutación global (de todo el segmento genómico), para documentar la distribución y tipo de mutaciones para cada posición codónica de cada segmento genómico así como para generar los árboles filogenéticos y calcular el MRCA.

## Resultados

Los archivos resultantes de nuestro análisis bioinformático así como las bases de datos, secuencias nucleotídicas/aminoácidas y los alineamientos que se han generado como parte de éste estudio se encuentran libremente disponibles a través de nuestro portal web (<https://www.genomica.uaslp.mx/Databases/Influenza.html>).

Del total de las 839 muestras correspondientes a la temporada invernal 2007-2008, 33 (3.9 %) resultaron positivas para el virus de la influenza A sin detectar la presencia de virus de la influenza B. La subtipificación de las 33 muestras positivas para influenza A reveló la presencia de 24 aislados H3N2 (72.8 %) y 8 aislados

del subtipo H1N1 (27.2 %). En el 100 % de las muestras se observó amplificación para el subtipo de hemaglutinina y neuraminidasa correspondiente. No se obtuvieron muestras con co-infección por ambos subtipos virales.

De manera adicional a este trabajo y con la necesidad de una técnica molecular de detección de la variante pandémica del virus de la influenza en el 2009, un total de 644 muestras fueron tamizadas de manera retrospectiva para la detección del virus de la influenza A(H1N1) 2009 con el objetivo de determinar el comportamiento de este patógeno en nuestra comunidad y evaluar la presencia del mismo durante los meses previos a la pandemia; las muestras se recabaron en el periodo comprendido entre agosto de 2008 y junio de 2009, encontrando 64 positivas (9.9 %).

Datos de secuenciación de alta calidad fueron generados para 10 aislados clínicos de virus de la influenza A, 5 del subtipo H1N1 y 5 del subtipo H3N2. De éstos, se obtuvieron 10 secuencias nucleotídicas completas para los segmentos 3(PA) y 5(NP), 9 para los segmentos 2(PB1), 6(NA) y 7(M), 8 para los segmentos 4(HA) y 8(NS) y 6 para el segmento 1(PB1). En total se obtuvieron 4 genomas completos y 6 genomas parciales. Tres de los 6 genomas parciales solamente carecen de información de secuencia para un solo segmento mientras que en los otros 3 falta información para dos o más segmentos. Ninguna de las secuencias generadas a partir de aislados locales presentó evidencia de heterocigocidad o ambigüedad en la secuenciación.

El análisis de homología de secuencias nucleotídicas realizado por BLAST demostró que la mayor parte de los segmentos genómicos de nuestros aislados H3N2 mostraban mayor similitud con secuencias derivadas de aislados norteamericanos y asiáticos de la misma temporada invernal (2007-2008). No obstante el segmento 7(M) de nuestros aislados H3N2 demostró poseer mayor similitud (>99 %) con los aislados asiáticos del 2004 en comparación a secuencias de la misma temporada invernal. En relación a los resultados BLAST realizados sobre secuencias aminoacídicas se encontró que las proteínas derivadas de aislados H3N2 locales poseían gran similitud con las secuencias proteicas de aislados virales norteamericanos, asiáticos y europeos y de un gran rango de fechas (principalmente de los años 1994 y del 2003-2008). En relación al análisis de nuestros aislados H1N1, tanto las secuencias nucleotídicas como las aminoacídicas de la mayor parte de los segmentos genómicos demostraron mayor similitud (> 99 % en promedio) con los virus americanos y los japoneses aislados entre el 2007 y 2008. No obstante, las secuencias aminoacídicas de la neuraminidasa de nuestras aislados H1N1 mostraron mayor similitud (>99 %) con virus europeos del 2008; mientras que la proteína del canal iónico M2 mostró mayor similitud con aislados virales norteamericanos del 2000.

El análisis de homología regional realizado con SimPlot para las secuencias locales confirmó los resultados del análisis global BLAST. Se demostró mayor similitud (>95%) entre nuestras secuencias con respecto a las secuencias del mismo subtipo y de la misma temporada invernal 2007-2008 como se ilustra para el segmento 1(PB2) de H3N2 de la figura 1. Adicionalmente, la figura 1 resalta la existencia de un gradiente en el cual nuestras secuencias muestran mayor homología con las secuencias más recientes en comparación a las de mayor antigüedad. Así pues, destaca una mayor homología con respecto a las secuencias consenso de la década de los 2000's (en rojo), seguido de la década de 1990's (en naranja), 1980's (amarillo), 1970's (verde claro) y 1960's (verde oscuro). Este tipo de comportamiento también se observó al analizar las secuencias nucleotídicas de otros segmentos genómicos de ambos subtipos virales.

El análisis de homología regional empleando el algoritmo BootScan hizo evidente la presencia de remanentes vestigiales de eventos recombinatorios para algunos segmentos genómicos de ambos subtipos virales. Evidencia de dichos eventos recombinatorios se presentó para los segmentos genómicos 1(PB2), 2(PB1), 7(M) y 8(NS) de los aislados H1N1 y para los segmentos 1(PB2), 5(NP), 6(NA), 7(M) y 8(NS) de los aislados H3N2. La tabla 3 resume la localización y características de los eventos recombinatorios que fueron detectados en cada uno de los subtipos virales aislados.

El segmento 1(PB2) de los aislados H1N1 poseía un evento de recombinación de 265 bp de longitud ubicado entre la posición 576 y 841 con mayor homología a secuencias de la temporada invernal 2003-2004 (ver figura 2, panel A). Este mismo segmento en las secuencias locales de H3N2 presentaba vestigios de un evento recombinatorio de 115 bp de longitud entre la posición 1555 y 1670 el cual mostraba mayor homología con secuencias de la temporada invernal 2004-2005 (ver figura 2, panel C). El segmento 2(PB1) de los aislados H1N1 presentó evidencia de un evento recombinatorio pequeño (82 bp) entre las posiciones 1452 y 1539 exhibiendo mayor homología con las secuencias derivadas de las temporadas invernales del 2002 al 2007. El segmento 5(NP) de los aislados H3N2 poseía evidencia de un evento recombinatorio ancestral en el extremo 5' en donde la homología era mayor con secuencias del 2004-2005. El segmento 6(NA) de los aislados H3N2 mostró mayor homología con secuencias del 2004-2005 en las primeras 217 bp del segmento, mientras que el resto del segmento era más parecido a las secuencias de la temporada invernal 2007-2008. El segmento 7(M) de los aislados H1N1 presentó evidencia de un evento recombinatorio de 103 bp de longitud entre las posiciones 728 y 831 con mayor homología a secuencias de la temporada invernal 2006-2007. Interesantemente, el segmento 7(M) de los aislados H3N2 exhibió poca homología (de tan solo 30%) con las demás secuencias consenso evaluadas, esto a lo largo de las primeras 276 bp. El resto del segmento siendo más similar a las secuencias de la temporada invernal 2007-2008 sin corresponder por BLAST a otros subtipos. El segmento 8(NS) de los aislados H1N1 mostró evidencia de un evento recombinatorio complejo en el cual las primeras 209 bp del fragmento eran más similares a secuencias del 2004-2005 mientras que siguientes 63 bp (entre las posiciones 210 y 273) correspondían más a secuencias del 2006-2007 y donde el final (entre las posiciones 274 y 890) era más parecido a secuencias de la temporada invernal del 2007-2008 (ver figura 2, panel B). El segmento 8(NS) de los aislados H3N2 demostró poseer una región altamente conservada entre los influenza virus A(H3N2) localizada entre las posiciones 215 y 484 en donde no destacó la homología por ninguna temporada invernal en especial (ver figura 2, panel D).

Las gráficas de las mutaciones dN versus dS (material suplementario disponible en <https://www.genomica.uaslp.mx/Databases/Pandemicflu.html>) demostró el claro predominio de mutaciones dS respecto a las dN para todos los segmentos genómicos. La significancia de este hallazgo se confirmó estadísticamente con el *Z test for substitutions*, ( $p=0.00001$ ) en apoyo a la hipótesis de selección purificadora, tanto para las secuencias locales 2007-2008 como para las mundiales de la misma temporada. Para el segmento 8(NS2) de nuestros aislados H3N2 el valor de la relación dN/dS fue superior a 1 (indicativa de selección positiva) con un nivel de significancia estadística de  $p=0.00032$  hecho que también se hizo evidente para otros segmentos genómicos (HA, NA y M2) aunque no a un nivel estadísticamente significativo. Las proteínas 1(PB2) y 2(PB1) de ambos subtipos virales y las proteínas M1, M2 y NS2 del subtipo H1N1

demonstraron poseer sustituciones sinónimas únicamente (que no se traducían en cambios aminoacídicos). No obstante, los fragmentos 4(HA) y 6(NA) de ambos subtipos virales, tanto locales como mundiales, exhibieron un mayor número de sustituciones dN, aunque sin rebasar al número de sustituciones sinónimas. Si bien la mayor parte de las sustituciones no-sinónimas que se encontraron en los aislados locales ya habían sido descritas previamente para aislados de virus de la influenza de otras partes del mundo, un repertorio limitado de ellas demostró ser exclusivo de los aislados mexicanos que secuenciamos.

El subtipo H1N1 no compartió la mayoría de las dN con los aislados de Norteamérica y Asia en HA y NA. No obstante las dN encontradas como únicas se localizaron tanto en proteínas de superficie, internas y no estructurales (PA, HA, NP, NA y NS1). Las proteínas PA, HA, NP, NA, M1, M2, NS1 y NS2 del subtipo H3N2 mostraron dN compartidas en su mayoría con los virus que circularon en Norteamérica y Asia. Aunque en menor número las dN únicas para las secuencias mexicanas de subtipo H3N2 correspondieron a la hemaglutinina, la nucleoproteína, la neuraminidasa y a la proteína que conforma el canal iónico (M2).

