



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

**INSTITUTO DE BIOTECNOLOGÍA**

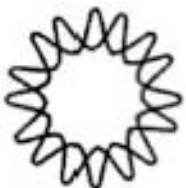
**PROGRAMA DE MAESTRÍA Y DOCTORADO  
EN CIENCIAS BIOQUÍMICAS**

**EVOLUCIÓN DE LOS VIRUS DE  
INFLUENZA A EN MÉXICO**

**T E S I S**  
QUE PARA OBTENER EL GRADO DE  
MAESTRA EN CIENCIAS BIOQUÍMICAS  
P R E S E N T A:

**MARINA ESCALERA ZAMUDIO**

TUTOR: DR. PAVEL ISA



CUERNAVACA, MORELOS

JUNIO, 2011



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## AGRADECIMIENTOS

Este trabajo fue realizado en el laboratorio del Dr. Carlos Arias y de la Dra. Susana López, en el Instituto de Biotecnología, UNAM. El trabajo fue apoyado por los donativos 5549 y I0110/184/09 del Consejo Nacional de Ciencia y Tecnología (CONACyT) de México, y PICOSI09-209 del Instituto de Ciencia y Tecnología del Distrito Federal. Asimismo, la alumna fue apoyada por la beca de posgrado 295899 CONACyT-México.

Este trabajo contó con la ayuda y apoyo técnico del Dr. Ricardo Grande y de la M.C. Verónica Jiménez, de la Unidad de Secuenciación Masiva de la Universidad Nacional Autónoma de México (IBT-UNAM), y gracias al personal de Biología Molecular y de la Unidad de Aislamiento de Virus de CPA y al laboratorio BSL-3 del Ministerio de Salud del Estado de Veracruz. Asimismo, se agradece al Dr. Joel Vásquez del Instituto Nacional De Enfermedades Respiratorias (INER), a la M.C. Irma López, del Instituto de Diagnóstico y Referencia Epidemiológicos (InDRE) y al personal del Hospital General Benito Juárez por su ayuda en la obtención de las muestras. Gracias también a la unidad de cómputo y cluster del Instituto de Biotecnología, UNAM.

Gracias a mi tutor Pavel Isa, por su paciencia y todo el apoyo. Gracias también, al Dr. Carlos Arias y a la Dra. Susana López por haberme dado la oportunidad de trabajar con ellos y conocerlos. Gracias a la Dra. Selene Zárate y al Dr. Lorenzo Segovia, parte de mi comité tutorial, y al Dr. Pablo Vinuesa por su apoyo y valiosas observaciones.

Gracias a todos los integrantes del laboratorio Arias-López por toda la paciencia y ayuda; en especial, agradezco a Ma. de los Dolores Soto y Ana Georgina Cobián, por que este trabajo también es de ellas. También agradezco a Ana Paola Carranco, Miguel Angel Martinez y Luis Paulin por su ayuda adicional en programación.

# ÍNDICE

<b>INTRODUCCION</b>	5
<b>ANTECEDENTES</b>	6
I. CLASIFICACIÓN Y ESTRUCTURA DE LOS VIRUS DE INFLUENZA A	6
I.1 PROTEÍNAS DEL COMPLEJO DE LA POLIMERASA VIRAL	8
I.2 LA PROTEÍNA PB1-F2 Y N40	8
I.3 PROTEÍNAS DE LA MATRIZ, MI Y M2	9
I.4 LA NUCLEOPROTEÍNA (NP)	9
I.5 LAS PROTEÍNAS NO ESTRUCTURALES NS1 y NEP/NS2	10
I.6 LA HEMAGLUTININA (HA) Y LA NEURAMINIDASA (NA)	10
II. TROPISMO VIRAL Y DETERMINANTES DE VIRULENCIA	11
III. ECOLOGÍA Y EVOLUCIÓN DE LOS VIRUS DE INFLUENZA A	13
IV. EPIDEMIAS Y PANDEMIAS DE INFLUENZA	16
IV.1 LOS VIRUS A/H1N1/2009 PANDÉMICOS	16
IV.2 LOS VIRUS A/H3N2 ESTACIONALES	18
<b>HIPOTESIS</b>	23
<b>OBJETIVOS</b>	23
<b>OBJETIVOS PARTICULARES</b>	23
<b>MÉTODOS</b>	24
I. OBTENCIÓN DE MUESTRAS, PROPAGACIÓN Y PURIFICACIÓN DE VIRUS	24
II. EXTRACCIÓN DE RNA GENÓMICO VIRAL Y AMPLIFICACIÓN POR RT-PCR MULTISEGMENTO (M-RTPCR)	24
III. CONSTRUCCIÓN DE BIBLIOTECAS Y SECUENCIACIÓN	25
IV. ANÁLISIS FILOGENÉTICO Y DETECCIÓN DE MUTACIONES	29
IV. 1 VIRUS A/H3N2 ESTACIONALES	30
IV. 2 VIRUS A/H1N1/2009 PANDÉMICOS	30

<b>RESULTADOS Y DISCUSIÓN</b>	33
I. ANÁLISIS DE LAS SECUENCIAS	33
II. ANÁLISIS EVOLUTIVO DE LOS VIRUS A/H1N1 PANDÉMICOS	35
II.1 LA TRAYECTORIA EVOLUTIVA DEL AISLADO PORCINO Y DE LOS VIRUS DE LA GLORIA	38
II.2 POLIMORFISMOS PRESENTES EN A/swine/4/Mexico/09	39
II.3 ANÁLISIS EVOLUTIVO DE LOS VIRUS A/H3N2 ESTACIONALES DE MÉXICO	41
<b>CONCLUSIONES</b>	46
I. ANÁLISIS DE LAS SECUENCIAS	46
II. VIRUS A/H1N1 PANDÉMICOS	46
III. VIRUS A/H3N2 ESTACIONALES	47
<b>BIBLIOGRAFÍA</b>	48
<b>ANEXO</b>	57

## INTRODUCCIÓN

La influenza es una enfermedad infecciosa respiratoria aguda, con un elevado impacto para la salud pública mundial, y genera un alto índice de morbilidad y de mortalidad a nivel global. Se estima que cada año el virus infecta al 5-15% de la población mundial, lo que resulta en aproximadamente 500,000 muertes (OMS, 2003). Aunque en México no existen estadísticas que reflejen el impacto de esta enfermedad, se estima que los índices de mortalidad y de morbilidad que se generan son similares a los reportados en Estados Unidos, ocasionando en promedio 100 mil hospitalizaciones y hasta 40 mil muertes por año (OMS, 2003).

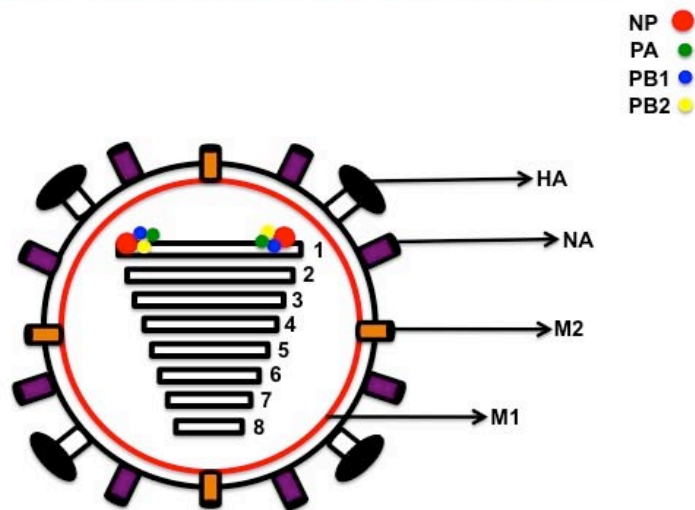
Con el fin de estar preparados ante el surgimiento de nuevos brotes de influenza con potencial pandémico, la OMS ha recomendado a todos los países contar con un sistema de vigilancia y de monitoreo en su territorio. En México, los estudios sobre los virus que han circulado en territorio nacional, son escasos. Entender la dinámica evolutiva de los virus de influenza A juega un papel clave en la vigilancia epidemiológica de los mismos, ya que a partir de este tipo de estudios es que podemos analizar los patrones en el comportamiento evolutivo de los virus y hacer predicciones de su futuro comportamiento biológico. El análisis molecular nos permite obtener información valiosa sobre las características moleculares particulares de ciertas cepas, como pueden ser las mutaciones asociadas a la virulencia y/o a la resistencia a los fármacos. Asimismo, nos permite monitorear el surgimiento de nuevos virus que se generan por el intercambio de segmentos génicos completos (rearreglantes).

## ANTECEDENTES

### I. CLASIFICACIÓN Y ESTRUCTURA DE LOS VIRUS DE INFLUENZA A

Los virus de influenza pertenecen a la familia Orthomyxoviridae, definida como virus envueltos que poseen un genoma de RNA de cadena sencilla fraccionado en segmentos con polaridad negativa (Palese & Shaw 2006; Strauss & Strauss 2008). Existen tres géneros distintos de virus de influenza: A, B y C. Aunque los tres grupos poseen similitudes morfológicas, difieren estructuralmente. El virus de influenza A es una de las causas más comunes de infecciones respiratorias severas de origen viral, y es el género viral de influenza con mayor impacto en la salud pública. Hay varios sub-tipos de virus de influenza A, que son clasificados de acuerdo a las proteínas de superficie que presentan: hemaglutinina (HA) y neuraminidasa (NA). Hasta ahora, se han identificado 16 sub-tipos distintos de hemaglutininas (H1-H16) y 9 distintas neuraminidasas (N1-N9), que se definen en base a ensayos de inhibición de hemaglutinación (para la HA), o de inhibición de la actividad de la sialidasa (para la NA) mediados por anticuerpos. Hasta ahora, no se han encontrado sub-tipos de influenza B o C (Palese & Shaw 2006).

Las partículas virales de Influenza A poseen una estructura compleja: al microscopio electrónico los virus tienen una apariencia pleomórfica, con partículas aproximadamente esféricas de un diámetro promedio de 100 nm y partículas elongadas de hasta 300 nm de extensión (Palese & Shaw 2006; Strauss & Strauss 2008). Las partículas virales poseen una envoltura lipídica, derivada de la membrana de la célula infectada, la cual tiene tres proteínas virales embebidas: la hemaglutinina (HA), la neuraminidasa (NA) y la proteína de matriz 2 (M2). La proteína de matriz 1 (M1) se encuentra justo por debajo de la envoltura viral. En el centro de las partículas virales, se encuentran los complejos de ribonucleoproteínas (RNPs) compuestos por un segmento del genoma viral asociado a varias nucleoproteínas (NP) (Fig.1). A estos complejos, se asocian a su vez, los complejos de la polimerasa viral que están constituidos por las polimerasas básicas 1 y 2 (PB1 y PB2) y la polimerasa ácida (PA). A partir de la purificación y análisis estructural de los RNPs, se ha concluido que cada RNP forma una doble hélice y se dobla sobre sí misma. En total, hay ocho RNP empaquetados en unidades individuales, y se distribuyen simétricamente en el centro de las partículas virales (Palese & Shaw 2006).



**Figura 1. Estructura General Del Virus De Influenza A**

El virus está cubierto por una capa lipídica, de la cual se proyectan las proteínas hemaglutinina (HA, en negro) y la neuraminidasa (NA, en púrpura), y la proteína de matriz 2 (M2, en naranja). Por debajo de la capa lipídica, se localiza la proteína de matriz 1 (M1, en rojo), la cual se asocia a los complejos ribonucleoprotéicos, compuestos por el RNA viral y proteínas asociadas (NP, en rojo; polimerasa ácida, PA, en verde; polimerasas básicas 1, PB1 en azul y polimerasas básicas 2, PB2 en amarillo).

El genoma del virus lo constituyen ocho segmentos de RNA, con tamaños que van desde los 2350 nucleótidos hasta los 890 nucleótidos (en orden de tamaño: PB2, PB1, PA, HA, NP, NA, M, NS). En total, el genoma mide aproximadamente 13,600 bases (Palese & Shaw 2006; Strauss & Strauss 2008). Cada uno de los segmentos genómicos codifica una o dos proteínas. El gen PB1 tiene dos marcos de lectura alternos que codifican las proteínas PB1-F2 y N40, mientras que el gen MP codifica dos proteínas: la M1 y M2. El gen más pequeño, NS, codifica las proteínas NS1 y NS2/NEP. Todos los genes virales tienen 13 y 12 nucleótidos hacia los extremos 5' y 3' que corresponden a UTRs (por sus siglas en inglés, *untranslated regions*) altamente conservadas y que contienen las señales de unión para la RNA polimerasa viral y para la encapsidación del genoma viral (Palese & Shaw 2006; Strauss & Strauss 2008).



## I.1 PROTEÍNAS DEL COMPLEJO DE LA POLIMERASA VIRAL

Las proteínas PA, PB1 y PB2 forman parte del complejo de la polimerasa viral, y participan en el proceso de la replicación y de la transcripción del genoma viral. La PB1 es una RNA polimerasa dependiente de RNA que cataliza la adición de nucleótidos durante la transcripción del genoma viral, y también posee actividad de endonucleasa. La PB2 se une al CAP de los pre-mRNA celulares y los recluta para usarlos para la transcripción de los mRNA virales en el proceso de *CAP-snatching*, además de que contribuye a la estabilización del complejo de la polimerasa viral, e induce la actividad de endonucleasa de PB1. Aunque no se ha descrito una función particular para PA, se sabe que las mutaciones en esta proteína disminuyen la eficiencia de la transcripción y de la replicación viral. Por lo tanto, se asume que PA juega un papel importante en ambos procesos.

## I.2 LA PROTEÍNA PB1-F2 Y N40

El gen PB1 posee también un marco de lectura alterno que codifica para la proteína PB1-F2. La proteína PB1-F2 es una proteína de 87-89 aminoácidos que no se expresa en todas las cepas de Influenza A, y se ha demostrado que es un inductor de la apoptosis en las células del sistema inmune y contribuye a la virulencia diferencial de las cepas (Palese & Shaw 2006; Strauss & Strauss 2008). Durante la infección de los macrófagos, PB1-F2 se localiza en la mitocondria e interactúa con factores que alteran la permeabilidad de la misma, como el canal de aniones dependiente de voltaje 1 (VDAC1) y el translocador de adenin-nucleótido 3 (ANT3). Dicha interacción promueve la desestabilización de la membrana mitocondrial y la liberación del citocromo C (Chen et al, 2001; Zamarin et al, 2005). También, se ha reportado que PB1-F2 exagera la respuesta pro-inflamatoria durante la infección viral primaria en ratones y aumenta la frecuencia y la severidad de las neumonías bacterianas secundarias (McAuley et al, 2007), aunque se desconoce el mecanismo de acción.

Recientemente, se demostró que existe un tercer polipéptido codificado por el segmento génico PB1, denominado N40 que se sintetiza a través del uso diferencial de codones del mRNA para PB1, lo que dirige la síntesis de un polipéptido con la región N-terminal truncada (Wise et al, 2009). Dicho polipéptido carece de la función de transcriptasa, pero es capaz de interactuar con el complejo de la polimerasa viral en el entorno celular (Wise et al, 2009). Aunque la función de dicho péptido no es

imprescindible para la viabilidad del virus, se postula que podría jugar un papel en la replicación viral

### I.3 PROTEÍNAS DE LA MATRIZ, M1 Y M2

El análisis estructural de M1 revela que ésta es una proteína dimérica, cuyos monómeros poseen dominios globulares compuestos por  $\alpha$ -hélices unidos entre sí por un dominio sensible a proteasas. Los dímeros de M1 se polimerizan formando una estructura reticular por debajo de la membrana lipídica. La M1 es una de las proteínas más abundantes en las partículas virales, y funciona como un puente entre los componentes del núcleo viral y las proteínas de la membrana (Palese & Shaw 2006; Strauss & Strauss 2008). Se sabe que M1 se une con los complejos de RNPs y con la proteína NEP/NS2, al mismo tiempo que interactúa con las colas citoplásmicas de las glucoproteínas de superficie HA y NA.