La mayoría de las sustituciones no-sinónimas detectadas en las proteínas HA y NA de nuestras secuencias locales pudieron ser mapeadas en modelos moleculares de referencia. De las 16 sustituciones no-sinónimas detectadas en nuestras secuencias de HA de aislados H1N1 locales, 12 pudieron ser mapeadas en el modelo molecular (Figura 3, panel A). Las sustituciones V236L e I256T demostraron encontrarse en los sitios antigénicos Ca2 y Ca1 de la hemaglutinina que son blancos importantes de la respuesta inmune humana contra el virus. Las sustituciones E240D, G233R y G262E demostraron encontrarse próximas a uno de los sitios antigénico conocidos de la proteína (aunque no directamente en ellos) mientras que dos de las sustituciones (I265K y T315P) demostraron localizarse en la interface de trimerización de la molécula. Adicionalmente, dos sustituciones demostraron estar demasiado lejos de cualquier sitio antigénico o funcional (N285K y Q341P) mientras que tres de las sustituciones (F483L, H485L y D489G) se encontraban ubicadas en la cadena HA2 implicada en la estabilización de la molécula. Cinco sustituciones restantes (T14A, K520R, V526A, V536G y C561R) no pudieron ser mapeadas por limitaciones impuestas por el cristal de referencia. Para las secuencia derivadas de aislados H3N2, siete de las nueve sustituciones no-sinónimas encontradas en la proteína HA fueron mapeadas hacia el modelo molecular correspondiente (Figura 3, panel B). Entre ellas se incluían dos sustituciones (I074N y L173S) en los sitios antigénicos C y B, dos más (D069N y E296K) localizadas en el sitio C, dos adicionales localizadas lejos de los sitios antigénicos (Q189E y D505E) y una localizada en la interface de trimerización (I316V).

En relación a las sustituciones detectadas en la proteína NA de nuestras secuencias derivadas de aislados H1N1, cinco (S446F, C448R, V450D, L465M y P466T) fueron mapeadas a regiones distantes de los sitios antigénicos o catalíticos conocidos (Figura 4, panel A). No obstante, la mayor parte demostraron localizarse en la región de la interface responsable de las interacciones que llevan a la formación de tetrámeros de la NA. Cuatro de las sustituciones no-sinónimas encontradas en las secuencias de NA de nuestros aislados H3N2 (N329T, S355G, E434K y D463E) fueron mapeadas en la superficie exterior de la proteína pero lejos del sitio catalítico mientras que la sustitución E221D demostró encontrarse próxima (7.8 Å) a dicho sitio (Figura 4, panel B).

Las propiedades fisicoquímicas de los aminoácidos con respecto a cambios de polaridad o carga fueron descritas para las sustituciones dN encontradas en HA y NA. En la hemaglutinina H1 de las 17 sustituciones dN encontradas en 5 no hubo cambio en las propiedades del aminoácido (V236L, E240D, F483L, K520R y V526A), un cambio de carga en el aminoácido se detectó en 7 sustituciones (G233R, G262E, I265K, N285K, H485L, D489G y C561R) y en 5 sustituciones se observó un cambio de polaridad (T014A, I256T, T315P, Q341P, V536G). En cinco de las sustituciones descritas para H3 no se encontró un cambio de propiedades fisicoquímicas (T003S, I074N, Q189E, I316V y D505E), en 2 de las posiciones hubo un cambio de polaridad (F015S y L173S), un cambio de carga se detectó en la sustitución D069N y únicamente un cambio de carácter ácido a básico en la posición E296K. En ambos casos de la neuraminidasa (N1 y N2) se encontraron 5 sustituciones dN. Para N1 uno de los cambios no presentó alteración en las propiedades fisicoquímicas (L465M) y en las cuatro restantes, existen 2 cambios de polaridad (S446F y P466T) y 2 de carga (C448R y V450D). De las cinco sustituciones en la N2, en cuatro no hay cambios fisicoquímicos (E221D, N329T, S335G y D463E) y en la posición restante (E435K) se observó un cambio de aminoácido ácido a básico.

La tasa de mutación entendida como el número de sustituciones por sitio codónico por año (s/s/año) fue calculada para las secuencias mundiales (incluyendo las nuestras) de cada segmento genómico y cada subtipo viral de la temporada invernal 2007-2008. La tasa de mutación fue mayor para los segmentos genómicos codificantes para las proteínas de superficie del virión tales como la HA, NA y M2 de ambos subtipos virales (ver figura 5). En cambio, los valores más bajos de la tasa de mutación se observaron en los segmentos genómicos codificantes para las proteínas internas y accesorias (aquellas que no se encuentran expuestas en la superficie del virión intacto) incluyendo a PB2, PB1, PA, NP, M1, NS1 y NS2. Al comparar la tasa de mutación entre ambos subtipos virales encontramos que las regiones codificantes para PB1, PA, NA, M y NS del subtipo H1N1 presentaban valores superiores a los encontrados en los mismos segmentos de virus con el subtipo H3N2. Por el contrario, la tasa de mutación de los segmentos codificantes para las proteínas PB2, HA y NP de H3N2 demostró ser superior a la de los mismos fragmentos de H1N1.

Proyecciones basadas en la tasa de mutación estimada para cada fragmento y subtipo viral considerando las relaciones filogenéticas estimadas por los árboles de máxima credibilidad generados por Beast permitieron establecer el tiempo al ancestro común más reciente de nuestras secuencias (tMRCA). Dicho análisis estableció que el tMRCA global (considerando a los 8 segmentos genómicos juntos) para los aislados mexicanos de influenza virus con subtipo H1N1 surgió hace aproximadamente 41 años. Para los aislados mexicanos de influenza virus con subtipo H3N2 dicho ancestro se estimó haber surgido hace aproximadamente 33 años. No obstante el tMRCA inferido para los segmentos individuales que codifican para PB1, HA y NP de influenza virus H1N1 mexicanos era más antiguo localizando el surgimiento del ancestro más cercano a una distancia de 45 años previos a la estación invernal 2007-2008. La tabla 4 concentra el tiempo estimado del surgimiento del ancestro común más reciente para cada segmento genómico de cada subtipo viral de nuestros aislados mexicanos.

El algoritmo Beast nos permitió generar 20 árboles filogenéticos ilustrando las relaciones evolutivas de las secuencias nucleotídicas de cada una de las 10 regiones codificantes de cada subtipo viral, incluyendo

nuestras secuencias locales. Para las secuencias de los aislados H1N1 se realizaron en promedio 60 millones de repeticiones, no obstante algunos segmentos como el 6(NA) necesitaron más (hasta 100 millones de iteraciones) mientras que otros como los 7(M1) y 2(PB1) requirieron menos (únicamente 30 millones). Para las secuencias de aislados H3N2 se realizaron 60 millones de iteraciones. Los árboles de máxima credibilidad fueron muestreados al azar cada 1000 iteraciones con lo que en promedio se generaron un total de 60,000 filogramas para cada segmento genómico de cada subtipo viral. De estos 60,000 filogramas, se descartó el 10% inicial y estimó el promedio del último 90% para las gráficas. Los filogramas de ambos subtipos virales adoptaron tanto topologías simples como complejas. La topología simple carece de cladas y hace evidente las relaciones genéticas directas de los ancestros inmediatos a cada secuencia. Este tipo de topología se evidenció para los filogramas de las proteínas PB1, M1, M2, NS1 y NS2 del subtipo H1N1 al igual que de las proteínas HA, NA y M2 de H3N2 (ver figura 6, panel A). Por otro lado, los filogramas con topología compleja presentan cladas con tres o más taxas que presentan mayor similitud entre ellas que con las demás secuencias (ver figura 6, panel B, C y D). Este tipo de topología se hizo evidente para los las regiones codificantes de las proteínas PB2, PA, HA, NP y NA del subtipo H1N1 al igual que para las de las proteínas PB2, PB1, PA, NP, M1, NS1 y NS2 del subtipo H3N2. El análisis de los filogramas de PB2, PA, HA y NA de las secuencias del subtipo H1N1 demostró el surgimiento de dos cladas recientes que divergieron hace aproximadamente  $8.2 \pm 1.19$  años antes de la temporada invernal 2007-2008 (alrededor del año 1999), ver figura 6, panel C. La primer clada demostró estar compuesta por virus de las temporadas invernales 2002-2003 y 2004-2007 mientras que la segunda clada (y la más reciente) se encontraba compuesta por virus de las temporadas 2003-2004 y 2007-2008 (entre ellas los aislados locales).

Un fenómeno similar se demostró para el filograma del segmento 5(NP) de los aislados H1N1, en el cual se hizo evidente la divergencia de las secuencias modernas de sus ancestros aproximadamente 31 años antes de la temporada 2007-2008 (ver figura 6, panel B). En este filograma la primer clada (en azul) incluía a los aislados de 1978-1979, 1980-1981, 1982-1983 y 1984-1985 mientras que la segunda clada correspondía a aislados virales obtenidos a partir de 1986 y hasta el presente (con la excepción de los aislados atípicos correspondientes a las temporadas invernales del 1988-1989 y 1991-1992).

Los filogramas generados para las secuencias de influenzavirus H3N2 se comportaron de manera similar. Dos cladas distintivas fueron observadas en los filogramas de PB2, PB1, PA, NP, M1, NS1 y NS2 de este subtipo viral. Para PB1 y PA una clada correspondió a las temporadas invernales de 1998-1999, 2000-2003 y la otra a temporadas invernales posteriores a 1999. Para las proteínas PB2, NP, M1, NS1 y NS2 las cladas fueron esencialmente similares con discretas variaciones en las temporadas invernales incluidas y en la estimación del tiempo de divergencia. En términos globales y tomando en cuenta a todas las proteínas del subtipo H3N2 el período estimado de divergencia para estas dos cladas fue de hace  $10.96 \pm 0.6$  años antes de la temporada invernal 2007-2008. Los árboles filogenéticos completos se encuentran en el sitio web <https://www.genomica.uaslp.mx/Databases/Influenza.html> para su consulta.

La presencia de temporadas invernales genéticamente atípicas fue identificada en los filogramas del H1N1 más no en los H3N2. Dichas temporadas atípicas aparecen en los filogramas como taxas genéticamente emparentadas a virus temporalmente distantes (usualmente dos o tres temporadas invernales por detrás en el

tiempo). Este comportamiento se hizo evidente en los filogramas de PB1, NP, M1, NS1 y NS2. En el caso de PB1, NP y M1 las temporadas invernales de 1988-1989 y 1991-1992 parecían derivarse de virus surgidos 5 temporadas atrás y alrededor de los años de 1971 y 1975 (figura 6, panel B). Para el segmento genómico 8(NS) la situación fue más compleja ya que la región codificante para NS1 de aislados de 1986-87 parecía derivarse de virus que circularon durante 1981 mientras que la de NS2 correspondiente a aislados de 1988-89 parecía derivarse de virus que circularon durante 1976.

## **Discusión**

La incidencia de detecciones de influenzavirus tipo A (3.9 %) fue semejante a los resultados reportados previamente en un estudio en la población pediátrica de San Luis Potosí (1). Durante esta misma temporada invernal no se detectó la presencia de ningún aislado del influenzavirus B, hecho que ha sido confirmado por reportes previos. No obstante, este hallazgo no debe ser tomado como evidencia de que los virus de la influenza B no circularon a nivel nacional durante esta temporada invernal ya que los síntomas de la infección causada por dichos suelen ser menos severos sin ameritar atención médica (y por ende su detección) (32, 33). Este manuscrito describe el éxito logrado al desarrollar y aplicar herramientas moleculares y de informática para caracterizar la diversidad genética de genomas completos de aislados de influenzavirus humanos. La caracterización genómica de virus de la influenza de circulación nacional no solo nos brinda la oportunidad de evaluar la extensión de la diversidad generada sino también los mecanismos de que hace uso el virus para cambiar y evadir al sistema inmune y las estrategias que utiliza para establecer una infección exitosa en el hospedero humano. De manera adicional, las técnicas moleculares y de bioinformática complementan las estrategias de epidemiología clásica permitiendo realizar estudios capaces de evaluar y/o guiar las estrategias médicas y de salud pública que se implementan para diagnosticar, tratar y prevenir las infecciones por estos virus. Así pues, nuestra iniciativa de caracterización genómica nos permite participar en la generación de información de secuencia requerida para guiar la formulación internacional de la vacuna anual así como con un set integral de habilidades técnicas y científicas para enfrentar a la próxima pandemia o brote de influenzavirus.

Si bien los mecanismos genéticos que determinan el surgimiento de una nueva cepa pandémica de influenzavirus humanos son bien conocidos, poco se sabe acerca de los mecanismos responsables de la diversificación de las cepas en los períodos interpandémicos. Dos mecanismos principales han sido postulados como los principales motores de la diversificación de virus de la influenza, el *shift* y el *drift* antigénico. En este estudio nosotros demostramos la existencia de un nuevo mecanismo de diversidad genética poco conocido y aparentemente muy importante para la diversificación de cepas interpandémicas, la recombinación de segmentos virales. Este mecanismo permite el intercambio de porciones discretas de material genético entre distintas cepas virales. No obstante, otros estudios virológicos serán necesarios para dilucidar si este fenómeno surge como resultado de recombinación homóloga con participación de la maquinaria de reparación del hospedero o si se tratara del resultado de un mecanismo de selección de copia dependiente de la polimerasa viral. De cualquier manera, dicho mecanismo contribuye de manera significativa a la generación de mayor diversidad genómica para estos patógenos virales.

Las secuencias mexicanas en su mayoría demostraron ser más similares a las de otras regiones geográficas pero de la misma temporada invernal 2007-2008. Si bien detectamos evidencia de eventos ancestrales de recombinación que llevaban a la incorporación de secuencias pertenecientes a aislados de temporadas invernales previas, el grado de certeza estadística fue bajo (indicado por porcentajes de árboles permutados entre 10 y 65%). Esto bien pudiera resultar del bajo número de secuencias locales que fueron incluidas en nuestro análisis por lo que esfuerzos futuros se encaminarán a incrementar el número de secuencias caracterizadas para re-evaluar este hallazgo. Aun así, diversos estudios han evidenciado eventos de recombinación similares. En una publicación se describió un evento en el cual el dominio globular de la hemaglutinina de un aislado de influenza A(H1N1) pandémico de 1918 se encontraba codificado por secuencias derivadas de influenza porcinos mientras que el tallo de la proteína correspondía a una secuencia aislada de un hospedero humano (20).

Adicionalmente otros autores identifican algunos eventos recombinatoriales entre diferentes subtipos de cepas aviarias en los segmentos que codifican para HA, PB2, PB1, PA y NP, colocando a estos eventos como importantes mecanismos evolutivos para virus aviarios (28). No obstante, otros estudios contradicen lo anterior sugiriendo en términos generales que los virus de ssRNA en sentido negativo no utilizan frecuentemente la recombinación homóloga como mecanismo evolutivo importante (4, 11).

Durante este estudio se caracterizó la diversidad genética y la dinámica evolutiva de los virus de la influenza A(H1N1) y A(H3N2) circulantes durante la temporada invernal 2007-2008 en el estado de San Luis Potosí, México. Las secuencias generadas representan la más completa y comprensiva información genética jamás realizada para aislados mexicanos que se haya depositado en las bases de datos internacionales hasta el surgimiento de la variante pandémica del 2009.

Debido a que el tipo de modelo evolutivo determina y justifica el tipo de interpretaciones que se pueden hacer en relación a la tasa de mutación viral y acumulo de mutaciones dN/dS, nuestro estudio se enfocó primero a determinar el modelo evolutivo más apropiado haciendo uso de una aplicación en línea previamente solo utilizada para el estudio de HIV. Este acercamiento nos brinda mayor confianza en el tipo de aseveraciones que podemos hacer a partir de nuestros resultados. Los virus de la influenza tienden a ser empujados hacia la diversidad antigénica a consecuencia de las presiones selectivas impuestas por la respuesta inmune de los hospederos comúnmente infectados, y ello lleva al surgimiento de nuevas cepas con mayor capacidad para colonizar a sus hospederos. No obstante, los aislados virales del mismo subtipo suelen permanecer relativamente conservados entre sí acumulando abundantes sustituciones dS en comparación a las dN. Es por ello que las secuencias analizadas en su mayoría mostraron evidencia estadística concordante con una evolución purificadora encaminada a conservar las propiedades funcionales. No obstante, para algunos segmentos genómicos como el 8(NS2) del subtipo H3N2, la mera presencia de tan solo una mutación dN en nuestras secuencias fue suficiente para sugerir una evolución purificante (hacia la diversidad). La inclusión de secuencias nucleóticas de otras partes del mundo pero de la misma temporada invernal nos permitió controlar este fenómeno de muestreo y aclarar el verdadero mecanismo evolutivo subyacente para dicho segmento.

Por otro lado, el análisis de la distribución de mutaciones en los segmentos 1(PB2) y 2(PB1) de ambos subtipos virales demostró la existencia de una muy baja tolerancia al cambio en estas subunidades de la polimerasa, lo cual concuerda con el papel crítico que dichas proteínas juegan en la replicación viral. Esto sugiere además que dichas proteínas en realidad no están sujetas a la presión selectiva impuesta por el sistema inmune por no encontrarse expuestas en la superficie viral como sucede con las proteínas HA y NA (62). Para el resto de las proteínas internas y accesorias codificadas por el genoma viral, la distribución de mutaciones indicó la preponderancia de dS con respecto a dN corroborando la tendencia a la selección purificante (conservación). Aun así las proteínas internas y no estructurales también se encuentran sometidas a cambios resultantes del rearrreglo o recombinación de segmentos genómicos (62). Los eventos recombinatoriales homólogos no se hicieron evidentes al limitar el análisis a secuencias mundiales y por ende podrían representar una característica genómica aparentemente exclusiva de los aislados mexicanos.

Los cambios de las propiedades fisicoquímicas de sitios estratégicos (sitios catalíticos, antigénicos o interfases de polimerización) inducidos por sustituciones no-sinónimas pueden tener repercusiones funcionales que se traduzcan en ventajas o desventajas para el virus. La adquisición de sustituciones en sitios reconocidos por medicamentos obviamente resultará en una ventaja al virus al brindarle resistencia a dicho fármaco. Interesantemente, en algunos aislados locales H1N1 detectamos la presencia de la sustitución I256T, la cual ocurre en uno de los sitios antigénicos mayores de la HA. En otros aislados H3N2 detectamos la presencia de la sustitución L173S en un sitio antigénico mayor de la HA. En ambos casos, la sustitución implica un cambio de las propiedades fisicoquímicas del sitio desde una superficie no-polar hacia una polar. Estudios recientes han demostrado como la afinidad de un anticuerpo por su antígeno (más no su especificidad) está determinada por el grado de superficie no-polar que interviene en la interacción. Subsecuentemente, cambios conformacionales de la superficie externa tanto del anticuerpo como del antígeno llevan a reforzar esta interacción(56). Debido a que las dos sustituciones previamente mencionadas involucran un cambio de superficie no-polar (necesaria para la afinidad del anticuerpo por el antígeno) por una polar, cabe la posibilidad de que dicho cambio haya logrado permearse hacia las cepas circulantes en respuesta a la ventaja biológica que ello supondría para el virus para evadir la respuesta inmune humoral.

La explicación del impacto funcional del resto de las sustituciones que fueron encontradas en HA, no obstante, se dificulta un poco más debido a que no inciden sobre sitios antigénicos ni incurrir en cambios de polaridad o carga. Es decir, a pesar de ser sustituciones no-sinónimas son sustituciones conservadoras (de las propiedades fisicoquímicas de superficie).

El análisis de las sustituciones que se detectaron en la neuraminidasa reveló la presencia de 4 cambios en sitios involucrados en su polimerización. Estas modificaciones se traducen en cambios de polaridad y carga eléctrica en estos aminoácidos pudiendo tener repercusiones importantes en la estabilidad y función del polímero de neuraminidasa. No obstante, en la NA de algunos virus del subtipo H3N2, solo en una de las modificaciones encontradas se localizó un cambio menor en las propiedades fisicoquímicas.