La proteína tetramérica transmembranal M2 es la menos abundante de la superficie viral y posee un ectodominio pequeño corto, un dominio transmembranal y un endodominio con motivos lipídicos que es blanco de fosforilación (Palese & Shaw 2006; Strauss & Strauss 2008). La M2 funciona como un canal iónico que se activa a un pH bajo, y cuya función principal es el transporte de protones del endosoma hacia el interior del virión. Este proceso es esencial para favorecer la liberación de los complejos de RNPs de los demás componentes virales, y permitir la importación nuclear para que se lleve a cabo la transcripción del genoma viral. La proteína M2 es el blanco de acción de los fármacos antivirales del tipo de los adamantanos, los cuales actúan al bloquear la actividad de canal iónico esta proteína (Palese & Shaw 2006).

### I.4 LA NUCLEOPROTEÍNA (NP)

La proteína NP es rica en argininas, y se une a RNA mediante interacciones electrostáticas no específicas para formar los RNPs. El RNA viral (vRNA) se enrolla alrededor de la NP, donde aproximadamente 24 nucleótidos completan una vuelta alrededor de cada monómero de NP. Asimismo, los monómeros de NP se oligomerizan otorgando un mayor nivel de organización a los RNPs. Asimismo, la NP posee señales de localización nuclear (NLS) que interactúan con proteínas celulares, como las importinas nucleares (como los complejos de karioferina  $\alpha/\beta$ ). Por lo tanto, la NP también participa en la transportación nuclear de los RNPs desde el citoplasma al núcleo, una vez que el proceso de desnudamiento se ha completado (Palese & Shaw 2006).

## I.5 LAS PROTEÍNAS NO ESTRUCTURALES NS1 y NEP/NS2

La NS1 y la NEP/NS2 son proteínas no estructurales que cumplen una función importante durante el ciclo replicativo viral. NEP/NS2 participa en el transporte de los nuevos complejos de RNPs del núcleo hacia el citoplasma, una vez que se ha completado la replicación del genoma viral, lo que permite el ensamblaje de nuevas partículas. Dicho proceso involucra la interacción de NEP/NS2 con receptores de exportación celulares como Crm1 y con porinas nucleares (Palese & Shaw 2006). Por otro lado, la NS1 es una proteína no estructural que se localiza en el núcleo de la célula durante la infección, y está constituida por dímeros con un dominio de unión a RNA de doble cadena, y un dominio efector PDZ capaz de unirse a varios factores celulares y modular la respuesta inmune del huésped (Jackson et al, 2008).

## I.6 LA HEMAGLUTININA (HA) Y LA NEURAMINIDASA (NA)

La HA es una proteína ampliamente estudiada, ya que juega un papel clave en el reconocimiento de la célula huésped y en la entrada de las partículas virales (Palese & Shaw 2006; Strauss & Strauss 2008). El ácido siálico (AS) es un monosacárido de 9 carbonos que se encuentra unido a muchos glucoconjugados de la membrana celular, y funciona como el principal receptor que los virus de influenza utilizan para entrar a las células. La HA cumple con tres funciones importantes que son: el reconocimiento del AS, mediando la entrada del virus hacia la célula; participa en el proceso de fusión de la membrana viral con la membrana endosomal; y es el principal blanco antigénico. La HA, como su nombre lo indica, tiene la capacidad de aglutinar eritrocitos. La hemaglutinina reconoce las sialoglucoproteínas presentes de forma abundante en la membrana de los eritrocitos. La interacción entre la HA y los AS da lugar a cúmulos de eritrocitos, fenómeno conocido como hemaglutinación. Esta propiedad puede ser bloqueada por anticuerpos. La inhibición de la hemaglutinación se ha utilizado para la clasificación de los diferentes subtipos del virus de Influenza A (Neumann et al, 2009; Wright et al, 2006).

La HA es una proteína trimérica transmembranal de 220 kDa, que posee una forma de bastón con una cabeza globular. La HA se sintetiza en forma del precursor H0, que presenta un péptido señal que le permite transportarse hacia el retículo endoplásmico. Ahí, dicha proteína es blanco de modificaciones post-traduccionales tales como glucosilación y lipidación. A un pH bajo, como el que se presenta dentro de las vesículas endosomales, ocurre un cambio conformacional en la proteína H0, que la vuelve

susceptible al corte por proteasas del tejido infectado, a partir de lo cual se generan los fragmentos HA1 y HA2. La proteólisis de H0 es esencial para la infección viral, ya que expone un motivo hidrofóbico localizado hacia el extremo amino-terminal de HA2, conocido como el péptido de fusión. Este motivo es responsable de mediar la fusión que ocurre entre la envoltura viral y la membrana endosomal, para permitir la salida de los RNPs hacia el citoplasma (Stegmann et al, 1990). La cabeza globular de la HA interviene en el reconocimiento de los receptores en las células huésped y contiene los epítopes que son reconocidos por anticuerpos neutralizantes. La HA de los virus tipo H1 presenta cuatro diferentes sitios antigénicos denominados Sa, Sb, Ca (Ca1 y Ca2) y Cb, mientras que la HA de los virus tipo H2 y H3 presentan cinco sitios antigénicos denominados A, B, C, D y E (Wiley et al, 1981).

La NA es una glucoproteína homotetramérica integral de membrana, que además de ser un segundo blanco antigénico, cumple un papel importante en la diseminación del virus, ya que es una sialidasa (Palese & Shaw 2006; Wright et al, 2006). La actividad de sialidasa de la NA permite la liberación de las partículas virales una vez que se completó el ciclo replicativo. La NA promueve la remoción de moléculas de AS unidas a las HA en las membranas virales, impidiendo la agregación de los virus en el proceso de gemación. También elimina las moléculas de AS presentes en la mucosa, evitando que los virus queden pegados a las células en el proceso de salida y de este modo favorece la diseminación viral (Palese & Shaw 2006; Wright et al, 2006). La NA posee cuatro motivos distintos de glucosilación y un sitio activo de unión a AS compuesto por nueve residuos conservados. Esta proteína también presenta una región transmembranal y una cola citoplasmática altamente conservada. Asimismo, la NA es el blanco de los antivirales oseltamivir y zanamivir, fármacos inhibidores de la actividad de sialidasa de esta proteína.

## II. TROPISMO VIRAL Y DETERMINANTES DE VIRULENCIA

El AS puede tener diferentes tipos de enlaces; por ejemplo,  $\alpha 2,3$  y  $\alpha 2,6$ . Dependiendo del sub-tipo de HA, AS con distintos tipos de enlace pueden ser reconocidos. Por lo tanto, el tropismo en la infección es determinado tanto por el tipo de enlace en los AS presentes en las células como el tipo de HA de la cepa viral (Palese & Shaw, 2007). Las células del tracto respiratorio superior en humanos poseen moléculas de AS con uniones tipo  $\alpha 2,6$ , mientras solo células del tracto respiratorio inferior expresan AS  $\alpha 2,3$  (Wright et al, 2006). Por otro lado, las células del tracto intestinal de aves presentan, en su mayoría, AS con uniones tipo  $\alpha 2,3$  (Palese & Shaw 2006; Wright et al, 2006). Esto ocasiona que los virus

que se han adaptado a la infección en humanos (H1, H2 y H3 principalmente) reconozcan principalmente AS  $\alpha 2,6$ , mientras que los virus que infectan a las aves reconocen AS  $\alpha 2,3$  (Palese & Shaw, 2007). La expresión diferencial del tipo de receptores en el tracto respiratorio de los humanos puede explicar la baja capacidad de infección de algunas cepas de origen aviar. Sin embargo, cuando ocurre una infección con un virus de origen aviar en humanos (por ejemplo, con virus tipo H5N1), puede resultar en una enfermedad más grave, dada la profunda localización de las células que se infectan (de Jong et al, 2006; Tran et al, 2004). De manera contrastante, los puercos poseen AS con ambos tipos de enlaces ( $\alpha 2,6$  y  $\alpha 2,3$ ) en el tracto respiratorio superior, por lo que son susceptibles a la co-infección tanto con cepas de origen aviar como de origen humano. Esto ha llevado a postular que los puercos podrían funcionar como un huésped intermediario donde se pueden generar nuevas variantes; asimismo, representan un nicho donde virus de origen aviar se pueden adaptar al reconocimiento de AS tipo  $\alpha 2,6$ , y por ende a la infección a humanos (Ma et al, 2009; Strauss & Strauss, 2008).

La virulencia es un rasgo genético complejo, y se sabe que está influenciada directamente por mutaciones en las proteínas virales. Hasta ahora, se han caracterizado diferentes mutaciones, particularmente en las proteínas PB1, PB2, HA, NA M2 y NS1, que juegan un papel importante en determinar la virulencia y la patogénesis (Lycett et al, 2009). Por ejemplo, diferentes mutaciones pueden otorgar una mayor eficiencia replicativa; la capacidad de adaptación a la infección de nuevos huéspedes; la resistencia a fármacos antivirales y la modulación de la respuesta inmune. Hay una amplia gama de trabajos que se han dedicado a explorar las mutaciones en los diferentes segmentos génicos que tienen consecuencias directas sobre la virulencia (Lycett et al, 2009). Sin embargo, es difícil obtener conclusiones generales sobre los factores de virulencia a partir de estudios individuales. Esto se debe, en parte, a que los resultados obtenidos para los virus aviares no son completamente extrapolables a virus humanos; además, existen discrepancias entre los datos teóricos y los datos experimentales. Por lo tanto, la interpretación de este tipo de datos debe ser cautelosa.

Se ha demostrado que los segmentos génicos PB1, HA y NA de la cepa pandémica de influenza A/H1N1/1918 son los determinantes clave en la virulencia en un modelo de infección en ratón (Pappas et al, 2008). Además, se conocen algunos aminoácidos en la HA que son responsables de mediar el reconocimiento de AS con diferentes tipos de enlace. Por ejemplo, los residuos de glutamina y de glicina en las posiciones 226 y 228 de las HA tipo H1, H2 y H3 permiten el reconocimiento de receptores AS  $\alpha 2,3$ , mientras que los residuos de leucina y de serina en las mismas

posiciones permiten que los mismos virus reconozcan AS  $\alpha$ 2,6 (Barnett et al, 1999; Connor et al, 1994; Naeve et al, 1984; Rogers et al, 1989). Por otro lado, la mutación más frecuente que confiere resistencia al oseltamivir en las NA tipo N1 y N2, es la H274Y (Colman, 2009). También, se ha demostrado que algunas mutaciones en PB2 son factores de virulencia; por ejemplo, se ha demostrado que la transmisión eficiente de virus de influenza A en humanos depende de la mutación E627K, que confiere una mejor capacidad de replicación en mamíferos a 37°C (Gastaminza et al, 2003, Hatta et al, 2007; Massin et al, 2001).

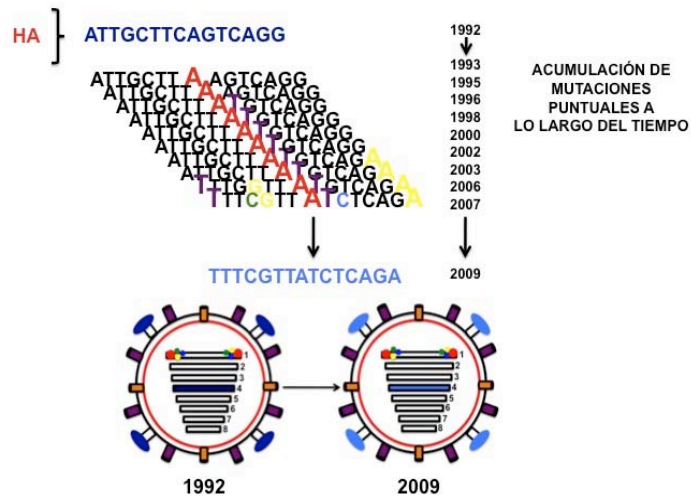
Asimismo, la NS1 puede inhibir mecanismos de la respuesta inmune del huésped, favoreciendo la replicación viral. Por ejemplo, se sabe que la NS1 inhibe la señalización de interferón tipo 1 (INF- $\alpha/\beta$ ) al interactuar con RNA de doble cadena y bloquear la respuesta inmune (Hatada et al, 1999). Se ha descrito que la presencia de un ácido glutámico en la posición 92 de esta proteína en algunos virus, está relacionada con la inhibición de la respuesta de INF- $\alpha/\beta$  (Seo et al, 2002). Asimismo, los últimos cuatro aminoácidos en el carboxilo terminal de la proteína NS1 (Glu-Ser-Glu-Val) constituyen un dominio PDZ que le confiere alta patogenicidad. Los virus de baja patogenicidad poseen una proteína truncada que carece del dominio PDZ, o tienen la secuencia no funcional Arg-Ser-Glu-Val (Jackson et al, 2008). Por otro lado, las mutaciones más frecuentes en M2 asociadas a la resistencia a adamantanos son S31N L26F, V27A, A30T, G34E (Deyde et al, 2007, Schmidtke et al, 2006). Se ha encontrado que estas mutaciones, particularmente la S31N, suelen ser fijadas y ser predominantes en ciertas poblaciones virales (Bright et al, 2006).

### III. ECOLOGÍA Y EVOLUCIÓN DE LOS VIRUS DE INFLUENZA A

Los virus de influenza A infectan a una gran variedad de vertebrados, incluyendo humanos, aves, porcinos, equinos, caninos, felinos y mamíferos acuáticos (Palese & Shaw 2006; Wright et al, 2006). La tasa de mutación en los virus se define como el número de cambios por genoma por ciclo replicativo (Drake 1993), mientras que la tasa de evolución o de sustitución ( $\mu$ ) se refiere al número de cambios que son fijados en la población viral, medida en sitio por sitio por año (Futuyma, 1998). La baja fidelidad de la replicación del RNA polimerasa viral resulta en una tasa de mutación del genoma de influenza A de aproximadamente  $2.0 \times 10^{-6}$  y una tasa de evolución de  $7.3 \times 10^{-5}$  (Nobusawa et al, 2006; Drake 1993). Se ha propuesto que las aves acuáticas silvestres son el reservorio natural de estos virus, ya que virus de todos los sub-tipos de HA y NA se

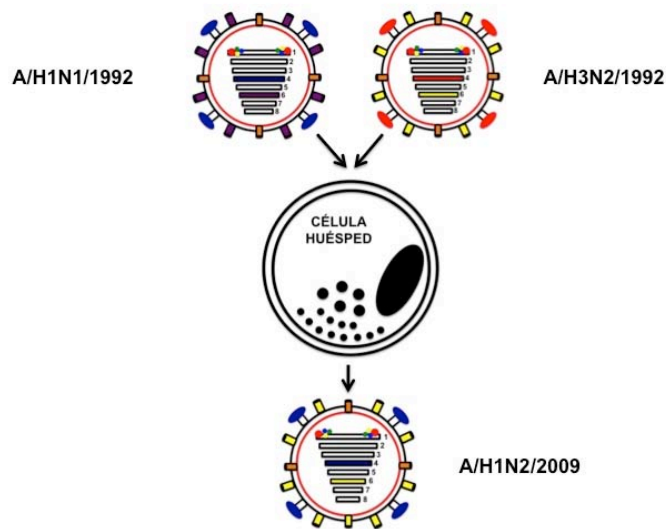
mantienen en circulación en dichos huéspedes. El análisis filogenético de los virus de influenza A permite postular que los virus que infectan a mamíferos surgieron a partir de cepas virus de origen aviar, ya que la tasa de evolución es mucho menor para virus que circulan en aves silvestres en comparación con aquellos que circulan en mamíferos (Palese & Shaw 2006; Wright et al, 2006). Esto sugiere que hay estasis evolutiva en estos huéspedes. Las cepas que circulan en mamíferos, al presentar mayores tasas de evolución, presentan un mayor número de sustituciones no sinónimas que puede fijarse (Webster et al, 1992). Para las cepas de influenza A que circulan en humanos, la tasa de evolución en cada segmento del genoma difiere, lo que refleja una presión de selección diferencial sobre cada proteína viral. Por ejemplo, en los virus tipo H1N1, las proteínas HA y NA, y en menor medida M2, tienen mayor tasa de evolución que las proteínas PB2, PB1, PA, NP y M1 (Webster et al, 1992).

Los virus de influenza A evolucionan a partir de selección clonal de sub-linajes que co-circulan en periodos de tiempo reducidos (Webster et al, 1992). Existen dos mecanismos principales de evolución en los virus de influenza que permiten cambiar y re-infectar continuamente a sus huéspedes; dichos mecanismos son la deriva antigénica y el cambio antigénico (Palese & Shaw 2006; Wright et al, 2006). La deriva antigénica ocurre como resultado de la acumulación lenta y progresiva de mutaciones puntuales que confieren cambios menores en las proteínas virales. Las mutaciones pueden ocurrir en diferentes proteínas, pero se fijan principalmente en los blancos antigénicos HA y NA, por la selección positiva de variantes capaces de evadir al sistema inmune (Fig. 2). Por otro lado, en el proceso de cambio antigénico, nuevas variantes surgen como consecuencia de re-arreglos genéticos entre dos o más cepas virales que co-circulan en una misma población (Fig. 3). La co-infección de un mismo huésped promueve el intercambio de segmentos entre cepas.



**Figura 2. Deriva Antigénica**

La deriva antigénica ocurre como el resultado de la acumulación lenta y progresiva de mutaciones puntuales que confieren cambios menores en las proteínas virales. Algunas mutaciones (representadas como cambios de color en los nucleótidos de la secuencia), pueden resultar ventajosas y podrían fijarse en las proteínas virales (por ejemplo, en la HA representada en azul claro).



**Figura 3. Cambio Antigénico**

El cambio antigénico ocurre cuando dos virus (por ejemplo, A/H1N1/1992 y A/H3N2/2009) co-infectan una misma célula huésped, e intercambian segmentos completos. Esto resulta en el surgimiento de un nuevo virus rearrreglante (en este caso, A/H1N2/2009), que posee una combinación de los segmentos génicos de las cepas parentales (H1 de A/H1N1/1992, NA de A/H3N2/2009).



## IV. EPIDEMIAS Y PANDEMIAS DE INFLUENZA

Una pandemia se define como la ocurrencia de una enfermedad infecciosa que afecta manera simultánea a personas en varios países, con una mortalidad elevada significativa en relación a la proporción de la población infectada (OMS, 2009). Por otro lado, una epidemia se define como la ocurrencia de una enfermedad infecciosa que afecta a un número de individuos superior al esperado en una población durante un tiempo determinado (OMS, 2009). Sin embargo, una epidemia se puede extender para dar lugar a una pandemia. Las pandemias de influenza en humanos se han asociado a cepas del tipo H1N1, H2N2, y H3N2; pero, desde 1997, la introducción a humanos vía zoonótica de cepas aviarias de sub-tipo H5, H7, H9 y H7, han generado brotes con potencial pandémico. Aunque aún no se ha reportado la transmisión vía humano-humano, la continua circulación y amplia distribución de estas cepas virulentas, resulta en una amenaza (Stephenson et al, 2006).

Históricamente, hay registro de tres pandemias de influenza A que ocurrieron en el siglo XX: la de 1918, la de 1957 y la de 1968. Dichas pandemias tuvieron grandes repercusiones en la sociedad al causar altas tasas de morbilidad y de mortalidad (Palese & Shaw 2006). La pandemia de 1918 fue ocasionada por un virus tipo H1N1, mientras que la de 1957 fue originada por un virus tipo H2N2. La pandemia de 1968 fue provocada por un virus tipo H3N2, que reemplazó a los virus tipo H2N2 que circularon en años anteriores. En 1977, después de 20 años, los virus tipo H1N1 resurgieron y desde entonces, co-circulan con los virus de influenza tipo H3N2 de manera continua en las poblaciones humanas. En 2009, se identificó la circulación de un virus nuevo de origen porcino tipo H1N1, que dio lugar a la primera pandemia de influenza de este siglo.

De cada tipo viral, se han detectado diferentes clados o linajes que han circulado a lo largo del tiempo. Un clado o linaje se define como un grupo de especies (o cepas virales) que descienden directamente de un ancestro (Futuyma, 1998). Asimismo, dos clados que divergen de un mismo descendiente se denominan linajes hermanos, y un grupo de clados o linajes hermanos que agrupan juntos se denominan un supergrupo o cluster.

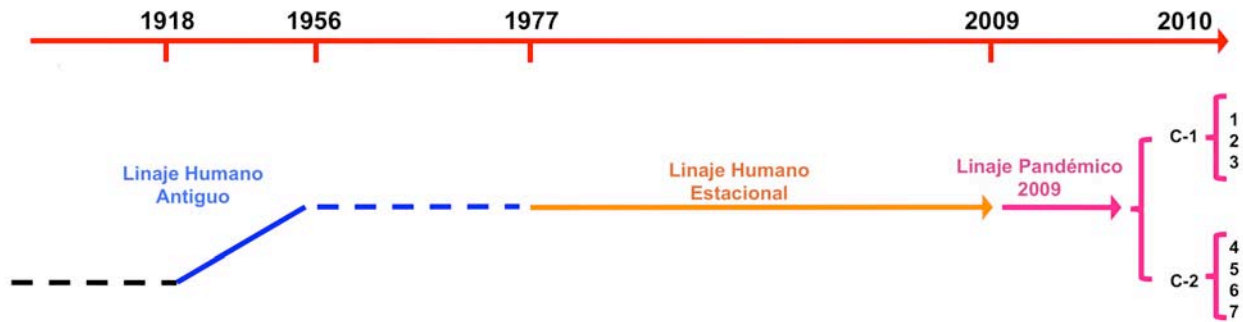
### IV.1 LOS VIRUS A/H1N1/2009 PANDÉMICOS

Durante marzo y principios de abril de 2009, los virus de origen porcino tipo A/H1N1/2009 se extendieron por todo el mundo, provocando la primera pandemia de influenza del siglo

XXI (Fig. 4) (Garten et al, 2009; Fraser et al, 2009). Se ha estimado que el virus pandémico A/H1N1/2009 probablemente se transmitió a los seres humanos durante el segundo semestre de 2008, y que cada segmento génico podría haber circulado en puercos desde hace 9 a 17 años, dependiendo del gen, antes de detectarse en el brote inicial (Smith et al, 2009). Asimismo, Smith et al. también encontraron virus H1N2 rearreglantes que circularon en Asia de manera previa al 2009, los cuales ya tenían 7 de los 8 genes virales pertenecientes a los virus A/H1N1/2009 pandémicos (Fig. 4) (Smith et al, 2009).

Poco después que emergieron, los virus A/H1N1/2009 pandémicos se diferenciaron en dos grupos o *clusters* (Fereidouni et al, 2009), y en siete diferentes sub-clados definidos (Nelson et al, 2009); sin embargo, la relevancia biológica de esta divergencia no es clara en términos de la adecuación evolutiva del virus (Fig. 4). Los virus de ambos *clusters* circularon a nivel global, aunque de manera gradual los virus del *Cluster 2* fueron reemplazando a los del *Cluster 1* (Nelson et al, 2009). La mayoría de los virus de México, California y Texas colectados durante la primera etapa de la pandemia (entre Abril y Mayo del 2009) agrupaban dentro del *Cluster 1* (o Texas), mientras que los virus colectados en Nueva York en meses posteriores (entre Junio del 2009 y Febrero del 2010) agrupaban dentro del *Cluster 2* (o Nueva York) (Fereidouni et al, 2009). Además, Nelson et al, encontraron que la mayoría de los virus colectados a principios de la pandemia en México pertenecían al *Cluster 1*, sub-clado 2, mientras que los de California pertenecían al sub-clado 1 (Anexo; Fig. 1; Nelson et al, 2009).

Sin embargo, dada la deficiencia en el muestreo y la vigilancia epidemiológica de los virus en las poblaciones porcinas en México y en todo el mundo, existen pocos datos que puedan contribuir en la resolución de cuándo, cómo y dónde emergieron los virus A/H1N1/2009 pandémicos. Por lo tanto, es de particular interés encontrar virus que estén estrechamente relacionados con el linaje pandémico o cepas de transición que puedan ayudar a resolver esta incógnita.



**Figura 4. Evolución y Predominancia de los Virus H1N1 a Nivel Global**

El siguiente esquema representa un línea del tiempo con la circulación de los principales linajes de virus tipo H1N1 humanos, desde 1918 hasta el 2010. Los virus tipo H1N1 se detectaron en humanos por primera vez en 1918, y fueron divergiendo en diferentes clados a lo largo del tiempo hasta el presente. El linaje humano antiguo o clásico corresponde a virus tipo 1918, que circularon hasta 1956, donde dejaron de circular. En 1977, los virus H1N1 resurgieron dando origen al linaje humano estacional, que circuló hasta el 2009. En este año, se detectaron por primera vez los virus pandémicos A/H1N1/2009 que reemplazaron a los virus estacionales, y que se diferenciaron a su vez en diferentes Clusters y sub-clados.

#### IV.2 LOS VIRUS A/H3N2 ESTACIONALES

En humanos, los virus H3N2 suelen ser más virulentos que los virus H1N1. Por ejemplo, las variantes antigénicas de los virus H3N2 han generado un mayor número de epidemias; desde 1990, nueve de once brotes estacionales se han asociado a virus tipo H3N2, mientras que sólo dos se asocian a virus tipo H1N1 (Subbarao et al, 2006). También, suelen ser más frecuentes y severos los casos por infecciones con virus tipo H3N2 que por H1N1. Por otro lado, la deriva antigénica parece ser más frecuente en los virus tipo H3N2 que en los virus H1N1, por lo que la tasa de evolución de los virus H3N2 es más elevada (Palese & Shaw 2006, Creanza et al, 2010). El análisis de la diversidad genética de los virus tipo H3N2 en una misma estación, revela que las epidemias son producto de la expansión clonal de una cepa que funciona como semilla al inicio de la misma (Creanza et al, 2010). Asimismo, se sabe que la diversidad génica entre las cepas en una misma epidemia es mayor al inicio de la estación, y que durante el pico estacional y hacia el final de la misma tiende a disminuir, lo que sugiere que hay virus que se seleccionan positivamente al inicio de la estación que predominan a lo largo de la misma.

Los virus H3N2 que infectan a humanos circulan a nivel global. En climas templados, las epidemias se asocian al invierno, mientras que en climas tropicales, ocurren principalmente en la época de lluvias (Cox y Subbarao, 2002). Sin embargo, la pregunta de cómo se propagan los virus de influenza A por todo el mundo para generar epidemias cada estación sigue sin respuesta. Una cuestión fundamental detrás de esta pregunta es resolver si las epidemias son consecuencia de la persistencia local o global de virus en niveles bajos, o si los linajes epidémicos se originan a partir de cepas semilla que salen de regiones específicas cada año.

Russell et al. han propuesto que los virus H3N2 se distribuyen desde Asia hacia América del Norte y Europa, como consecuencia de las migraciones humanas, y ocasionan epidemias locales en cada temporada. Por lo tanto, postula que Asia funciona como un foco de generación de nuevas cepas semilla que dan lugar a epidemias anuales, y que estos virus van cambiando por deriva antigénica conforme se diseminan en distintas poblaciones (Russell et al, 2008). Russell y colaboradores también propone que la persistencia global de los virus no contribuye a la evolución a largo plazo de los virus H3N2 humanos, pues al final de cada estación la cepa que da lugar a la epidemia deja de circular.

Sin embargo, se ha demostrado en trabajos previos que, si bien la deriva antigénica es el mecanismo dominante que ejerce mayor presión selectiva sobre los virus de influenza tipo H3N2, el surgimiento de nuevos linajes antigénicamente distintos también se da de manera frecuente por eventos de re-arreglos génicos entre linajes virales, y que la persistencia local y global es importante para la evolución a largo plazo de estos virus (Bedford et al, 2010; Nelson et al, 2010; Holmes et al, 2005).

Aunque se han caracterizado antigénicamente algunos de los virus H3N2 que circularon en México durante las estaciones de 2002 hasta 2009, el origen evolutivo y las características moleculares particulares de estas cepas se desconoce. A partir del análisis evolutivo molecular de los virus H3N2 colectados en México, podemos confirmar la identidad antigénica, detectar re-arreglos y mutaciones puntuales asociadas a resistencia a fármacos en estos virus.

La temporada de influenza 2003-2004 comenzó antes que la mayoría de las epidemias anuales reportadas anteriormente, y fue moderadamente grave (CDC, 2004). Mas del 99% de los virus que circularon correspondían a aislados tipo H3N2, del cual sólo el 11,2% fueron antigénicamente similares a la cepa de vacunación A/Panamá/2007/99 (que circuló la temporada anterior 2002-2003), mientras que el resto fueron similares a la variante divergente, A/Fujian/411/2002 (Fig. 5). Esto ocasionó una disminución en la

eficacia de la vacuna por un desajuste entre la cepa de vacunación y la predominante en circulación durante esa temporada. Para la estación del 2004-2005, los virus tipo A/Fujian/411/2002 fueron desplazados por los virus divergentes a los que se les denominó tipo A/California/07/2004 (Fig. 5), que circularon de manera predominante. Interesantemente, una pequeña parte de los virus que circularon en la estación 2004-2005 correspondían a virus rearreglantes tipo A/Wisconsin/67/2005, que para la estación 2005-2006, fueron los que predominaron en circulación (Fig. 5; CDC, 2005).