En su mayoría para los aislados mexicanos las proteínas expresadas en la superficie del virión (HA, NA y M2) presentaron una tasa de mutación más elevada respecto a las proteínas internas o accesorias. Dicha tasa mutacional elevada condiciona los niveles de diversidad proteica que le permiten al virus superar a las

defensas del hospedero o restricciones impuestas por fármacos antivirales. La tasa de mutación global para los influenzavirus ha sido estimada en un rango que comprende entre  $1.3 \times 10^{-3}$  a  $3.7 \times 10^{-3}$  s/s/a (36). Las secuencias locales se analizaron en conjunto con las secuencias consenso mundiales de cada temporada invernal y se puede observar valores mucho menores de la tasa de mutación  $1.21 \times 10^{-3}$  a  $4 \times 10^{-3}$  s/s/a para el subtipo H3N2 respecto a lo reportado previamente (desde  $3.8 \times 10^{-3}$  a  $5.72 \times 10^{-3}$  s/s/a para secuencias norteamericanas aisladas entre los años de 1993 y 2005) (50). De manera más específica para este subtipo viral nuestros valores de la tasa mutacional fueron similares a lo reportado previamente para las regiones codificantes de NA y NP. Encontramos valores inferiores de la tasa mutacional con respecto a reportes previos para las regiones de HA, NS1 y NS2 mientras que nuestros valores de la tasa mutacional para PB2, PB1, PA, M1 y M2 fueron superiores a los reportados (17, 23, 30, 40, 41, 43, 64). Estas diferencias se deben al mayor número de secuencias incluidas en los otros estudios brindándole mayor diversidad al análisis de las secuencias nucleótidas. En el caso de nuestras estimaciones para el subtipo H1N1 encontramos valores similares a los reportados para PB2, HA, NP, NS1 y NS2 (24, 25, 50) y valores superiores a los reportados para PB1, PA, NA, M2, y M1 (30, 31, 46, 50). Interesantemente nuestra estimación de la tasa de mutación de la región de la NA de subtipos H1N1 demostró ser la más alta de todos los segmentos ( $7.2 \times 10^{-3}$  sustituciones/sitio/año) y a un nivel equiparable a lo reportado para la variante pandémica 2009 (22). Además en nuestra opinión consideramos que las diferencias encontradas entre nuestras estimaciones y las publicadas previamente para cada segmento pudieran surgir de sesgo en el uso de las distintas bases de datos nucleótidas disponibles. En nuestro caso el análisis incluyó secuencias de distintas regiones del mundo, solamente secuencias únicas eliminando duplicados, de cepas virales aisladas de hospederos humanos e, importantemente, solamente empleando secuencias altamente curadas tomadas de *The Flu Project*. Contrariamente, las estimaciones realizadas por otros estudios previamente publicados emplearon criterios de selección de secuencias más arbitrarios y/o regionales que no permiten obtener una perspectiva global de la velocidad evolutiva de virus mundiales.

La inferencia filogenética es una importante herramienta que permite estimar las relaciones evolutivas que se dan tras el surgimiento y diversificación de patógenos virales emergentes. La tasa de mutación de cualquier segmento genómico (viral, procariota o eucariota) permite calibrar un reloj molecular que permite entonces estimar las fechas en que ocurrieron fenómenos de divergencia o incluso la fecha que hipotéticamente pudo haber surgido el ancestro común más reciente (MRCA) de dicho genoma (13, 49).

Las estimaciones filogenéticas basadas en el análisis de todos los segmentos genómicos pertenecientes a cada subtipo viral colocan a los ancestros comunes más recientes de nuestros aislados virales alrededor del año de 1967 para el subtipo H3N2 y en 1975 para el H1N1. Estos datos concuerdan perfectamente con el surgimiento de la pandemia de 1968 en Hong Kong a causa del subtipo H3N2 y el resurgimiento del subtipo H1N1 alrededor de 1977 en Rusia (55). El virus pandémico de 1918 dejó de circular en el año de 1957 al ser reemplazado por el subtipo H2N2. No obstante tras veinte años el subtipo H1N1 re-emergió en los humanos permaneciendo en circulación como tal hasta el año 2009 en que surgió la nueva variante pandémica triple-rearreglante. Estudios filogenéticos llevados a cabo por otros grupos de investigación basados en los segmentos genómicos 2(PB1), 4(HA) y 5(NP) de la variante pandémica del 2009 colocaron al MRCA de este

subtipo alrededor de 1947, casi 28 años antes de que surgiera la re-emergencia del subtipo H1N1 en 1977. Otros datos reportados recientemente muestran evidencia de que dicho subtipo viral permaneció en latencia en otros hospederos durante al menos 27 años (63).

La tendencia evolutiva del subtipo H3N2 a comportarse como un linaje viral relativamente estable de cambios graduales que permiten vincular en estudios filogenéticos a una temporada estacional con la inmediatamente anterior o subsecuente ha sido reportada previamente por otros grupos de investigación. De manera similar, el surgimiento de cladas dispersas en los árboles filogenéticos del subtipo H1N1 tampoco es un hallazgo novedoso (6, 16, 17, 27, 35, 38, 47). Los filogramas del subtipo H3N2 en su mayoría muestran dos cladas definidas donde destaca una sola clada viral alrededor del 2001 y el 2004 (en algunas secuencias entre 1998 y el 2003) que parece divergir de los aislados virales más recientes entre 10 y 12 años antes de la temporada invernal del 2007 y 2008. Interesantemente nuestras secuencias muestran evidencia de mayor similitud en algunas regiones genómicas con esta clada distintiva en comparación a secuencias más recientes. Nuevamente resaltando la participación de fenómenos recombinatoriales en la evolución de virus circulantes en México.

Si bien los sistemas de vigilancia epidemiológica actualmente disponibles integran la capacidad de distinguir entre los distintos subtipos virales, a la fecha pocos si es que algunos sistemas nacionales poseen la capacidad de detectar el surgimiento de nuevas variantes virales, que sin modificar el subtipo antigénico, incorporen regiones subgenómicas que les brinden ciertas ventajas biológicas a los virus (39). Este fenómeno se hizo patente durante el surgimiento del brote de la influenza pandémica del 2009 durante el cual reacciones anómalas fueron reportados para algunas pruebas moleculares y serológicas. Nuestro grupo de investigación adoptó la responsabilidad de desarrollar herramientas moleculares y de bioinformática que nos permitiesen evaluar y vigilar el surgimiento de precisamente este tipo de eventos no detectables por otros medios convencionales. Uno de los objetivos principales de la vigilancia epidemiológica en materia de influenza tiene que ver con la caracterización genómica de las cepas circulantes para detectar la aparición de nuevas variantes rearrreglantes que pudieran representar una amenaza pandémica. Actualmente las herramientas que hemos desarrollado colocan al estado de San Luis Potosí en la capacidad de detectar no solo variantes rearrreglantes sino también como generador de información importante a nivel mundial en la vigilancia del surgimiento de cepas recombinantes del virus. Hoy en día nuestra aportación a la aplicación de las técnicas moleculares ha venido a mejorar la capacidad regional de respuesta a contingencias epidemiológicas y en concordancia con las recomendaciones y lineamientos expuestos como resultado de la pandemia del 2009 (60, 61). Un ejemplo de eso fue nuestra participación temprana en dicho evento puso a prueba la competencia del grupo multidisciplinario que se había establecido años atrás para la detección molecular de patógenos respiratorios. En respuesta a dicho evento se desarrollaron varias técnicas moleculares para detectar y caracterizar el genoma de éste patógeno viral. Dichas técnicas permitieron no solo detectar al patógeno entre las muestras del brote sino también el llevar a cabo un estudio retrospectivo que buscaba determinar la existencia de dicho virus en los meses previos al surgimiento del brote (21).

Nuestros resultados resaltan la importancia de contar con sistemas de vigilancia epidemiológicos competentes, de alta resolución, redundantes y versátiles para mantener al día la información nacional generada en relación al a diversidad genética de influenzavirus y/o el surgimiento de nuevas cepas o subtipos.

## Referencias

1. **Aranda-Romo, S., A. Comas-Garcia, C. A. Garcia-Sepulveda, A. E. Hernandez-Salinas, M. Pina-Ramirez, and D. E. Noyola.** 2010. Effect of an immunization program on seasonal influenza hospitalizations in Mexican children. *Vaccine* **28**:2550-5.
2. **Bao, Y., P. Bolotov, D. Dernovoy, B. Kiryutin, L. Zaslavsky, T. Tatusova, J. Ostell, and D. Lipman.** 2008. The influenza virus resource at the National Center for Biotechnology Information. *J Virol* **82**:596-601.
3. **Bedford, T., S. Cobey, P. Beerli, and M. Pascual.** 2010. Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). *PLoS Pathog* **6**:e1000918.
4. **Boni, M. F., Y. Zhou, J. K. Taubenberger, and E. C. Holmes.** 2008. Homologous recombination is very rare or absent in human influenza A virus. *J Virol* **82**:4807-11.
5. **Brown, E. G.** 2000. Influenza virus genetics. *Biomed Pharmacother* **54**:196-209.
6. **Bush, R. M., C. A. Bender, K. Subbarao, N. J. Cox, and W. M. Fitch.** 1999. Predicting the evolution of human influenza A. *Science* **286**:1921-5.
7. **Claas, E. C., A. J. van Milaan, M. J. Sprenger, M. Ruiten-Stuiver, G. I. Arron, P. H. Rothbarth, and N. Masurel.** 1993. Prospective application of reverse transcriptase polymerase chain reaction for diagnosing influenza infections in respiratory samples from a children's hospital. *J Clin Microbiol* **31**:2218-21.
8. **Contreras-Trevino, H. I.** 2010. Laboratorio de Genómica Viral y Humana.
9. **Creanza, N., J. S. Schwarz, and J. E. Cohen.** 2010. Intraseasonal dynamics and dominant sequences in H3N2 influenza. *PLoS One* **5**:e8544.
10. **Chang, S., J. Zhang, X. Liao, X. Zhu, D. Wang, J. Zhu, T. Feng, B. Zhu, G. F. Gao, J. Wang, H. Yang, and J. Yu.** 2007. Influenza Virus Database (IVDB): an integrated information resource and analysis platform for influenza virus research. *Nucleic Acids Res* **35**:D376-80.
11. **Chare, E. R., E. A. Gould, and E. C. Holmes.** 2003. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J Gen Virol* **84**:2691-703.
12. **Drake, J. W.** 1993. Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci U S A* **90**:4171-5.
13. **Drummond, A. J., O. G. Pybus, A. Rambaut, R. Forseberg, and A. G. Rodrigo.** 2003. Measurably evolving populations. *Trends Ecol Evol* **18**:8.
14. **Drummond, A. J., and A. Rambaut.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**:214.
15. **Edgar, R. C.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**:113.
16. **Ferguson, N. M., A. P. Galvani, and R. M. Bush.** 2003. Ecological and immunological determinants of influenza evolution. *Nature* **422**:428-33.
17. **Fitch, W. M., R. M. Bush, C. A. Bender, and N. J. Cox.** 1997. Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc Natl Acad Sci U S A* **94**:7712-8.
18. **Fouchier, R. A., V. Munster, A. Wallensten, T. M. Bestebroer, S. Herfst, D. Smith, G. F. Rimmelzwaan, B. Olsen, and A. D. Osterhaus.** 2005. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. *J Virol* **79**:2814-22.
19. **Garten, R. J., C. T. Davis, C. A. Russell, B. Shu, S. Lindstrom, A. Balish, W. M. Sessions, X. Xu, E. Skepner, V. Deyde, M. Okomo-Adhiambo, L. Gubareva, J. Barnes, C. B. Smith, S. L. Emery, M. J. Hillman, P. Rivaller, J. Smagala, M. de Graaf, D. F. Burke, R. A. Fouchier, C. Pappas, C. M. Alpuche-Aranda, H. Lopez-Gatell, H. Olivera, I. Lopez, C. A. Myers, D. Faix, P. J. Blair, C. Yu, K. M. Keene, P. D. Dotson, Jr., D. Boxrud, A. R. Sambol, S. H. Abid, K. St George, T. Bannerman, A. L. Moore, D. J. Stringer, P. Blevins, G. J. Demmler-Harrison, M. Ginsberg, P. Kriner, S. Waterman, S. Smole, H. F. Guevara, E. A. Belongia, P. A. Clark, S. T. Beatrice, R. Donis, J. Katz, L. Finelli, C. B. Bridges, M. Shaw, D. B. Jernigan, T. M. Uyeki, D. J. Smith, A. I. Klimov, and N. J. Cox.** 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* **325**:197-201.
20. **Gibbs, M. J., J. S. Armstrong, and A. J. Gibbs.** 2001. Recombination in the hemagglutinin gene of the 1918 "Spanish flu". *Science* **293**:1842-5.
21. **Gomez-Gomez, A., M. Magana-Aquino, C. Garcia-Sepulveda, U. R. Ochoa-Perez, R. Falcon-Escobedo, A. Comas-Garcia, S. Aranda-Romo, H. I. Contreras-Trevino, P. V. Jimenez-Rico, M. A. Banda-Barbosa, F. Dominguez-Paulin, J. M. Bernal-Blanco, L. F. Perez-Gonzalez, and**