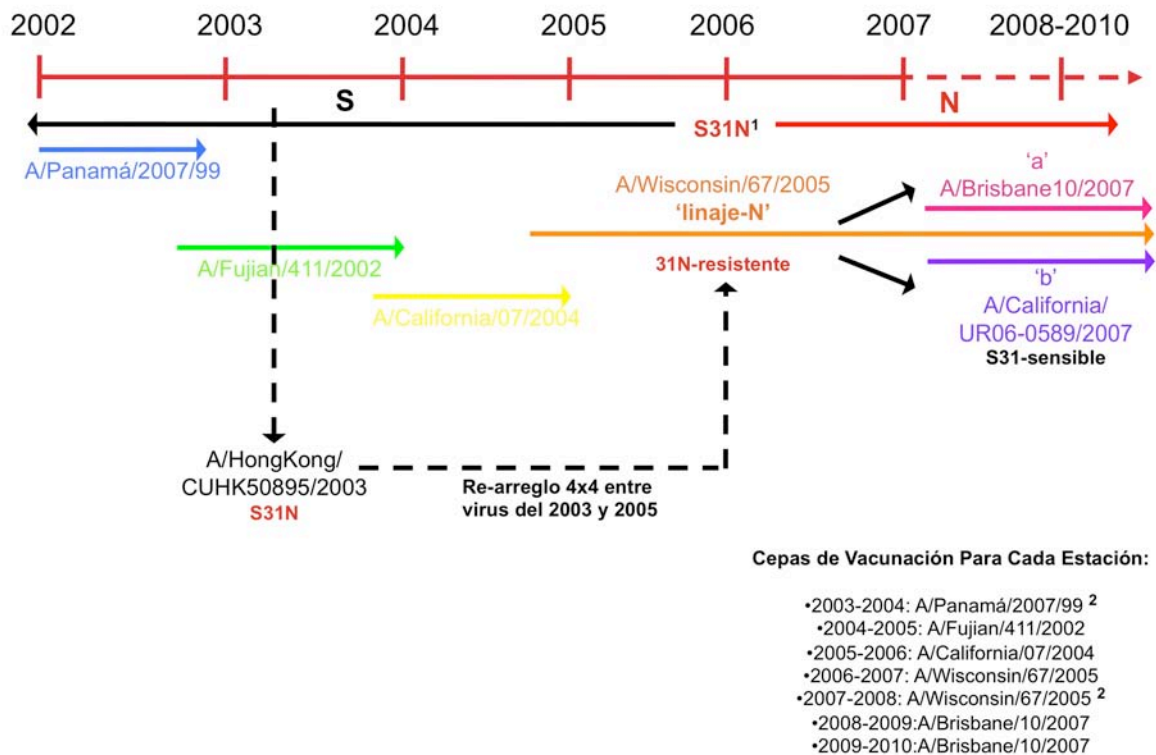
Los virus tipo A/Wisconsin/67/2005 se denominan linaje-N (N-lineage). Estos virus se generaron por un evento de rearreglo génico entre virus resistentes a amantadanos que circularon durante la temporada del 2003 y los virus sensibles a amantadanos que circularon en 2004-2005 (Nelson et al, 2008, Simonsen et al, 2007). Se ha postulado que el linaje-N posee los genes PB1, PA, NP y M de virus tipo A/HongKong/CUHK50895/2003, que circularon ampliamente en Asia en 2003 (Nelson et al, 2009), mientras que los genes HA, NA, PB2 y NS provienen de virus A/Canterbury/127/2005 que circularon durante 2004-2005 (Simonsen et al, 2007). El linaje-N se caracteriza por tener la mutación S31N en la proteína de canal iónico M2, que confiere resistencia a los amantadanos (Nelson et al, 2009, Nelson et al, 2008). Además de esta mutación, el linaje-N posee 17 cambios no silenciosos distribuidos a lo largo de siete proteínas virales (PB1, PA, NP, NA, M1 y M2). La introducción del linaje-N a la población representó un aumento importante en la resistencia al tratamiento con amantadanos (Fig. 5).

En las temporadas de influenza desde 1995 hasta 2002, se detectó que menos del 2% de los virus A/H3N2 a nivel mundial eran resistentes a amantadanos. Sin embargo, el incremento drástico en la resistencia a los amantadanos se detectó por primera vez en Asia, particularmente en China, donde la prevalencia en la resistencia aumentó de un 8,2% en 2002 a un 73,8% en 2004-2005. Recientemente, la mutación S31N se detectó en el 100% de los virus, en un muestreo que se realizó en múltiples países de Asia (Nelson et al, 2009, Deyde et al, 2007). Este incremento en la resistencia a los amantadanos se atribuye principalmente al surgimiento del linaje-N y a la dominancia de estos virus sobre los tipo A/California/07/2004, que habían circulado de manera previa en 2002-2003 (Fig. 5). Asimismo, no se considera que la mutación S31N se haya fijado en las diferentes poblaciones virales como consecuencia de la presión de selección que ejerce el uso de fármacos antivirales sobre la proteína M2, si no a partir de un proceso de *hitch-hiking* evolutivo, en el que una mutación particular es acarreada a partir de la selección de otras

mutaciones ventajosas en diferentes segmentos génicos, tal como HA (Simonsen et al, 2007).

Durante la estación de influenza de 2007-2008, nuevamente la mayoría de los virus que co-circularon en Estados Unidos correspondían a virus tipo H3N2 (CDC,2007). A partir de la tipificación genética y antigénica de los virus que circulaban, se escogió la cepa estacional H3N2 A/Wisconsin/67/2005 para preparar la vacuna para el 2007-2008. Sin embargo, hacia la mitad de la estación se detectó en circulación una cepa antigénicamente divergente, la cual se denominó como A/Brisbane/10/2007 (Fig. 5). Hacia el final de la estación, más del 87% de los virus circulantes eran tipo A/Brisbane/10/2007, por lo que no fue sorprendente que la vacuna sólo tuviera un 44% de efectividad (calculado a partir de los datos serológicos al final de la estación; CDC, 2008).

Hoy en día, además del linaje-N, se han detectado en circulación dos clados principales. El primero corresponde al clado mayor compuesto por virus tipo A/Brisbane/10/2007, denominado 'a'. En éste, agrupan virus resistentes a los amantadanos. El segundo linaje corresponde a un clado menor llamado 'b', donde agrupan virus tipo A/California/UR06-0589/2007 (Nelson et al, 2008). Aunque los clados 'a' y 'b' descienden directamente del linaje-N, el clado 'b' se caracteriza por ser sensible a los amantadanos, al poseer una serina (S) en la posición 31 de la proteína M2, lo que indica que ha ocurrido una reversión adquirida por medio de un rearreglo génico. Por otro lado, los clados 'a' y 'b' pueden variar antigénicamente, ya que difieren en varios aminoácidos en la proteína HA, cinco de los cuales están localizados en los sitios antigénicos A, B y C (Nelson et al, 2008). En total, se han reportado diez marcadores en HA que distinguen al linaje-N del clado 'a' y 'b', que se distribuyen principalmente en los sitios antigénicos A, B y C de HA (Nelson et al, 2008).



**Figura 5. Evolución y Predominancia de los Virus H3N2 a Nivel Global**

El siguiente esquema representa un línea del tiempo con la circulación de los principales linajes virales H3N2 por cada estación desde el 2002 hasta el 2010. Cada linaje está representado con su cepa de referencia en distintos colores; A/Panamá2007/199 en azul, A/Fujian411/2002 en verde, A/California/07/2004 en amarillo, el linaje-N, o virus tipo A/Wisconsin/67/2005 en naranja y los linajes 'a' tipo A/Brisbane/10/2007 en rosa y 'b' tipo A/California/UR06-0589/2007 en morado. La flecha punteada en negro, indica el evento de re-arreglo génico entre virus del 2003 y del 2005 que dio lugar al linaje-N. Las flechas que salen del linaje-N (en negro), muestran que los linajes 'a' y 'b' divergieron del mismo. Por debajo de la línea temporal está representado el aumento en la prevalencia de virus resistentes a amantadanos, que se asocia a la mutación S31N en la proteína M2. La línea negra representa virus sensibles (S31 en M2), mientras que la línea roja representa virus resistentes (con la mutación 31N en M2). <sup>1</sup> La prevalencia en la resistencia a amantadanos aumentó de un 8,2% en 2002 a un 73,8% en 2004-2005 (CDC). <sup>2</sup> Indica que para esa estación, la cepa escogida para la vacuna no coincidió con la cepa en circulación.

## HIPOTESIS

- Existen cepas que son intermediarios entre el último ancestro común de los virus pandémicos y estos, y podrían pertenecer a linajes hermanos del clado de virus A/H1N1 pandémicos.
- La persistencia local de cepas y los eventos de rearrreglos génicos contribuyen a la evolución a largo plazo de los virus tipo H3N2

## OBJETIVOS

- Optimizar diferentes métodos de preparación de bibliotecas de muestras de RNA viral, para la secuenciación masiva en la plataforma de Illumina.
- Estudiar el patrón evolutivo de los diferentes genes virales de las muestras secuenciadas en relación a otras secuencias en la base de datos.

## OBJETIVOS PARTICULARES

- 1) Construir las bibliotecas para secuenciación por diferentes métodos, para optimizar la cobertura y la resolución de las secuencias obtenidas.
- 2) Caracterizar los genomas de aislados clínicos de virus de la influenza A/H1N1/2009 y A/H3N2 que han circulado en México y analizar la presencia de mutaciones y/o polimorfismos en las secuencias.
- 3) Mediante el análisis filogenético, detectar patrones de evolución y virus rearrreglantes entre las muestras secuenciadas.



## MÉTODOS

### I. OBTENCIÓN DE MUESTRAS, PROPAGACIÓN Y PURIFICACIÓN DE VIRUS.

Se obtuvieron dieciséis muestras de virus A/H1N1/2009 pandémicos humanos que se colectaron en la ciudad de México en el Hospital General Benito Juárez y el Instituto Nacional de Enfermedades Respiratorias (INER), tres aislados de La Gloria, Veracruz que se colectaron en el laboratorio BSL-3 del Ministerio de Salud del Estado de Veracruz, y uno en el Instituto de Biotecnología, UNAM en Cuernavaca, Morelos. También se obtuvo la muestra A/swine/4/Mexico/09, que corresponde a un aislado porcino que se colectó en la Unidad de Aislamiento de Virus de CPA en Querétaro, Querétaro (Tabla 1). Para los virus H3N2, el Instituto de Diagnóstico y Referencia Epidemiológicos (InDRE) donó veintiún aislados clínicos de virus tipo H3N2 adaptados a cultivo. Las muestras se colectaron en un periodo de ocho años, en distintos lugares de la República Mexicana (Tabla 2). Algunos de los virus tipo H3N2 se caracterizaron antigénicamente de manera previa en el InDRE/CDC. Todos los virus se colectaron por el personal entrenado de cada institución en condiciones de anonimato.

Las muestras cultivables se propagaron en cultivos de células MDCK (American Type Culture Collection CCL34) en medio MEM a 37° C con tripsina TPCK por aproximadamente 48 horas y la presencia del virus se determinó por la presencia de efecto citopático (ECP). Una vez observado el ECP, los sobrenadantes del cultivo enriquecidos con partículas virales se colectaron y almacenaron a -70° C para su uso posterior. Algunas muestras no cultivables se propagaron en huevo y se extrajo el RNA para secuenciación directa. Para otras, únicamente se extrajo el RNA para secuenciación directa.

### II. EXTRACCIÓN DE RNA GENÓMICO VIRAL Y AMPLIFICACIÓN POR RT-PCR MULTISEGMENTO (M-RTPCR).

A partir de los sobrenadantes de cultivo o de las muestras clínicas, se extrajo el RNA total de cada muestra, usando un kit comercial (PureLink RNA/ DNA mini Kit, Invitrogen), tratando previamente con Turbo DNAsa (Ambion) por 1 hora a 37°C. La presencia de RNA correspondiente al virus de influenza se verificó mediante un RT-PCR (OneStep RT-PCR Kit, Qiagen) con oligonucleótidos específicos para amplificar un fragmento del gen M. Algunas de las muestras de RNA a partir de las cuales se obtuvieron amplicones para

M, se seleccionados para hacer un RT-PCR multi-segmento (M-RTPCR), en el que se amplificó el genoma completo usando oligonucleótidos universales específicos para influenza, que son complementarios a los extremos 3' y 5' conservados de cada uno de los segmentos génicos (MBT<sub>uni-12</sub> 5'ACGCGTGATCAGCAAAGCAGG3', MBT<sub>uni-13</sub> 5'ACGCGTGATCAGTAGAAACAAGG 3') (Zhou et al, 2009). Se utilizó un kit comercial (Superscript III High-Fidelity One-step RT-PCR, Invitrogen) que usa la enzima de alta fidelidad DNA polimerasa *Taq Platinum* bajo las siguientes condiciones de amplificación: 42°C 60 min, 94°C 2 min, (94°C 30 seg, 45°C 30 seg, 68°C 3 min) X 5 ciclos, (94°C 30 seg, 57°C 30 seg, 68°C 3 min) por 31 ciclos (Zhou et al, 2009).

### III. CONSTRUCCIÓN DE BIBLIOTECAS Y SECUENCIACIÓN

Para determinar cuál era el mejor sistema para generar bibliotecas de óptima calidad para la secuenciación masiva, se probaron cuatro métodos distintos: 1) PCR, amplificación del genoma completo al azar con oligos degenerados y amplificación por PCR anidado (Wang et al. 2002); 2) FR, fragmentación del genoma completo amplificado previamente, mediante la enzima fragmentasa; 3) HA, síntesis de cadena complementaria de cDNA de los genomas completos mediante el uso de hexámeros al azar; y 4) mRNA: procesamiento directo de RNA total extraído, sin enriquecimiento previo del material genético (Diagrama 1). A continuación, se describe cada método de manera detallada:

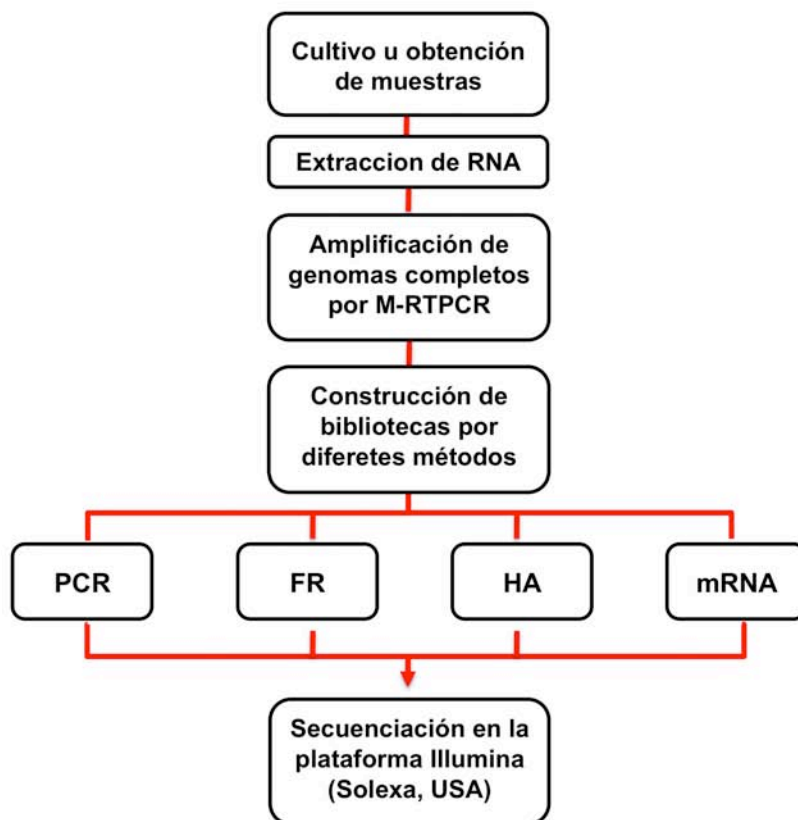
**PCR:** En la primera parte de este método, se utilizó un oligonucleótido degenerado compuesto por 12 nucleótidos de secuencia conocida hacia el extremo 5' y con una secuencia de 9 nucleótidos al azar hacia el extremo 3' (5'CACTGGAGGATA-NNNNNNNN3'). Este oligonucleótido se pega en cualquier secuencia de DNA y se utiliza para sintetizar la cadena complementaria de DNA de todos los segmentos del genoma viral con la enzima secuenasa 2.0 (USB, USA), en dos rondas seguidas. Las condiciones de síntesis fueron las siguientes: 94°C 2 min, 10°C 10 min, 10-37°C gradiente 8 min, 37°C 8 min dos ciclos. En la segunda parte de este método, los productos del primer paso se purificaron por columnas de afinidad (ZYMO, EUA), y se amplificaron de manera específica mediante un PCR anidado. Para esto, se uso un oligonucleótido que reconoce la cola de 12 nucleótidos conservada en el oligonucleótido utilizado en el paso previo (5'GTTTCC-CACTGGAGGATA3'). Las condiciones de amplificación fueron las siguientes: 94°C 2 min, (94°C 30 seg, 50°C 1 min, 72°C 1 min) por 10-15 ciclos, 72°C 8 min. Una vez obtenido el producto, nuevamente se purificó como

se describió anteriormente. El oligonucleótido que se usó en la segunda parte del método tiene un sitio de reconocimiento para la enzima de restricción *GsuI* (marcado en rojo), la cual corta 16 y 14 nucleótidos después del sitio de reconocimiento. El DNA se digirió con la enzima *GsuI* (Fermentas, EUA) a 30°C por 2 horas e incubando de manera posterior a 65°C por 20 min para inactivar la enzima. La digestión del producto tiene como propósito eliminar las secuencias correspondientes a los oligonucleótidos, dejando únicamente secuencias virales. Una vez llevado a cabo los pasos anteriores, las bibliotecas se purificaron como se describió anteriormente.

**FR:** En este proceso, cada genoma completo se trató con la enzima fragmentasa (NEBnext dsDNA Fragmentase, NEB, EUA), la cual rompe al azar el DNA de doble cadena para generar fragmentos de 100-800 pb. Las muestras se incubaron con la enzima durante 12 min. a 37°C para obtener fragmentos de 100-800 pb. Posterior a la fragmentación, cada producto se limpió por columnas de concentración y de limpieza de DNA (ZYMO research, EUA).

**HA:** En este tercer método de preparación de bibliotecas, se utilizaron hexámeros al azar (Fermentas, EUA) para iniciar la síntesis de cDNA de cada genoma completo obtenido previamente. Las condiciones de la síntesis fueron las siguientes: 94°C 2 min., 10°C 10 min., 10-37°C gradiente 8 min., 37°C 8 min., por dos rondas. Cada producto se purificó por columnas de concentración y de limpieza de DNA (ZYMO, EUA). En este método no hay una amplificación del material genético por PCR.

**mRNA:** De manera alterna, para los virus propagados en huevo, el RNA total se obtuvo como se describió previamente y se procesó con el kit de preparación de Illumina mRNA-Seq 8, sin ningún paso de amplificación o enriquecimiento del material genético previo.



**Diagrama 1.**

Diagrama de flujo con los diferentes pasos y métodos usados para la obtención de las bibliotecas para hacer secuenciación masiva.

Finalmente, el DNA obtenido por los cuatro diferentes métodos se procesó de manera individual, siguiendo los protocolos de Solexa (Illumina, EUA). Las bibliotecas obtenidas de aproximadamente 350 pares de bases se sometieron a secuenciación masiva por la técnica de dNTP's terminadores reversibles marcados fluorescentemente (Solexa, Illumina, EUA). La secuenciación se realizó en las instalaciones de la Unidad de Secuenciación Masiva de la Universidad Nacional Autónoma de México (IBT-UNAM), y la captura de imágenes se hizo mediante el secuenciador Illumina Genome Analyzer II, (Illumina, EUA). Todas las muestras se secuenciaron por *Single-Read*, con una sola muestra por carril o haciendo *multiplexing* con dos a cinco muestras por carril, en cuatro corridas distintas. En las primeras dos corridas se hizo secuenciación con una muestra por carril, secuenciado 36nt por *read* o lectura. En la segunda y en la tercera corrida se usó el método *multiplexing*, con dos a cinco muestras por carril, secuenciando 46nt y 60nt por *lectura*, respectivamente. En las muestras en las que se usó el método de

*multiplexing*, las lecturas idénticas de cada muestra se colapsaron para prevenir el posible sesgo introducido por la amplificación de PCR. El análisis computacional se hizo usando el cluster del Instituto de Biotecnología, UNAM.

Los genomas completos de cada muestra se ensamblaron y analizaron mediante los programas MAQ versión 0.7.1 y R/Bioconductor instalados en el cluster de computo (IBT-UNAM), utilizando el genoma de referencia de la cepa A/H1N1/Netherlands/602/2009 para los virus pandémicos A/H1N1/2009, y el virus A/New York/392/2004 para los virus estacionales H3N2. Los SNPs en las secuencias se identificaron usando MAQ versión 0.7.1 y Mapview (Bioinformatics, 2009), en las posiciones con al menos 10X de cobertura.

MUESTRA	TIPO VIRAL	FECHA COLECTA	LUGAR DE COLECTA
A/swine/4/Mexico/09	A/H1N1/2009 pandémico	22/04/09	Querétaro, Querétaro
A/Mexico/La Gloria-3	A/H1N1/2009 pandémico	30/03/09	La Gloria, Veracruz
A/Mexico/La Gloria-4	A/H1N1/2009 pandémico	30/03/09	La Gloria, Veracruz
A/Mexico/La Gloria-8	A/H1N1/2009 pandémico	30/03/09	La Gloria, Veracruz
4	A/H1N1/2009 pandémico	13/10/09	Cuernavaca, Morelos
22	A/H1N1/2009 pandémico	07/05/09	Hosp.Gen., D.F.
23	A/H1N1/2009 pandémico	07/05/09	Hosp.Gen., D.F.
47	A/H1N1/2009 pandémico	12/10/09	INER, D.F.
96	A/H1N1/2009 pandémico	03/11/09	INER, D.F.
137	A/H1N1/2009 pandémico	19/11/09	INER, D.F.
250	A/H1N1/2009 pandémico	11/01/10	INER, D.F.
306	A/H1N1/2009 pandémico	09/02/10	INER, D.F.
307	A/H1N1/2009 pandémico	09/02/10	INER, D.F.
96	A/H1N1/2009 pandémico	03/11/09	INER, D.F.
137	A/H1N1/2009 pandémico	19/11/09	INER, D.F.
250	A/H1N1/2009 pandémico	11/01/10	INER, D.F.

**Tabla 1. Muestras Secuenciadas de Virus Tipo A/H1N1/2009**

Virus A/H1N1/2009 pandémicos secuenciados en este trabajo. En la tabla, se especifica el nombre de la muestra, el tipo viral, fecha y lugar de colecta. Todos los virus colectados son de origen humano, excepto la muestra A/swine/4/Mexico/09, que corresponde a un aislado porcino.

MUESTRA	TIPO VIRAL	AÑO DE COLECTA	LUGAR DE COLECTA
756	H3N2 A/Korea/770/2002	2003	D.F.
835	H3N2	2003	D.F.
940	H3N2	2003	D.F.
2662	H3N2 A/Korea/770/2002	2003	D.F.
2664	H3N2 A/Korea/770/2002	2003	D.F.
2112	H3N2 A/California/07/2004	2005	D.F.
2160	H3N2 A/California/07/2004	2005	D.F.
2227	H3N2 A/Wisconsin/67/2005	2005	Tlaxcala
2246	H3N2 A/Wisconsin/67/2005	2005	D.F.
2270	H3N2	2005	Querétaro
2408	H3N2	2005	San Luis Potosí
2601	H3N2 A/California/07/2004	2005	D.F.
29	H3N2 A/Wisconsin/67/2005	2006	D.F.
2250	H3N2	2006	D.F.
2640	H3N2	2006	Edomex
2090	H3N2	2007	Nayarit
2362	H3N2	2007	Colima
499	H3N2	2008	Baja California
25206	H3N2	2009	Jalisco
25216	H3N2	2009	Jalisco
25565	H3N2	2009	Jalisco

**Tabla 2. Muestras Secuenciadas de Virus Tipo A/H3N2**

Virus H3N2 colectados en la Republica mexicana a lo largo de ocho años secuenciados para este trabajo. En la tabla se especifica el nombre de la muestra, el tipo viral al que corresponde si la muestra fue caracterizada antigénicamente de manera previa año y lugar de colecta. Todos los virus colectados fueron de origen humano.

#### IV. ANÁLISIS FILOGENÉTICO Y DETECCIÓN DE MUTACIONES

Para caracterizar la historia filogenética de los virus H3N2 y H1N1 pandémicos, se obtuvieron las regiones que codifican para proteínas de los ocho genes virales. Además de las secuencias obtenidas en este trabajo, se seleccionó un sub conjunto representativo de secuencias de cada tipo viral (H1N1 y H3N2) disponible en la base de datos NCBI-IVR. Las secuencias se alinearon mediante el algoritmo MAFFT (Kato y Toh, 2009) implementado en SeaView (Gouy et al, 2010). Posteriormente, se aplicó ModelTest (Posada et al, 1998) a cada alineamiento para determinar el modelo de sustitución con mejor ajuste para el análisis filogenético de cada sub conjunto de secuencias.

#### IV. 1 VIRUS A/H3N2 ESTACIONALES

Para los virus estacionales H3N2, se extrajeron las regiones que codifican para cada una de las proteínas virales de las secuencias obtenidas. Cada grupo de secuencias se alineó de manera individual junto con otras secuencias de la base de datos, como se describió anteriormente. Para esto, se seleccionaron 100-125 secuencias en la base de datos NCBI-IVR, que corresponden a virus de diferente origen animal colectados en diferentes países del mundo a lo largo de varios años (1968-2010); se incluyeron las siguientes cepas de referencia: A/Panamá/2007/99, A/Korea/770/2002, /Fujian/411/2002, A/HongKong/CUHK50895/2003, A/California/07/2004, A/Wisconsin/67/2005, A/Brisbane/10/2007 y A/California/UR06-0589/2007, así como otras cepas mexicanas secuenciadas previamente (A/Yucatán/ME6057/2003, A/Yucatán/844/2003). Una vez obtenidos los alineamientos, se hizo análisis evolutivo de cada gen bajo el criterio de máxima verosimilitud (ML) en programa PhyML (Guindon & Gascuel, 2003). Los árboles de ML se infirieron bajo el modelo GTR+I+G, con aLRT (tipo-SH) para el soporte de las ramas. Todos los árboles se enraizaron por punto medio para mayor claridad.

#### IV. 2 VIRUS A/H1N1/2009 PANDÉMICOS

Para el estudio filogenético de los virus A/H1N1/2009 pandémicos, se extrajeron las regiones que codifican para cada una de las proteínas virales de las secuencias obtenidas. De manera paralela, se seleccionó un subconjunto de 55 secuencias de virus pandémicos disponibles en la base de datos NCBI-IVR hasta septiembre de 2010. Las secuencias se eligieron para representar homogéneamente aislados de todo el mundo, con fechas de colección desde abril 2009 hasta febrero 2010. El análisis previo de las secuencias con algoritmos de alineamiento local y global (BLAST, MAFFT) mostró que el porcentaje de identidad es muy alto en todos los genes (%ID: 99,8%) entre los virus A/H1N1/2009 pandémicos, incluyendo las cepas de este estudio. Por lo tanto, se usó una muestra pequeña pero representativa de 55 de las secuencias virales, ya que permite una mejor resolución del árbol. Las secuencias se alinearon como se describió anteriormente.

El análisis filogenético se realizó a partir de las regiones que codifican para las proteínas virales, individuales o concatenadas. Se utilizó análisis Bayesiano bajo un modelo de reloj molecular relajado y de cadenas de Markov de Monte Carlo, implementado en BEAST v1.4.8 (Drummond y Rambaut, 2007). Los árboles se infirieron bajo el modelo SRD06, que parte los datos según la posición de las bases en cada codón.

La primera y segunda posición se toman como una sola partición, mientras que la segunda partición corresponde a la tercera posición en el codón, y se estiman parámetros individuales para ésta. Se utilizó el modelo de reloj molecular relajado UCLD, que permite la variación de la tasa de evolución a lo largo de las ramas, dentro de una distribución logarítmica normal. Se llevaron a cabo al menos dos análisis independientes de 30 millones de generaciones con un muestreo adecuado para producir 10 mil árboles, garantizando un tamaño efectivo de muestra  $> 200$ ; finalmente se eliminó el correspondiente al 10% del muestro inicial de cada corrida. Cada árbol corresponde a un árbol de máxima credibilidad (MCC), seleccionado con un límite de valor de probabilidad posterior de 0.5.

De manera alterna, se realizó análisis evolutivo bajo el criterio de máxima verosimilitud (ML) con los programas PhyML (Guindon & Gascuel, 2003) o PHYLIP (Felsenstein, 1993). Los árboles de ML se infirieron bajo el modelo de sustitución de nucleótidos HKY85, con aLRT (tipo-SH) para el soporte de las ramas. Todos los árboles se enraizaron con secuencias de virus de origen porcino y/o aviar identificadas como las más cercanas en análisis evolutivos previos (Smith et al, 2009). Para reducir la probabilidad de artefactos en el análisis filogenético debido a la atracción de ramas largas (LBA), se aseguró que las secuencias usadas para el enraizamiento tuvieran un porcentaje de identidad de entre el 85 y el 98% a nivel de nucleótidos.

Para la detección de mutaciones y de polimorfismos en los diferentes genes virales de los virus H1N1, se usaron dos diferentes conjuntos de datos; el primer conjunto (set uno) se utilizó para identificar polimorfismos entre nuestros aislados en relación a los demás virus pandémicos. Este set se compuso de secuencias individuales correspondientes a 1958 genomas completos de virus A/H1N1/2009 pandémicos disponibles en la base de datos NCBI-IVR, hasta septiembre de 2010. El segundo conjunto (set dos) se usó para la detección de polimorfismos en relación a los virus H1N1 no-pandémicos, por lo que contenía secuencias individuales para todos los genes virales de virus tipo H1N1 estacional, sin importar el origen biológico, lugar o fecha de colección, disponibles en la base de datos NCBI IVR-hasta septiembre de 2010. Las regiones que codifican para proteínas se extrajeron y se alinearon para cada set como se describió anteriormente. Los sitios variables se detectaron de forma manual en los virus secuenciados en este trabajo, con el uso de los editores de alineamientos SeaView (Gouy et al, 2010) y Jalview (Clamp et al, 2004).

En este trabajo, se utilizaron las siguientes definiciones de polimorfismo y mutación para poder distinguir entre ambos. Un polimorfismo se define como la existencia de dos o



mas genotipos en una población, en donde la frecuencia del genotipo más raro en la población es de más del 1% (Futuyma, 1998). Una mutación, por otro lado, se define como un error en la replicación de una secuencia de nucleótidos, que resulta en la sustitución de una base, y que ocurre en una frecuencia del 1% o menos en la población (Futuyma, 1998). Por lo tanto, un polimorfismo difiere de una mutación puntual en la frecuencia. Por lo tanto, un polimorfismo difiere de una mutación puntual única en su frecuencia. Las frecuencias de los polimorfismos se calcularon con SequenceLogo (Crooks et al, 2004), implementado en el editor JalView.

## RESULTADOS Y DISCUSIÓN

### I. ANÁLISIS DE LAS SECUENCIAS

Los métodos de secuenciación masiva han sido utilizados para caracterizar los genomas de poblaciones complejas y para detectar mutaciones dentro de éstas, convirtiéndose en una herramienta útil para analizar la diversidad genética (Yongfeng et al, 2010; Nakamura et al, 2009). El uso de este método se ha limitado a muestras de tipo viral y en particular para virus de RNA, puesto que el material genético suele ser menos abundante y estable (Kuroda et al, 2010; Ramakrishnan et al, 2009; Greniger et al, 2009). En varios trabajos se ha utilizado la secuenciación masiva para caracterizar el genoma de los virus de influenza, obteniendo resultados variables (Kuroda et al, 2010; Ramakrishnan et al, 2009; Greniger et al, 2009). En muchos de estos trabajos, no se ha logrado secuenciar el genoma completo del virus de influenza o la cobertura de la secuenciación no fue suficiente. Para este trabajo, se usaron cuatro diferentes métodos para la preparación de bibliotecas para secuenciación masiva (Solexa, Illumina), con el fin de tratar de enriquecer el material inicial, y así poder maximizar la cobertura y la resolución en el proceso de secuenciación.