- D. E. Noyola.** 2010. Severe pneumonia associated with pandemic (H1N1) 2009 outbreak, San Luis Potosi, Mexico. *Emerg Infect Dis* **16**:27-34.
22. **Goni, N., A. Fajardo, G. Moratorio, R. Colina, and J. Cristina.** 2009. Modeling gene sequences over time in 2009 H1N1 influenza A virus populations. *Virology* **6**:215.
23. **Gorman, O. T., W. J. Bean, Y. Kawaoka, I. Donatelli, Y. J. Guo, and R. G. Webster.** 1991. Evolution of influenza A virus nucleoprotein genes: implications for the origins of H1N1 human and classical swine viruses. *J Virol* **65**:3704-14.
24. **Gorman, O. T., W. J. Bean, Y. Kawaoka, and R. G. Webster.** 1990. Evolution of the nucleoprotein gene of influenza A virus. *J Virol* **64**:1487-97.
25. **Gorman, O. T., R. O. Donis, Y. Kawaoka, and R. G. Webster.** 1990. Evolution of influenza A virus PB2 genes: implications for evolution of the ribonucleoprotein complex and origin of human influenza A virus. *J Virol* **64**:4893-902.
26. **Gutierrez, J. P., and S. M. Bertozzi.** 2005. [Influenza vaccination in the elderly population in Mexico: economic considerations]. *Salud Publica Mex* **47**:234-9.
27. **Hay, A. J., V. Gregory, A. R. Douglas, and Y. P. Lin.** 2001. The evolution of human influenza viruses. *Philos Trans R Soc Lond B Biol Sci* **356**:1861-70.
28. **He, C. Q., Z. X. Xie, G. Z. Han, J. B. Dong, D. Wang, J. B. Liu, L. Y. Ma, X. F. Tang, X. P. Liu, Y. S. Pang, and G. R. Li.** 2009. Homologous recombination as an evolutionary force in the avian influenza A virus. *Mol Biol Evol* **26**:177-87.
29. **Hoffmann, E., J. Stech, Y. Guan, R. G. Webster, and D. R. Perez.** 2001. Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* **146**:2275-89.
30. **Ito, T., O. T. Gorman, Y. Kawaoka, W. J. Bean, and R. G. Webster.** 1991. Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. *J Virol* **65**:5491-8.
31. **Kawaoka, Y., S. Krauss, and R. G. Webster.** 1989. Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *J Virol* **63**:4603-8.
32. **Kim, H. W., C. D. Brandt, J. O. Arrobio, B. Murphy, R. M. Chanock, and R. H. Parrott.** 1979. Influenza A and B virus infection in infants and young children during the years 1957-1976. *Am J Epidemiol* **109**:464-79.
33. **Kim, Y. H., H. S. Kim, S. H. Cho, and S. H. Seo.** 2009. Influenza B virus causes milder pathogenesis and weaker inflammatory responses in ferrets than influenza A virus. *Viral Immunol* **22**:423-30.
34. **Kumar, S., M. Nei, J. Dudley, and K. Tamura.** 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* **9**:299-306.
35. **Landry, M. L., and D. Ferguson.** 2003. Suboptimal detection of influenza virus in adults by the Directigen Flu A+B enzyme immunoassay and correlation of results with the number of antigen-positive cells detected by cytospin immunofluorescence. *J Clin Microbiol* **41**:3407-9.
36. **Lavenu, A., M. Leruez-Ville, M. L. Chaix, P. Y. Boelle, S. Rogez, F. Freymuth, A. Hay, C. Rouzioux, and F. Carrat.** 2006. Detailed analysis of the genetic evolution of influenza virus during the course of an epidemic. *Epidemiol Infect* **134**:514-20.
37. **Leitner, T., and J. Albert.** 1999. The molecular clock of HIV-1 unveiled through analysis of a known transmission history. *Proc Natl Acad Sci U S A* **96**:10752-7.
38. **Levin, S. A., J. Dushoff, and J. B. Plotkin.** 2004. Evolution and persistence of influenza A and other diseases. *Math Biosci* **188**:17-28.
39. **Lin, Y. P., V. Gregory, M. Bennett, and A. Hay.** 2004. Recent changes among human influenza viruses. *Virus Res* **103**:47-52.
40. **Lindstrom, S., A. Endo, S. Sugita, M. Pecoraro, Y. Hiromoto, M. Kamada, T. Takahashi, and K. Nerome.** 1998. Phylogenetic analyses of the matrix and non-structural genes of equine influenza viruses. *Arch Virol* **143**:1585-98.
41. **Lindstrom, S. E., Y. Hiromoto, R. Nerome, K. Omoe, S. Sugita, Y. Yamazaki, T. Takahashi, and K. Nerome.** 1998. Phylogenetic analysis of the entire genome of influenza A (H3N2) viruses from Japan: evidence for genetic reassortment of the six internal genes. *J Virol* **72**:8021-31.
42. **Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray.** 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**:152-60.

43. **Ludwig, S., U. Schultz, J. Mandler, W. M. Fitch, and C. Scholtissek.** 1991. Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. *Virology* **183**:566-77.
44. **McHardy, A. C., and B. Adams.** 2009. The role of genomics in tracking the evolution of influenza A virus. *PLoS Pathog* **5**:e1000566.
45. **Nicholson, K. G., J. M. Wood, and M. Zambon.** 2003. Influenza. *Lancet* **362**:1733-45.
46. **Okazaki, K., Y. Kawaoka, and R. G. Webster.** 1989. Evolutionary pathways of the PA genes of influenza A viruses. *Virology* **172**:601-8.
47. **Plotkin, J. B., J. Dushoff, and S. A. Levin.** 2002. Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc Natl Acad Sci U S A* **99**:6263-8.
48. **Poddar, S. K.** 2002. Influenza virus types and subtypes detection by single step single tube multiplex reverse transcription-polymerase chain reaction (RT-PCR) and agarose gel electrophoresis. *J Virol Methods* **99**:63-70.
49. **Pybus, O. G., and A. Rambaut.** 2009. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet* **10**:540-50.
50. **Rambaut, A., O. G. Pybus, M. I. Nelson, C. Viboud, J. K. Taubenberger, and E. C. Holmes.** 2008. The genomic and epidemiological dynamics of human influenza A virus. *Nature* **453**:615-9.
51. **Russell, C. A., T. C. Jones, I. G. Barr, N. J. Cox, R. J. Garten, V. Gregory, I. D. Gust, A. W. Hampson, A. J. Hay, A. C. Hurt, J. C. de Jong, A. Kelso, A. I. Klimov, T. Kageyama, N. Komadina, A. S. Lapedes, Y. P. Lin, A. Mosterin, M. Obuchi, T. Odagiri, A. D. Osterhaus, G. F. Rimmelzwaan, M. W. Shaw, E. Skepner, K. Stohr, M. Tashiro, R. A. Fouchier, and D. J. Smith.** 2008. Influenza vaccine strain selection and recent studies on the global migration of seasonal influenza viruses. *Vaccine* **26 Suppl 4**:D31-4.
52. **Salomon, R., and R. G. Webster.** 2009. The influenza virus enigma. *Cell* **136**:402-10.
53. **Scholtissek, C., W. Rohde, V. Von Hoyningen, and R. Rott.** 1978. On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* **87**:13-20.
54. **Shaw, M. W., N. H. Arden, and H. F. Maassab.** 1992. New aspects of influenza viruses. *Clin Microbiol Rev* **5**:74-92.
55. **Smith, G. J., J. Bahl, D. Vijaykrishna, J. Zhang, L. L. Poon, H. Chen, R. G. Webster, J. S. Peiris, and Y. Guan.** 2009. Dating the emergence of pandemic influenza viruses. *Proc Natl Acad Sci U S A* **106**:11709-12.
56. **Sundberg, E. J.** 2009. Structural basis of antibody-antigen interactions. *Methods Mol Biol* **524**:23-36.
57. **Suzuki, Y.** 2006. Natural selection on the influenza virus genome. *Mol Biol Evol* **23**:1902-11.
58. **Taubenberger, J. K., and D. M. Morens.** 2006. 1918 Influenza: the mother of all pandemics. *Emerg Infect Dis* **12**:15-22.
59. **Taubenberger, J. K., A. H. Reid, R. M. Lourens, R. Wang, G. Jin, and T. G. Fanning.** 2005. Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**:889-93.
60. **Tenorio-Abreu, A., J. M. Eiros, E. Rodriguez, J. F. Bermejo, M. Dominguez-Gil, T. Vega, J. Castrodeza, and R. O. de Lejarazu.** 2009. [Influenza surveillance by molecular methods.]. *Rev Esp Quimioter* **22**:214-20.
61. **Wang, R., and J. K. Taubenberger.** 2010. Methods for molecular surveillance of influenza. *Expert Rev Anti Infect Ther* **8**:517-27.
62. **Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers, and Y. Kawaoka.** 1992. Evolution and ecology of influenza A viruses. *Microbiol Rev* **56**:152-79.
63. **Wertheim, J. O.** 2010. The re-emergence of H1N1 influenza virus in 1977: a cautionary tale for estimating divergence times using biologically unrealistic sampling dates. *PLoS One* **5**:e11184.
64. **Xu, X., N. J. Cox, C. A. Bender, H. L. Regnery, and M. W. Shaw.** 1996. Genetic variation in neuraminidase genes of influenza A (H3N2) viruses. *Virology* **224**:175-83.