Se comprobó que los cuatro métodos probados son efectivos para obtener la secuencia genómica completa del virus de influenza (Tabla 3). El método a utilizar debe seleccionarse en relación a la cantidad de RNA inicial en la muestra, y sobre todo en la calidad del producto obtenido de la amplificación del genoma por M-RTPCR. Importantemente, en todas las secuencias obtenidas se observaron SNPs; la mayoría de los virus analizados presentaban menos de 20 SNPs en todo el genoma viral, sin importar el método por el cual se obtuvieron, con algunas excepciones en donde se detectaron más de 50 (Tabla 3). Para esta parte del trabajo se tomo la definición de SNP como una variación en una secuencia de bases que ocurre en cualquier posición del genoma (Futuyma, 1998).

Muestra	Método	Reads Totales	Reads Únicos (%)	Reads Infl. (%)	Cob. Gen. (%)	SNPs
4	PCR	11, 252, 936	939, 183	369, 462	13, 632 (100)	20
22	FR	7, 196, 055	725, 937 (10.09)	104, 475 (14.4)	13, 848 (99.9)	6
23	FR	7, 606, 144	861, 913 (11.33)	32, 142 (3.7)	12, 835 (94.2)	25
29	mRNA	1, 804, 703	210, 069 (11.64)	35, 645 (16.9)	13, 550 (99.4)	10
47	FR	3, 769, 418	649, 594 (17.23)	16, 876 (2.6)	12, 971 (95.2)	12
96	PCR	1, 145, 281	167, 785 (14.65)	33, 362 (19.9)	12, 218 (89.6)	6
120	PCR	5, 513, 830	670, 629 (12.16)	4, 252 (0.6)	11, 100 (81.4)	16
137	PCR	4, 865, 616	1, 001, 792 (20.59)	224, 268 (22.4)	13, 631 (99.9)	3
250	PCR	5, 444, 679	1, 021, 111 (18.75)	43, 141 (4.2)	10, 245 (75.2)	12
306	PCR	6, 699, 437	897, 343 (13.39)	10, 488 (1.2)	12, 290 (90.2)	6
307	FR	6, 681, 663	380, 564 (5.70)	190, 024 (49.9)	13, 632 (100)	6
307	PCR	23, 917, 782	4, 164, 027 (17.41)	798, 468 (19.2)	13, 632 (100)	32
307	RH	4, 662, 243	1, 003, 268 (21.52)	230, 198 (22.9)	13, 632 (100)	4
499	FR	4, 809, 690	793,739 (16.5)	218,716 (27.6)	13, 512 (99.2)	17
756	PCR	106, 091	29, 563 (27.87)	14, 381 (48.6)	13, 581 (99.7)	3
835	FR	7, 428, 939	694, 502 (9.35)	405, 459 (58.4)	13, 627 (100)	8
940	mRNA	1, 170, 726	145, 244 (12.41)	22, 312 (15.4)	13, 505 (99.1)	6
2090	FR	8,429,818	622,062 (7.4)	288,345 (46.4)	13, 604 (99.8)	29
2112	FR	9, 564, 287	448, 404 (4.69)	204, 777 (45.7)	13, 609 (99.9)	10
2160	FR	8, 388, 114	841, 094 (10.03)	411, 575 (48.9)	13, 618 (99.9)	8
2227	PCR	6, 035, 458	787, 564 (13.05)	235, 162 (29.9)	13, 624 (99.9)	9
2246	PCR	22, 552, 895	2, 113, 683 (9.37)	403, 047 (19.1)	13, 627 (100)	10
2250	FR	6,214,786	561,535 (9.0)	191,289 (34.1)	13, 572 (99.6)	31
2270	FR	6,374,530	784,974 (12.3)	447,902(57.1)	13, 599 (99.8)	12
2362	FR	8,895,019	601,477 (6.8)	254,305 (42.3)	13, 532 (99.3)	25
2408	FR	7,757,037	1,277,943 (16.5)	62,148 (4.9)	13, 580 (99.7)	88
2601	FR	4, 778, 676	299, 053 (6.26)	130, 651 (43.7)	13, 609 (99.9)	1
2640	FR	7,041,205	614,391 (8.7)	314,987 (51.3)	13, 583 (99.7)	18
2662	FR	11, 198, 199	617, 408 (5.51)	344, 132 (55.7)	13, 624 (99.9)	58
2662	RH	4, 030, 487	727, 468 (18.05)	143, 492 (19.7)	13, 595 (99.8)	53
2664	FR	6,433,919	935,048 (14.5)	440,459 (47.1)	13, 589 (99.7)	37
25206	FR	17,068,231	1,203,421 (6.9)	564,761 (46.9)	13, 627 (99.9)	100
25216	FR	8,647,987	635,392 (7.3)	335,005 (52.7)	13, 588 (99.7)	17
25565	FR	11,559,829	1,010,957 (8.7)	408,684 (40.4)	13, 577 (99.6)	26
A/swine/4	mRNA	10, 593, 972	ND <sup>1</sup>	101, 038 (1.0)	13, 564 (99.5)	6
La Gloria-3	mRNA	8, 576, 322	ND	380, 093 (4.4)	13, 592 (99.7)	4
La Gloria-4	mRNA	7, 609, 171	ND	242, 307 (3.2)	13, 556 (99.4)	5
La Gloria-8	mRNA	10, 421, 316	ND	44, 998 (0.4)	13, 543 (99.3)	5

**Tabla 3. Resultados de la Secuenciación**

En la tabla se indica la muestra, el método de preparación de biblioteca usado por muestra, el número total de lecturas (*reads*) obtenidos por cada muestra secuenciada, el número de *reads* únicos, el número de *reads* que corresponden al genoma de influenza (en porcentaje), así como la cobertura total del genoma de influenza (en porcentaje) y el número de SNPs en cada muestra. Las bibliotecas de las muestras 307 y 2662 fueron obtenidas por diferentes métodos. La cobertura obtenida para misma muestra preparada por distintos métodos fue la misma, y las secuencias sólo difieren en el número de SNPs encontrados. <sup>1</sup> No Determinado.

Para comparar la calidad de las bibliotecas obtenidas por los diferentes métodos, las bibliotecas de las muestras 307 y 2662 fueron procesadas por al menos dos métodos (307 por FR, PCR y RH; 2662 por FR y RH) (Tabla 3). A partir del análisis de la secuenciación, se determinó que la cobertura de las secuencias de una misma muestra procesada por diferentes métodos fue la misma. La resolución de la secuenciación también fue igual, pues se determinó que todos los nucleótidos resueltos coinciden, y las secuencias obtenidas sólo difieren en el número de SNPs encontrados (Tabla 3). Para la muestra 307 se observó que en los tres casos la cobertura del genoma fue del 100%. Asimismo, para la muestra 307, procesada por el método de FR se encontraron 6 SNPs; por el método de RH se observaron 4; mientras que por el método de PCR se hallaron 32. La cobertura del genoma de la muestra 2662 procesada por los métodos de FR y RH fue de 99.8 y 99.9 %, respectivamente. En la muestra 2662 procesada por el método de RH se observaron 53 SNPs, mientras que la procesada por el método de FR presentó 58.

Finalmente, la comparación de secuencias del genoma completo de los virus A/H1N1/2009 pandémicos mostró que dos de los tres aislados de La Gloria son idénticos entre sí a nivel de nucleótidos (La Gloria-3 y La Gloria-4). Sólo hay dos diferencias entre estos dos virus y el tercer aislado de La Gloria (La Gloria-8), a nivel de aminoácidos. Por lo tanto, sólo se utilizaron los virus de La Gloria-4 y La Gloria-8 para el análisis filogenético posterior.

## II. ANÁLISIS EVOLUTIVO DE LOS VIRUS A/H1N1 PANDÉMICOS

La caracterización genética y el análisis filogenético de los virus A/H1N1/2009 pandémicos demostró que son de origen porcino, y representan una nueva cepa que contiene los genes PB1, PB2, PA, HA, NP y NS derivados de virus porcinos triple rearreglantes de América del Norte, mientras que los genes de NA y M provienen del linaje porcino euroasiático (Garten et al, 2009). Se describió anteriormente que los virus A/H1N1/2009 pandémicos que circularon a nivel global se diferenciaron en dos grupos o clusters (Fereidouni et al, 2009) y en siete diferentes sub-clados definidos (Nelson et al, 2009).

De manera previa se determinó que los virus colectados en México a principios de la pandemia, pertenecen al *Cluster 1*/sub-clado 2 (Nelson et al, 2009). Se buscaron en las secuencias de las muestras los marcadores moleculares distintivos para cada *Cluster* y sub-clado descritos anteriormente. A partir de esto se determinó que los virus con fechas de colección temprana (virus 22 y los tres aislados de La Gloria, colectados entre Marzo y

Mayo del 2009, Ver Tabla 1) agrupan dentro del *Cluster 1*/ sub-clado 2. Por otro lado, se observó que la mayoría de los virus colectados en México con fecha de colección tardía (después de Junio del 2009), agrupan entre el sub-clados 7 del *Cluster 2* (Tablas 4, 5 y 6; Anexo; Figuras 1 y 2). Las muestra 307, 306, 250, 137, 96, 47, y 4 pertenecen al sub-clado 7, pues tienen la mutación I123V en la proteína NS1, 220T en HA y V06I en NA, que definen a dicho sub-clado (Tablas 4, 5 y 6). Por otro lado, la muestra 23 pertenece al sub-clado 5. La distribución de los marcadores moleculares de *Clusters* y sub-clados en las secuencias coincide con el análisis filogenético (Anexo; Figuras 1 y 2).

Virus	Gene (nt/aa)	PB2 (2163/721)	HA (658/220) (1408/470)		NP (298/100) (1143/381)		NA (742/248)	M (492/164) (600/200)		NS (367/123)
307		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	A/Asn	A/Gln	A/Ala	G/Val
306		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	A/Asn	A/Gln	A/Ala	G/Val
47		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	G/Asp	A/Gln	A/Ala	G/Val
22		G/Lys	T/Ser	C/Leu	G/Val	G/Ala	A/Asn	G/Gln	G/Ala	A/Ile
23		G/Lys	T/Ser	T/Leu	A/Ile	A/Ala	G/Asp	G/Gln	G/Ala	A/Ile
4		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	G/Asp	A/Gln	A/Ala	G/Val
96		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	G/Asp	A/Gln	A/Ala	G/Val
137		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	G/Asp	A/Gln	A/Ala	G/Val
250		A/Lys	A/Thr	T/Leu	A/Ile	A/Ala	G/Asp	A/Gln	A/Ala	G/Val
A/Mexico/La Gloria-4		G/Lys	T/Ser	C/Leu	G/Val	G/Ala	A/Asn	G/Gln	G/Ala	A/Ile
A/Mexico/La Gloria-8		G/Lys	T/Ser	C/Leu	G/Val	G/Ala	A/Asn	G/Gln	G/Ala	A/Ile
A/swine/4/Mexico/09		G/Lys	T/Ser	C/Leu	G/Val	G/Ala	A/Asn	G/Gln	G/Ala	A/Ile
<b>Cluster (1) Texas</b>		<b>G/Lys</b>	<b>T/Ser</b>	<b>C/Leu</b>	<b>G/Val</b>	<b>G/Ala</b>	<b>A/Asn</b>	<b>G/Gln</b>	<b>G/Ala</b>	<b>A/Ile</b>
<b>Cluster (2) New York</b>		<b>A/Lys</b>	<b>A/Thr</b>	<b>T/Leu</b>	<b>A/Ile</b>	<b>A/Ala</b>	<b>G/Asp</b>	<b>A/Gln</b>	<b>A/Ala</b>	<b>G/Val</b>

**Tabla 4. Marcadores Moleculares que definen a los Clusters 1 y 2 de los Virus A/H1N1 Pandémicos.**

Los Clusters 1 y 2 en los que divergen los virus pandémicos están definidos por nueve cambios en nucleótidos y aminoácidos distribuidos en seis genes. En rojo se indican los marcadores del Cluster 1 y en azul los del Cluster 2. La mayoría de los aislados pandémicos secuenciados pertenecen al Cluster 2, mientras que los virus 22, los aislados de La Gloria y A/swine/4/Mexico/09 poseen los marcadores del Cluster 1, que se ha caracterizado como el más temprano (Fereidouni et al, 2009).

SUB-CLADO	PB2	PB1	PA	HA	NP	NA	M1/2	NS1	NS2
1			S224P*	S100P* A214T* V338I*					
2			M581L*		T373I				
3									
4	V649I*	I667T			V100I	V106I			E63K
5					V100I*	V106I* N248D			
6				K2E Q310H*	V100I	V106I N248D			
7				S220T	V100I*	V106I N248D*		I123V	

**Tabla 5. Marcadores Moleculares que definen a los Sub-Clados de los Virus A/H1N1 Pandémicos.**

Se han descrito siete diferente sub-clados definidos por 24 cambios no sinónimos en los ocho genes virales. El asterisco indica que la presencia de esa mutación no es determinante para la agrupación dentro un sub-clado particular (Modificado de Nelson et al, 2009).

MUESTRA	PB2	PB1	PA	HA	NP	NA	M1/2	NS1	NS2	SUB-CLADO
A/swine/4/Mexico/09				214T*						-
A/Mexico/La Gloria-4			581L*		373I					2
A/Mexico/La Gloria-8			581L*		373I					2
307				220T	373I	V106I N248D*		I123V		7
306				220T	373I	V106I		I123V		7
47				220T	373I	V106I N248D*		I123V		7
22			581L*		373I					2
23					373I	V106I N248D*				5
4				220T	373I	V106I N248D*		I123V		7
96				220T	373I	V106I N248D*		I123V		7
137				220T	373I	V106I N248D*		I123V		7
250				220T	373I	V106I N248D*		I123V		7

**Tabla 6. Marcadores Moleculares de los Sub-Clados presentes en los Virus Secuenciados.**

Los virus analizados poseen los marcadores moleculares que definen los diferentes sub-clados descritos previamente. En la tabla se indica los marcadores presentes en los virus analizados y a que sub-clado corresponde cada uno.

Es interesante notar que aunque los virus 22 y 23 se colectaron en la misma fecha y en la misma localidad (07/05/09, Hospital General, D.F), el virus 22 pertenece al *Cluster 1*/sub-clado 2, mientras que el virus 23 al *Cluster 2*/sub-clado 5 (Anexo; Fig. 1 y 2). Estos hallazgos son congruentes con los trabajos previos en donde se ha determinado la co-circulación de ambos *Clusters* virales desde el inicio de la pandemia. Sin embargo, el virus 23 comparte algunos marcadores del *Cluster 1* (Tabla 4). En la posición 2163 del gen PB2, que corresponde al aminoácido 721, tiene una Gly/Lys, en la posición 658 del gen HA (aminoácido 220) tiene una Thr/Ser, y en la posición 367 del gen NS (aminoácido 123) tiene una Ala/Ile. Esto sugiere que el virus 23 podría representar una cepa transitoria entre el *Cluster 1* y el *Cluster 2*.

Se ha postulado que en un principio, el *Cluster 1* circulaba en mayor proporción que el *Cluster 2*, pero que éste fue desplazado gradualmente hasta que los virus del *Cluster 2* predominaron. Se ha estimado que para mediados del 2009, solo el 30% de los virus que circulaban correspondían al *Cluster Texas*, mientras que más del 70% corresponde al *Cluster NY* (Nelson et al, 2009). Finalmente, la proporción en la distribución de ambos grupos que se determinó en trabajos previos (Nelson et al, 2009), es congruente con nuestros resultados, ya que aproximadamente el 80% de los virus analizados se agrupan dentro del grupo NY, y sólo un 20% en el grupo TX.

## II.1 LA TRAYECTORIA EVOLUTIVA DEL AISLADO PORCINO Y DE LOS VIRUS DE LA GLORIA.

En este trabajo, se determinó la secuencia completa del genoma de un aislado viral obtenido durante la fase inicial de la pandemia de un cerdo que presentaba enfermedad respiratoria aguda en una granja ubicada en Querétaro, Querétaro. La relación filogenética de esta cepa en relación a otros virus pandémicos se infirió como se ha descrito en materiales y métodos. De interés, se encontró que al darle polaridad a los árboles de los genes HA, NA, NP y PA, el aislado porcino A/swine/4/Mexico/09 representa una rama basal para el resto de los virus pandémicos con un alto valor de soporte para el nodo divergente (aLRT:1; Anexo; Fig. 3). Posteriormente, se buscaron en las secuencias del aislado los marcadores moleculares distintivos para cada *Cluster* y sub-clado descritos anteriormente (Tabla 4, 5 y 6). Se encontró que aunque A/swine/4/Mexico/09 tiene los nueve nucleótidos característicos del *Cluster 1*, este virus sólo posee uno de los cuatro marcadores descritos para clasificar dentro del sub-clado 1: por lo tanto, no puede clasificarse (Tabla 4, 5 y 6). Dichos hallazgos son congruentes con el análisis filogenético.

La información obtenida a partir del análisis individual de los genes fue bastante limitada, ya que la diferencia genética entre todos los virus pandémicos es muy pequeña (porcentaje de identidad de entre virus pandémicos > 99.9%). Por lo tanto, la resolución de los árboles es baja y no permite aclarar la relación filogenética entre las secuencias virales. Para tratar de maximizar la resolución filogenética, se realizó análisis bayesiano con las regiones concatenadas que codifican para las proteínas HA y NA. Esto nos permitió integrar la señal filogenética de los dos genes con la tasa de evolución más alta en un solo árbol. En el árbol de Máxima Credibilidad de Clado obtenido (MCC), se observa que los virus de La Gloria descienden de la cepa *A/swine/4/Mexico/09*, y juntos forman un clado basal para el resto de las cepas pandémicas, soportado con un alto valor de probabilidad posterior (PP:1; Anexo; Fig. 4). El árbol de las regiones de HA-NA concatenadas construido bajo ML también sostiene la observación de que *A/swine/4/Mexico/09* es basal al resto de los virus pandémicos, con un valor de soporte alto (aLRT:1; Anexo; Fig. 4).

## II.2 POLIMORFISMOS PRESENTES EN *A/swine/4/Mexico/09*

De manera paralela, se identificaron todos los cambios presentes en los ocho genes virales de *A/swine/4/Mexico/09* con respecto a las secuencias de los otros virus pandémicos. Para ello, las secuencias de genes individuales se analizaron por separado en Seaview y Jalview, como se describió anteriormente. Se detectaron catorce cambios no sinónimos distribuidos a lo largo de seis genes: PB1 R126H, E636G; PA N222K, V407I; HA P200S, A214T, T220S; NP K77R, D53E ; NA S35G, I106V, P126H, D248N y NS1 V123I (el primer aminoácido antes de la posición representa el polimorfismo predominante en la población viral pandémica, mientras que el último aminoácido es el presente en *A/swine/4/Mexico/09*). Siete de los catorce cambios observados en el aislado de cerdo están presentes en menos del 1% de la población viral pandémica y no pandémica, por lo que corresponden a mutaciones. Por otro lado, los cambios V407I (PA) P200S, S220T (HA), D53E (NP), P126H y 106V (NA), y V123I (NS1) están presentes en más del 1% de la población viral y se consideran polimorfismos. Los sitios polimórficos en HA S220T, NA I106V, y NS1 V123I están presentes en 13-30% de la población viral pandémica 2009, y representan la minoría poblacional. Es importante señalar que los cambios 200, 214 y 220 en la proteína HA, se encuentran localizados en el sitio de reconocimiento de receptor. La posición 220 corresponde al sitio antigénico Ca y se considera un sitio variable para las cepas pandémicas (Lin et al, 2009). Ninguno de los



otros cambios identificados en las demás proteínas virales están localizados en una posición biológicamente relevante, o en estudios previos se han detectados de estar bajo presión de selección positiva.

De manera interesante y de acuerdo con nuestro análisis genético, hemos identificado que cuatro de los catorce polimorfismos (V407I- PA, P200S-HA, D53E-NP y P126H-NA) que se encuentran en el virus *A/swine/Mexico/4/2009* están altamente conservados entre la población del virus *A/H1N1* no pandémicos, independientemente de su origen animal, la procedencia geográfica o el año de su aislamiento (Tabla 7). Dichos polimorfismos ocurren en el 83 y el 99% de las secuencias estacionales analizadas. Por otra parte, mientras que estos cuatro cambios no sinónimos se fijaron en la mayoría de los virus estacionales *H1N1* humanos y animales, estos cambios están ausentes en la mayoría de la población viral pandémica (~1%; Tabla 7). Además de que se encontró que muy pocos de los virus pandémicos poseen alguno de los polimorfismos ancestrales, se observó que solamente el virus *A/swine/Mexico/4/2009* posee los cuatro cambios distribuidos en cuatro genes. Es importante notar que estos cuatro polimorfismos se distribuyen en los cuatro genes virales en cuyo análisis filogenético, *A/swine/4/Mexico/09* resulta en una rama basal al resto de los virus pandémicos.

Por otra parte, existe evidencia de que los polimorfismos encontrados no se deben a un artefacto originado por la adaptación al crecimiento en huevo. En primer lugar, los virus utilizados para este estudio que fueron crecidos en huevo no comparten los polimorfismos encontrados en el aislado de origen porcino. En segundo lugar, las mutaciones que son resultado de la adaptación a crecimiento en huevo suelen alterar las propiedades de unión al receptor y ocurren principalmente en las proteínas HA y NA (Gambaryan et al, 1999; Brookes et al, 2010; Mochalova et al, 2003; Matrosovich et al, 2004). Estas mutaciones se han caracterizado adecuadamente en dichos trabajos; por ejemplo, se ha encontrado que las mutaciones presentes en los virus *H1N1* crecidos en huevo ocurren principalmente en las posiciones 190, 225, 226 163, 187, 189 de la proteína HA. Cabe destacar que los virus de La Gloria tienen la mutación Q226R, típica de la adaptación a crecimiento en huevo. La mutación Q226R se ha observado con anterioridad en cepas de virus humanos *H1N1* estacionales y pandémicos que se han cultivado en huevo (Brookes et al, 2010; al Mochalova et al, 2003; al Matrosovich et al, 2004).

Los polimorfismos hallados apoyan la hipótesis de que el aislado porcino *A/swine/4/Mexico/09* posee rasgos ancestrales no compartidos con otros virus pandémicos, y sugiere que aunque el virus de origen porcino está estrechamente

relacionado con todos los otros virus pandémicos de 2009, éste podría pertenecer a un linaje hermano que co-circuló junto con el linaje pandémico durante la fase inicial de la pandemia. Es probable que este clado hermano se originara a partir de un evento de transmisión epizoonótico de hombre a cerdo que ocurrió de manera temprana, como se ha descrito en trabajos previos (Wells et al, 1991; Shinde et al 2009). Sin embargo, dada la deficiente vigilancia epidemiológica y muestreo en las poblaciones virales porcinas en México y en el mundo, la ausencia de otros virus tipo A/swine/4/Mexico/09 no nos permite sostener dicha conclusión.

Muestra	Proteína (aa)	PA (407)	HA (200)	NP (53)	NA (126)
A/swine/4/Mexico/09		Ile (I)	Pro (P)	Glu (E)	His (H)
A/H1N1/2009-2010 pdm		Val >99%	Ser >99%	Asp >98%	Pro >99%
		Ile <1%	Pro <1%	Glu <2%	His <1%
Otros virus A/H1N1		Ile >94%	Pro >83%	Glu >99%	His >98%
		Val <6%	Ser >15%	Asp <1%	Pro <1%

**Tabla 7. Polimorfismos Ancestrales en el Aislado Viral A/swine/4/Mexico /09**

Polimorfismos en las proteínas HA, NA, NP y PA de A/swine/4/Mexico/09, conservados entre el linaje de virus A/H1N1 no pandémicos. Los polimorfismos están presentes en aprox. el 1% de las secuencias pandémicas.

### II.3 ANÁLISIS EVOLUTIVO DE LOS VIRUS A/H3N2 ESTACIONALES DE MÉXICO

De los virus secuenciados para la estación 2002-2003, las muestras 756/2003, 2662/2003 y 2664/2003 fueron caracterizadas de manera previa como virus tipo A/Korea/770/2002, (INDRE, CDC). Aunque se ha determinado que los virus tipo A/Korea/770/2002 y A/Fujian/411/2002 son antigénicamente equivalentes, nuestro análisis filogenético de HA y NA revela que estos virus se distinguen en dos clados definidos (Anexo, Fig. 6 y 7). El clado Fujian se caracteriza por estar relacionado de manera cercana con la cepa A/California/07/2004 y ser sensible a los amantadanos. Por otro lado, en el linaje Korea agrupan los virus 940/2003, 756/2003, 835/2003, junto con los virus de referencia A/Yucatán/ME6057/3003, A/Yucatán/844/2003 y el virus resistente a los amantadanos A/HongKong/CUHK50895/2003.

De manera interesante, mientras que el gen HA del virus 2662/2003 corresponde a virus del linaje Korea (como se caracterizó antigénicamente), la proteína NA corresponde

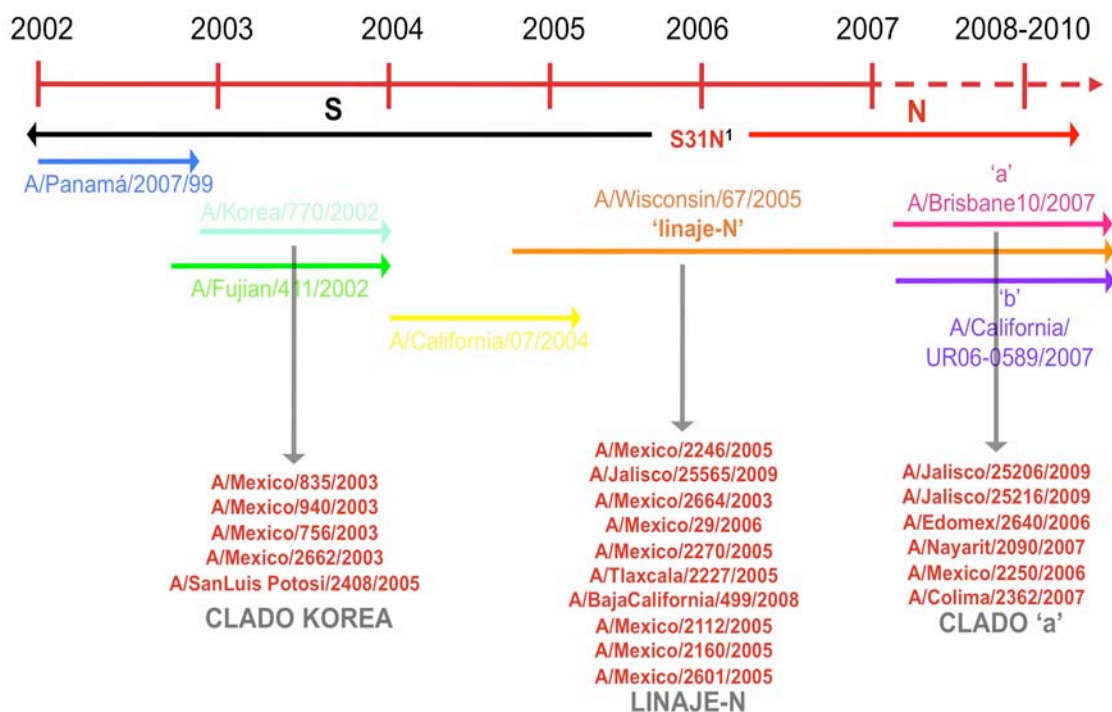
a virus tipo Fujian. Este virus posee 6 de 8 genes de los virus tipo A/Panamá/2007/99 (PB2, PB1, PA, NP, M1/M2 y NS1/NS2) (Anexo; Figuras 5, 6, 7 y 8). Asimismo, para el gen NS1/NS2, el virus 2662 agrupa de manera cercana con virus porcinos de origen euroasiáticos de 2004-2006. Esto sugiere que este virus se originó a partir de un evento de rearreglo génico entre virus de distintas estaciones, y que el gen NS1/NS2 del virus 2662/2003 es de origen porcino, lo que podría indicar un evento de re-arreglo génico entre cepas de origen humano y porcino (Anexo; Fig. 8). Por otro lado, el virus 2408/2005 también agrupa dentro del clado Korea en el análisis filogenético de HA y M, mientras que en NA y NS1/2 pertenece al linaje-N. Asimismo, este virus tiene proteínas PB2, PB1, PA, NP tipo A/Panamá/2007/99, agrupando de manera cercana el virus 2662/2003 en estos genes, lo que sugiere que este virus podría ser un rearreglante entre virus de tres distintos linajes (Anexo; Figuras 5, 6, 7 y 8). Estos hallazgos son congruentes con algunos trabajos previos que proponen que la evolución de los virus de influenza es más compleja de lo que reflejan los modelos estándar de deriva antigénica. En las poblaciones pueden persistir múltiples linajes de virus antigénicamente diferentes, y los eventos de rearreglos génicos son frecuentes, lo que contribuye de manera directa a la evolución viral a largo plazo (Holmes et al 2005, Nelson et al, 2008).

También, se encontró que el virus 2664/2003, aunque similar antigénicamente a los virus tipo Korea, posee la mutación S31N y agrupa dentro del linaje-N en el análisis de cada uno de los genes virales, incluyendo HA (Anexo; Figuras 5, 6, 7 y 8). Por lo tanto, pertenece al linaje-N. Esto es congruente con las observaciones de que los virus tipo A/Wisconsin/67/2005 circularon desde 2003-2004, y confirma la estrecha relación de los virus del linaje-N con algunos aislados del 2003. Asimismo, la cepa de referencia A/HongKong/CUHK50895/2003 agrupa en NA y HA dentro del clado Korea, mientras que en M2 agrupa dentro del linaje-N, lo que es afín a la hipótesis de que los virus de este tipo fueron los que dieron origen a la mutación S31N en el linaje-N mediante un evento de rearreglo génico.