**TABLAS**

**Tabla 1.** Oligonucleótidos para la subtipificación del virus estacional de influenza A.

Nombre	Segmento	Secuencia (5'-3')	Posición	Amplicón	Referencia
Flu-HA1-F	HA1	GAA-ATT-TGC-TAT-GGC-TGA-CGG-GR	490-509	171	(21)
Flu-HA1-R		GAC-ACT-ACA-GAG-ACA-TAA-GCA-TTT-TC	655-635		(48)
Flu-HA3-F	HA3	CAG-CAA-AGC-CTA-CAG-CAA-MTG-TT	272-292	236	(21)
Flu-HA3-R		GGC-ATA-GTC-ACG-TTC-AAT-GCT-G	503-485		(48)
Flu-HA5-F	HA5	AAA-CTC-CAA-TRG-GGG-CGA-TAA-AC	894-914	344	(21)
Flu-HA5-R1		CAA-CGG-CCT-CAA-ACT-GAG-TGT	1245-1225		(48)
Flu-HA5-R2		CCA-ACA-GCC-TCA-AAC-TGA-GTG-T	1245-1226		(21)
Flu-NA1Nw-F	NA1	ACT-CAR-GAG-TCT-GAA-TGT-G	696-714	409	(21)
Flu-NA1Nw-R1		GTC-CTT-CCT-ATC-CAA-ACA-CC	1086-1105		
Flu-NA1Nw-R2		GTT-CTC-CCG-AGC-CAG-ATA-CC	1086-1105		
Flu-NA2-F1	NA2	GGA-AAA-TCG-TTC-ATA-CTA-GCA-MAT-TG	779-804	176	(21)
Flu-NA2-F2		GGG-AAA-ATC-GTT-CAT-ATT-AGC-ACA-TTG	779-805		(48)
Flu-NA2-R		AGC-ACA-CAT-AWC-TGG-AAA-CAA-TGC	955-932		(21)

Las secuencias de los oligonucleótidos reversa mostrados en esta tabla representan el complemento reversa de la secuencia presente en los alineamientos.

**Tabla 2.** Oligonucleótidos para la amplificación genómica y sub-genómica de los fragmentos.

Nombre	Segmento	Secuencia	Posición	Amplicón (bp)	Referencias
PB2-F	PB2	TAT-TGG-TCT-CAG-GGA-GCG-AAA-GCA-GGT-C	24	2366	(29)
SwPB2-F		ATG-GAG-AGA-ATA-AAA-GAA-C	56	833	(8)
PB2-R1		GCT-AGT-GGA-TCT-GCY-G	855		
PB2-F2		RAT-GTA-CAC-TCC-AGG-T	779	942	
PB2-R2		RAT-TTC-TGA-TGA-TCC-A	1692		
PB2-F3		GGA-RGT-MAG-TGA-AAC-AC	1602	791/2366 <sup>A</sup>	
PB2-R		ATA-TGG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GTC-GTT-T	2358		
PB1-F		PB1	TAT-TCG-TCT-CAG-GGA-GCG-AAA-GCA-GGC-A	21	2342
SwPB1-F1	ATG-GAT-GTC-AAT-CCG-ACT-C		50	833	(8)
PB1-R1	CAT-TAY-CYC-CAA-CYG		847		
PB1-F2	CAC-RAT-GAC-CAA-AGA-Y		723	849	
PB1-R2	CTC-CAT-GCT-RAA-ATT-RG		1540		
PB1-F3	GAG-CAA-AAA-GAA-GTC-Y		1477	887/2342 <sup>B</sup>	
PB1-R	ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GCA-TTT		2332		
PA-F	PA		TAT-TCG-TCT-CAG-GGA-GCG-AAA-GCA-GGT-AC	19	2233
SwPA-F1		GGA-AGA-CTT-TGT-GCG-AC	47	856	(8)
PA-R1		CCA-TCA-GSA-GGA-ATT-TKG	868		
PA-F2		GCT-RCA-TTG-AGG-GCA-AG	763	897	
PA-R2		TCC-CAT-TTR-TGT-GGY-TC	1626		
PA-F3		AAG-AGG-GAA-GGM-GAA-A	1514	741/2233 <sup>C</sup>	
PA-R		ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GTA-CTT	2222		
HA-F		HA	TAT-TCG-TCT-CAG-GGA-GCA-AAA-GCA-GGG-G	32	862
HA1-R1	H1	AAG-CCT-CTA-CTC-ART-GCG	864	986	
HA1-F2		CCR-GGG-ATA-CWA-TAA-TA	820		
NS-R	HA	ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GGT-GTT-TT	1783	881	(29)
HA-F		TAT-TCG-TCT-CAG-GGA-GCA-AAA-GCA-GGG-G	22		
HA3-R1	H3	ATT-ATT-GAG-CTT-TTC-CC	872	952	(8)
HA3-F2		AAC-AGC-ACA-GGG-AAT-C	832		
NS-R	HA	ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GGT-GTT-TT	1750	816	(29)
NP-F	NP	TAT-TCG-TCT-CAG-GGA-GCA-AAA-GCA-GGG-TA	23		1566
SwNP-F		ATG-GCG-TCT-CAA-GG	67		
NP-R1		TGA-GCA-ACT-GAT-CCT-CTC	853	870/1566 <sup>D</sup>	(8)
NP-F2		GGA-YCA-AGT-GAG-AGA-AAG	722		
NP-R		ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GGT-ATT-TTT	1555		(29)
NA-F	NA	TAT-TGG-TCC-AGG-GAG-CAA-AAG-CAG-GAG-T	23	821	(8)
NA-R1		AYY-TTY-CCC-TYY-TCR-AT	812		
NA-F2		ACM-CAR-GAG-TCW-GAA-T	732	780	
NA-R		ATA-TGG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GAG-TTT-TTT	1475		(29)
M-F		M	TAT-TCG-TCT-CAG-GGA-GCA-AAA-GCA-GGT-AG	21	1126
M-R	ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GTA-GTT-TTT		1117		
NS-F	NS	TAT-TCG-TCT-CAG-GGA-GCA-AAA-GCG-GGT-G	19	965	
NS-R		ATA-TCG-TCT-CGT-ATT-AGT-AGA-AAC-AAG-GGT-GTT-TT	949		

Las secuencias de los oligonucleótidos reversa mostrados en esta tabla representan el complemento reversa de la secuencia presente en los alineamientos. Las letras en negritas corresponde a la secuencia viral y el resto corresponde a las secuencias de reconocimiento para las enzimas de restricción *BsmBI* (Bm) y *BsaI* (Ba). El tamaño de amplicón indicados con las letras A, B, C y D resultan de la combinación de los oligonucleótidos F y R respectivamente para los fragmentos genómicas completos PB2, PB1, PA y NP utilizados en la primera PCR de la estrategia de caracterización sub-genómica.

**Tabla 3.** Resumen de homología regional y de eventos recombinatorios encontrados en las secuencias de influenza virus H1N1 y H3N2 mexicanos.

H1N1			H3N2		
Segmento	Localización	Homología	Segmento	Localización	Homología
1(PB2)	1-575, 842-2277	07-08	1(PB2)	1-1554,1671-2277	07-08
	576-841	03-04		1555-1670	04-05
2(PB1)	1-1451, 1540-2277	07-08	2(PB1)	1-2277	07-08
	1452-1539	02-03	3(PA)	1-2148	07-08
3(PA)	1-2148	07-08	4(HA)	1-1695	07-08
4(HA)	1-1695	07-08	5(NP)	1-1154	07-08
5(NP)	1-1494	07-08		1155-1494	04-05
6(NA)	1-1410	07-08	6(NA)	1-217	04-05
7(M)	1-727, 832-1007	07-08		218-1410	07-08
	728-831	06-07	7(M)	1-276	00-01
8(NS)	1-209	04-05		277-1007	07-08
	210-273	06-07	8(NS)	1-214, 485-864	07-08
	274-864	07-08		215-484	00-08

La homología corresponde a la temporada invernal a la cual las secuencias mexicanas generadas mostraron mayor similitud al resto del mundo.

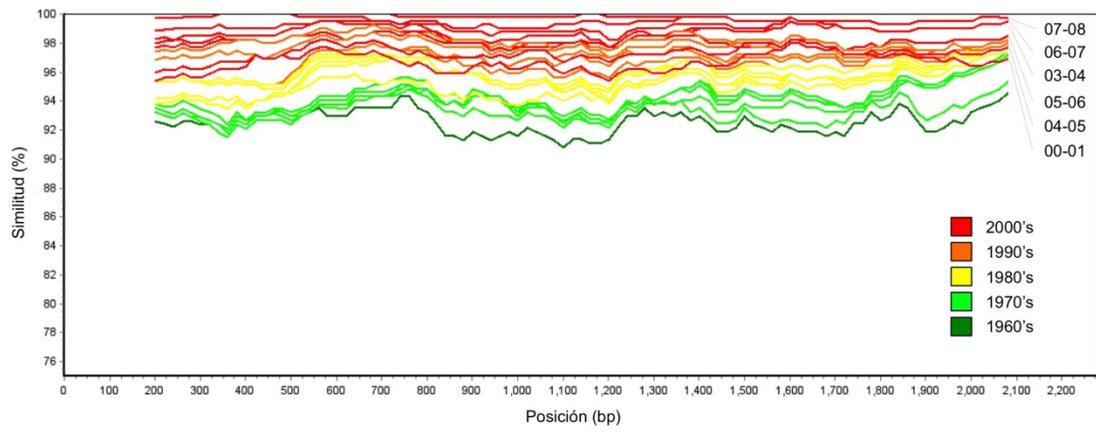
**Tabla 4.** tMRCA expresado en años de los aislados mexicanos para ambos subtipos

Proteína	tMRCA (años)	
	H1N1	H3N2
PB2	33.1 (32.0-35.2)	41.2 (40.0-42.9)
PB1	45.8 (34.2-60.0)	40.4 (40.0-41.2)
PA	32.7 (32.0-34.2)	42.0 (40.0-44.6)
HA	45.0 (37.1-52.1)	40.2 (40.0-40.7)
NP	45.8 (34.8-60.9)	41.4 (40.0-43.5)
NA	34.3 (32.0-38.9)	42.5 (40.0-45.4)
M1	34.5 (32.0-38.9)	41.6 (40.0-43.6)
M2	32.2 (32.0-33.2)	40.3 (40.0-41.6)
NS1	32.7 (32.0-33.4)	41.4 (40.0-44.0)
NS2	34.1 (32.0-36.9)	42.1 (40.0-45.2)

El valor encontrado corresponde al periodo de tiempo estimado en años en un intervalo de confianza al 95 %.

# FIGURAS

## Figura 1



**Figura 2**

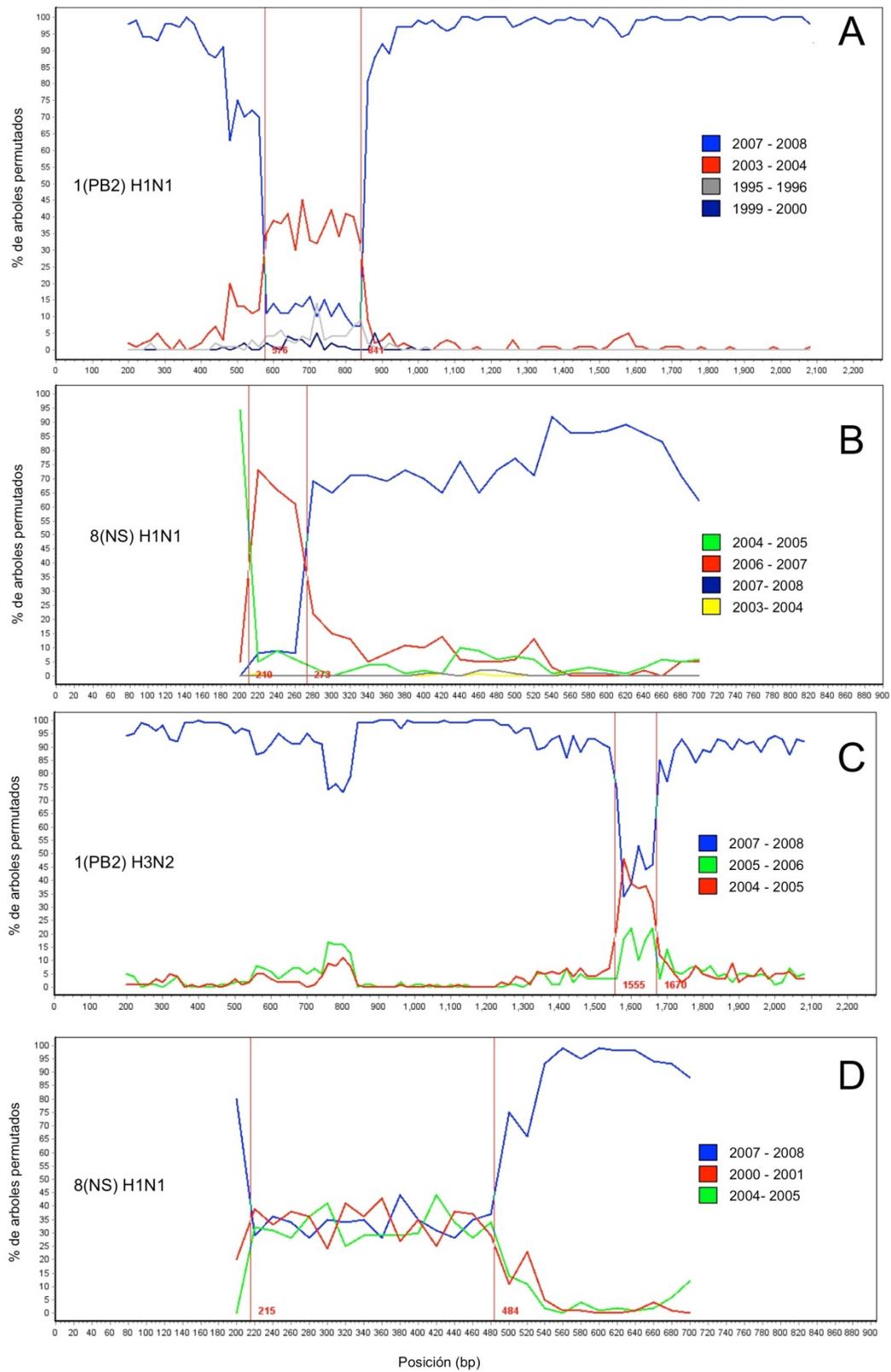


Figura 3

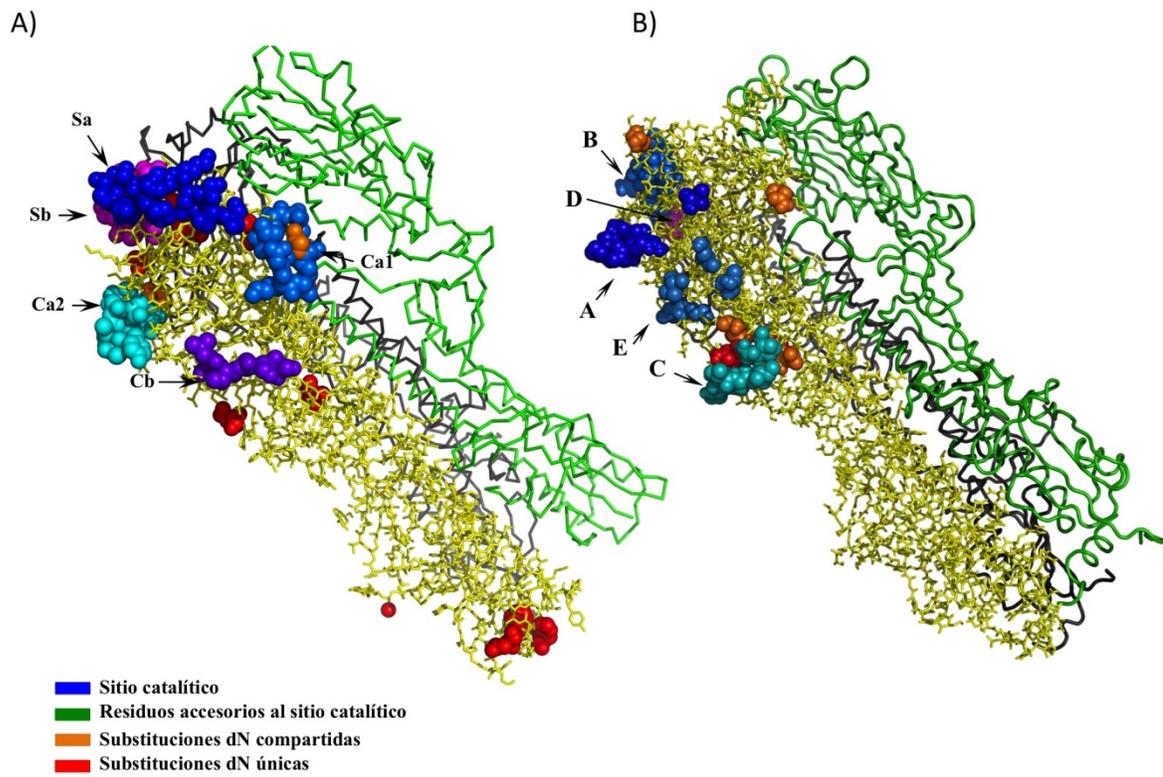
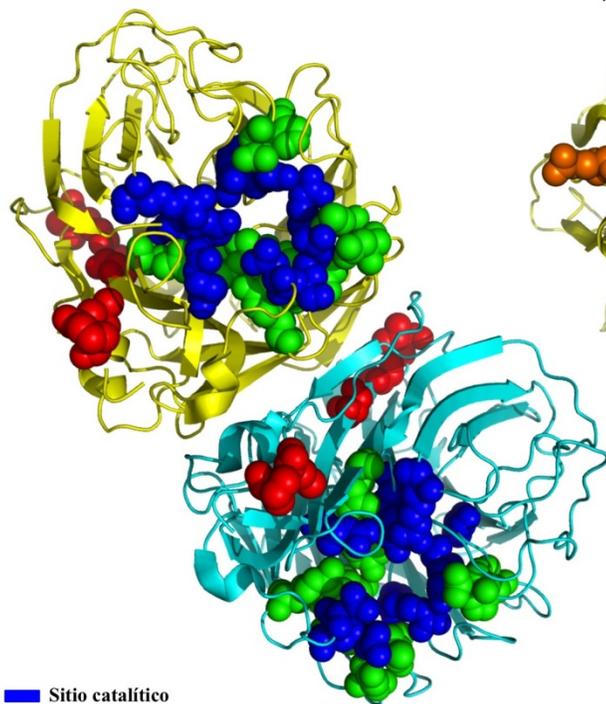
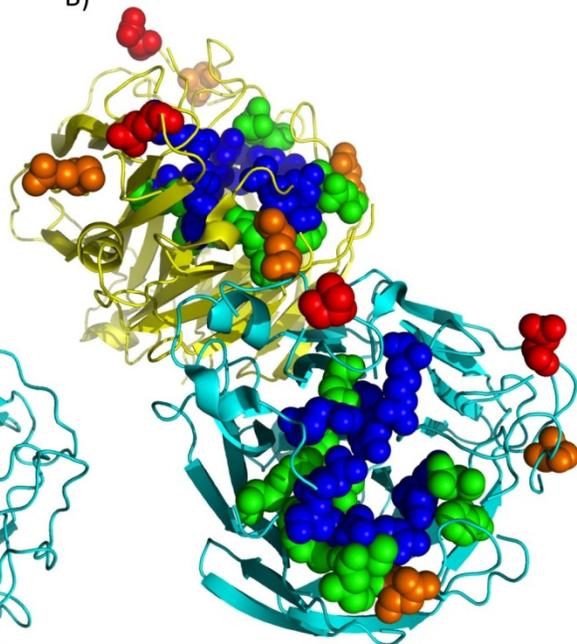


Figura 4

A)



B)



- Sitio catalítico
- Residuos accesorios al sitio catalítico
- Substituciones dN compartidas
- Substituciones dN únicas

Figura 5

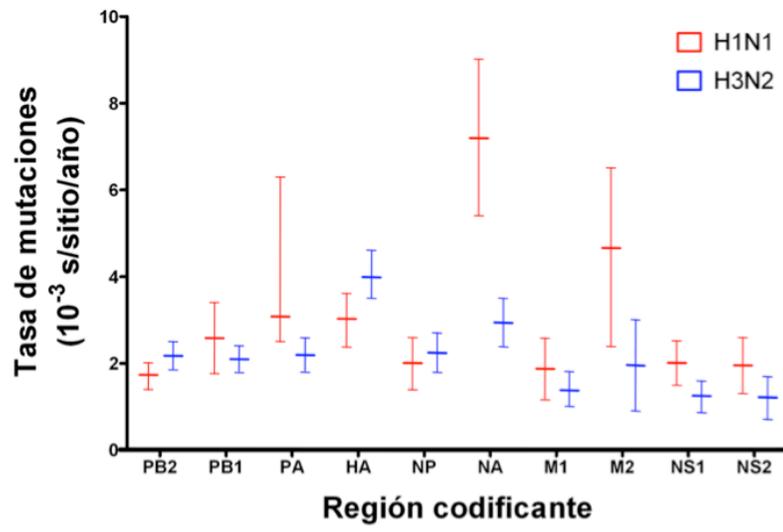
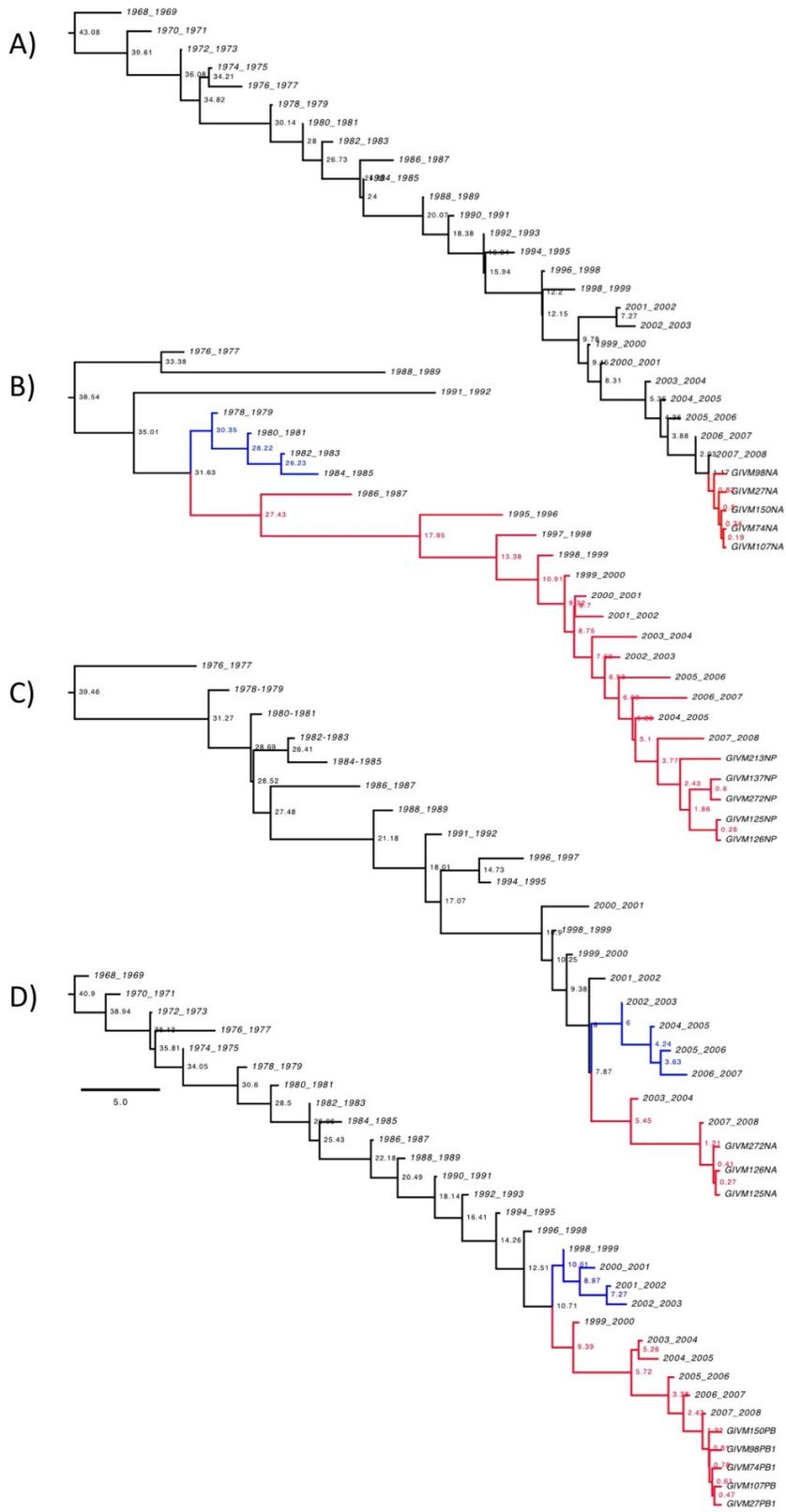


Figura 6



## Leyendas de tablas y figuras

**Figura 1. Análisis de homología temporal.** Análisis realizado por ventana desplazante con SimPlot comparando las secuencias del fragmento genómico 1(PB2) de influenza virus H3N2 mexicanos contra las secuencias consenso de influenza virus H3N2 de distintas temporadas invernales.

**Figura 2. Eventos de recombinación en las secuencias mexicanas.** a) Evento de recombinación encontrado en el fragmento 1 codificante para la subunidad PB2 de la polimerasa de influenza virus H1N1 mexicanos b) Evento de recombinación encontrado en el fragmento 8 codificante para la proteína NS de influenza virus H1N1 mexicanos. c) Evento de recombinación encontrado en el fragmento 1 codificante para la subunidad PB2 de la polimerasa de influenza virus H3N2 mexicanos. d) Evento de recombinación encontrado en el fragmento 8 codificante para la proteína NS de influenza virus H3N2 mexicanos.

**Figura 3. Hemaglutinina de los subtipos H1 y H3.** Mapeo tridimensional de las sustituciones no sinónimas dN de los aislados virales mexicanos. En color rojo se muestran las dN únicas y en naranja las sustituciones compartidas. a) Hemaglutinina H1, en diferentes tonos de azul y magentas se indican los sitios antigénicos Sa, Sb, Ca1, Ca2 y Cb. b) Hemaglutinina H3, en diferentes tonos de azul y magentas se indican los sitios antigénicos A, B, C, D y E.

**Figura 4. Neuraminidasa de los subtipos N1 y N2.** Mapeo tridimensional de las sustituciones no sinónimas dN de los aislados virales mexicanos. En color azul se muestran los residuos del sitio catalítico y en verde los residuos accesorios del sitio catalítico. Las sustituciones dN únicas se muestran en color rojo y en naranja las sustituciones encontradas en otros aislados. a) Neuraminidasa N1, b) Neuraminidasa N2.

**Figura 5. Tasa de mutaciones expresada como sustituciones por sitio por año para las 10 regiones codificantes de los influenza virus A(H1N1) y A(H3N2).**

**Figura 6. Filogramas representativos de los subtipos H1N1 y H3N2.** a) Filograma de influenza virus con topología simple de cladas de máxima credibilidad generado para 25 secuencias consenso del fragmento 6(NA) para cada temporada invernal de influenza virus A(H3N2) y las 5 secuencias localmente generadas (GIVM, clada en rojo). b) Filograma de influenza virus con topología compuesta de cladas de máxima credibilidad generado para 20 secuencias consenso del fragmento 5(NP) para temporadas invernales de influenza virus H1N1 así como 5 secuencias localmente secuenciadas (GIVM). La presencia de dos cladas (en rojo y azul) es clara al igual que la existencia de dos outsiders (1988-1989 y 1991-1992). c) Cladas terminales del gen codificante para la neuraminidasa de influenza virus A(H1N1). Filograma de cladas de máxima credibilidad generado para 20 secuencias consenso del fragmento 6(NA) para temporadas invernales de influenza virus A(H1N1) además se incluyen 3 secuencias localmente secuenciadas (GIVM). La topología de los filogramas generados para los fragmentos PB2, PA y HA de influenza virus H1N1 fue similar claramente

distinguiendo a dos cladas terminales (en rojo y azul) que divergieron aproximadamente 8 años antes de la temporada invernal 2007-2008. d) Filograma ejemplar de las cladas presentes en influenzavirus A(H3N2) generado para 25 secuencias consenso del fragmento 2(PB1) de influenzavirus A(H3N2). Se incluyen 5 secuencias localmente generadas (GIVM). Se distinguen dos cladas terminales (en rojo y azul) que divergieron poco más de 10 años antes de la temporada invernal 2007-2008.