Se identificó previamente que los virus 29/2006, 2227/2005, 2246/2005 eran antigénicamente similares a la cepa A/Wisconsin/67/2005, mientras que los virus 2112/2005, 2160/2005 y 2601/2005 a la A/California/07/2004. En el análisis filogenético de los ocho genes virales (Anexo; Figuras 5, 6, 7 y 8), ninguno de estos virus agrupa dentro del clado Fujian. En los árboles obtenidos con los genes HA y NA, se observa que los virus 2270/2005, 2664 /2003, 25565/2009, 29/2006, 2112/2005, 2160/2005, 2227/2005, 2246/2005, 2601/2005 y 499/2008 agrupan dentro del linaje-N, mientras que los virus 2640/2006, 2090/2007, 25216/2009, 25206/2009, 2362/2007 y 2250/2006

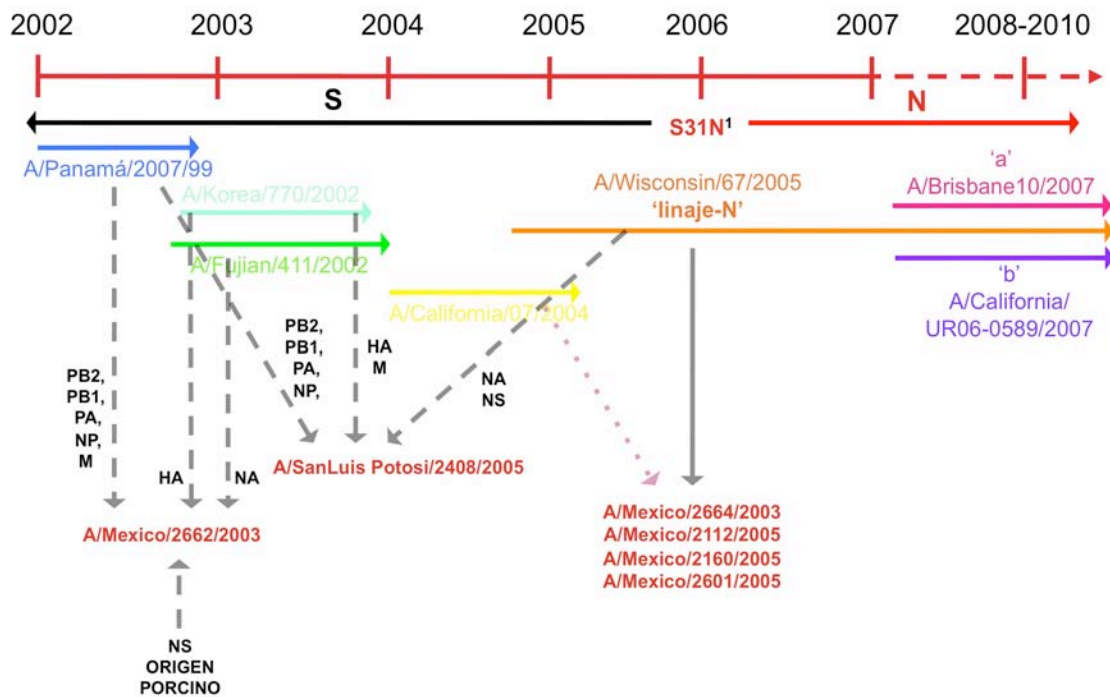
pertenecen al clado 'a', que deriva del linaje-N. El clado 'a' se caracteriza por estar relacionado de manera cercana con la cepa de vacunación A/Brisbane/10/2007 y ser resistente a los amantadanos, al poseer la mutación S31N en la proteína M2. Los virus 2664/2003 y 25565/2009 pertenecen al linaje-N en todos sus genes, lo que sugiere que los virus tipo A/Wisconsin/667/2005 llevan mas de 6 años en circulación, desde 2003 hasta 2009 (Anexo; Figuras 5, 6, 7 y 8).

Finalmente, para observar el cambio en la frecuencia y distribución de la mutación S31N a lo largo del tiempo y de la historia evolutiva de los virus H3N2 muestreados, se proyectaron sobre el árbol de HA las secuencias que tienen el polimorfismo 31N, detectadas en el alineamiento en aminoácidos de M2 (secuencias en rojo, Anexo; Fig. 5). A partir de esto, podemos concluir, como se ha observado previamente, que la frecuencia de la mutación S31N ha ido en aumento a lo largo del tiempo. Desde 1999 hasta 2003, el cambio S31N no estaba presente en la mayoría de los aislados (detectado en cinco virus, dos de los cuales son de origen porcino), mientras que para los virus colectados en 2005 y en adelante, mas del 99% de los virus presentan el cambio. Asimismo, podemos observar que a partir del surgimiento de virus tipo A/HongKong/CUHK50895/2003, la frecuencia de S31N aumentó considerablemente. La gran mayoría de los virus resistentes a los amantadanos agrupa dentro de los clados A/Wisconsin/67/2005 (naranja) y A/Brisbane/10/2007 (rosa). De manera contrastante, en los linajes Panamá (azul), Korea (verde), Fujian (amarillo), la proporción de virus resistentes es mínima (Anexo; Fig. 5).



**Figura 6. Linajes de los Virus H3N2 Secuenciados en base al Análisis de HA y NA**

El siguiente esquema representa un línea del tiempo con los principales linajes virales H3N2 a los que pertenecen las muestras secuenciadas, en base al análisis filogenético de HA y NA. En este trabajo, se encontró que a pesar de ser equivalentes antigénicamente, los virus tipo A/Fujian411/2002 y A/Korea/770/2002 forman clados definidos. Cada linaje está representado con distintos colores; A/Panamá2007/199 en azul, A/Fujian411/2002 en verde, A/Korea/770/2002 en verde aqua, A/California/07/2004 en amarillo, el linaje-N, o virus tipo A/Wisconsin/67/2005 en naranja y los linajes 'a' tipo A/Brisbane/10/2007 en rosa y 'b' tipo A/California/UR06-0589/2007 en morado. <sup>1</sup>La prevalencia en la resistencia a amantadanos aumentó de un 8.2% en 2002 a un 73.8% en 2004-2005 (CDC).



**Figura 7. Eventos de Rearreglos Génicos entre los Virus H3N2 Secuenciados**

El siguiente esquema representa una línea del tiempo con los principales linajes virales H3N2. Las flechas punteadas en negro, indican los eventos de rearrreglos génicos que dieron lugar a los diferentes segmentos génicos de los virus analizados en este trabajo (A/Mexico/2662/2003 y A/SanLuisPotosi/2408/2005). La flecha punteada en rosa indica que la identidad antigénica de los virus A/Mexico/2664/2003, 2112/2005, 2160/2005 y 2601/2005 no coincide con su identidad evolutiva, pues estos fueron caracterizados como similares a la cepa A/California/07/2004, pero pertenecen al linaje-N en el análisis filogenético de sus genes. Cada linaje está representado con distintos colores; A/Panamá2007/199 en azul, A/Fujian411/2002 en verde, A/Korea/770/2002 en verde aqua, A/California/07/2004 en amarillo, el linaje-N, o virus tipo A/Wisconsin/67/2005 en naranja y los linajes 'a' tipo A/Brisbane/10/2007 en rosa y 'b' tipo A/California/UR06-0589/2007 en morado. <sup>1</sup>La prevalencia en la resistencia a amantadanos aumentó de un 8.2% en 2002 a un 73.8% en 2004-2005 (CDC).

## CONCLUSIONES

### I. ANÁLISIS DE LAS SECUENCIAS

En total, se obtuvieron quince muestras de virus A/H1N1/2009 pandémicos humanos y un aislado porcino, que se colectaron en diferentes puntos de la república Mexicana durante la pandemia de A/H1N1/2009. Se obtuvieron veintiún aislados clínicos adaptados a cultivo para los virus H3N2, que se colectaron en un periodo de ocho años en distintos lugares de la República Mexicana. Algunos de los virus tipo H3N2 se caracterizaron antigénicamente de manera previa en el InDRE/CDC. En este trabajo, se probaron cuatro diferentes métodos para la preparación de bibliotecas para secuenciación masiva (Solexa, Illumina), con el fin de tratar de enriquecer el material inicial, y así poder maximizar la cobertura y resolución en el proceso de secuenciación. Se comprobó que los cuatro métodos son efectivos para obtener la secuencia genómica completa del virus de influenza. El método a utilizar debe seleccionarse en relación a la cantidad de RNA inicial en la muestra, y sobre todo en la calidad del producto obtenido de la amplificación del genoma por M-RTPCR.

### II. VIRUS A/H1N1 PANDÉMICOS

A partir de las secuencias obtenidas, se llevó a cabo análisis filogenético de los virus A/H1N1/2009 pandémicos con métodos robustos como el criterio de máxima verosimilitud y análisis bayesiano, y se buscaron en las secuencias de las muestras los marcadores moleculares distintivos para cada *Cluster* y sub-clado descritos anteriormente. A partir de esto se determinó que los virus con fechas de colección temprana agrupan dentro del *Cluster 1/ sub-clado 2*. Por otro lado, se observó que la mayoría de los virus colectados en México con fecha de colección tardía, agrupan entre el sub-clado 7 del *Cluster 2*. Las muestras que pertenecen al sub-clado 7 tienen la mutación I123V en la proteína NS1, 220T en HA y V06I en NA, que definen a dicho sub-clado. La distribución de los marcadores moleculares de *Clusters* y sub-clados en las secuencias coincide con el análisis filogenético.

Al hacer análisis filogenético y darle polaridad a las topologías, se observa con valor de soporte alto en los árboles, que el aislado porcino A/swine/4/Mexico/09 representa una rama basal para el resto del clado pandémico en cuatro genes virales (HA, NA NP y PA). Por otro lado, el virus de La Gloria, parecen descender del aislado

porcino, formando juntos un clado basal para el resto de las cepas pandémicas. El análisis genético demostró que A/swine/4/Mexico/09 posee rasgos moleculares ancestrales no compartidos con otros virus A/H1N1/pandémicos, pero ampliamente conservados entre el linaje de virus H1N1 estacionales, sin importar el origen animal, año o lugar de colección. Esto sugiere que aunque el aislado de cerdo está estrechamente relacionado con los demás virus pandémicos de 2009, podría pertenecer a un linaje hermano que co-circuló con el clado pandémico en cerdos durante la fase inicial de la pandemia.

### III. VIRUS A/H3N2 ESTACIONALES

Una cuestión fundamental que permanece sin resolver es si la evolución de los virus H3N2 en las epidemias es consecuencia de la persistencia de virus a niveles bajos, o si los linajes epidémicos se originan a partir de cepas semillas que salen de regiones específicas cada año. Para observar el patrón evolutivo de los virus H3N2 que han circulado en México, se secuenciaron varios aislados de virus tipo H3N2 adaptados a cultivo y se hizo análisis filogenético bajo el criterio de máxima verosimilitud. A partir del análisis evolutivo se encontró que algunos aislados, como el 2662/2003 y 2408/2005, se originaron por múltiples eventos de re-arreglos génicos entre cepas de distintos linajes. Asimismo, el virus 2662 agrupa de manera cercana con virus porcinos de origen euroasiáticos de 2004-2006 en el gen NS1/NS2. Esto sugiere que el gen NS1/NS2 en el virus 2662/2003 es de origen porcino, y podría indicar una vez más un evento de re-arreglos génicos entre cepas de origen humano y porcino. Asimismo, otros virus como 2664/2003 y 25565/2009 pertenecen al linaje-N, lo que sugiere que los virus de este linaje llevan más de 6 años en circulación. Nuestros hallazgos apoyan las observaciones previas que proponen que la evolución de los virus de influenza es más compleja de lo que reflejan los modelos estándar de deriva antigénica. Múltiples linajes de virus antigénicamente diferentes pueden persistir en las poblaciones, mientras que los eventos de rearrreglos génicos son frecuentes y contribuyen de manera directa a la evolución viral a largo plazo (Holmes et al 2005, Nelson et al, 2008).



## BIBLIOGRAFÍA

- Barnett, J. M., A. Cadman, F. M. Burrell, S. H. Madar, A. P. Lewis, M. Tisdale, and R. Bethell. 1999. In vitro selection and characterisation of influenza B/Beijing/1/87 isolates with altered susceptibility to zanamivir. *Virology* 265:286-95.
- Bergmann, M., A. Garcia-Sastre, E. Carnero, H. Pehamberger, K. Wolff, P. Palese, and T. Muster. 2000. Influenza virus NS1 protein counteracts PKR-mediated inhibition of replication. *J Virol* 74:6203-6.
- Bioinformatics, 2009, 25(12) 1554-1555
- Bouvier NM, Palese P. The biology of influenza viruses. *Vaccine* 2008; 26 Suppl 4:D49-53.
- Bright, R. A., D. K. Shay, B. Shu, N. J. Cox, and A. I. Klimov. 2006. Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. *Jama* 295:891-4.
- Brookes SM, Núñez A, Choudhury B, Matrosovich M, Essen SC, et al. (2010) Replication, Pathogenesis and Transmission of Pandemic (H1N1) 2009 Virus in Non-Immune Pigs. *PLoS ONE* 5(2): e9068. doi:10.1371
- Brown IH. The epidemiology and evolution of influenza viruses in pigs. *Vet Microbiol* 2000; 74 (1-2):29-46.
- CDC, Abril 18 2008/57(15),393-398
- CDC, <http://www.cdc.gov/flu/about/qa/season.htm>
- CDC, <http://www.cdc.gov/flu/weekly/weeklyarchives20032004/0304summary.htm>
- CDC, [http://www.cdc.gov/flu/weekly/weeklyarchives2005-2006/05\\_06summary.htm](http://www.cdc.gov/flu/weekly/weeklyarchives2005-2006/05_06summary.htm)
- Chan AY, Vreede FT, Smith M, Engelhardt OG, Fodor E. Influenza virus inhibits RNA polymerase II elongation. *Virology*. 2006 Jul 20;351(1):210-7.
- Chen W, Calvo PA, Malide D et al, A novel influenza A virus mitochondrial protein that induces cell death. *Nat Med* 2001; 7 (12):1306-12.
- Colman PM. New antivirals and drug resistance. *Annu Rev Biochem* 2009; 78:95-118.

Conenello GM, Zamarin D, Perrone LA et al, A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. PLoS Pathog 2007; 3 (10):1414-21.

Connor, R. J., Y. Kawaoka, R. G. Webster, and J. C. Paulson. 1994. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. Virology 205:17-23.

Cox NJ, Subbarao K, Global epidemiology of influenza: past and present, Annu Rev Med. 2000; 51:407-21.

de Jong MD, Simmons CP, Thanh TT et al, Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. Nat Med 2006; 12 (10):1203-7.

Dharan NJ, Gubareva LV, Meyer JJ et al, Infections with oseltamivir-resistant influenza A(H1N1) virus in the United States. Jama 2009; 301 (10):1034-41.

Drake JW. Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci U S A 1993; 90 (9):4171-5.

Drummond AJ & Rambaut A (2007) "BEAST: Bayesian evolutionary analysis by sampling trees." BMC Evolutionary Biology 7, 214

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Fouchier RA, Munster V, Wallensten A et al, Characterization of a novel influenza A virus hemagglutinin sub-type (H16) obtained from black-headed gulls. J Virol 2005; 79 (5):2814-22.

Fraser C, Donnelly CA, Cauchemez S et al, Pandemic potential of a strain of influenza A (H1N1): early findings. Science 2009; 324 (5934):1557-61.

Futuyma, D.J. 1998. Evolutionary Biology. 3rd ed. Sinauer Associates, Sunderland, Massachusetts.

Garcia-Garcia, J., and C. Ramos. 2006. [Influenza, an existing public health problem]. Salud Publica Mex 48:244-67.

García-Sastre A, Mechanisms of inhibition of the host interferon alpha/beta-mediated antiviral responses by viruses, Microbes Infect. 2002 May;4(6):647-55.

Garten RJ, Davis CT, Russell CA et al, Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 2009; 325 (5937):197-201.

Gastaminza P, Perales B, Falcón AM, Ortín J. Mutations in the N-terminal region of influenza virus PB2 protein affect virus RNA replication but not transcription. *J Virol*. 2003 May;77(9):5098-108.

Glaser, L., J. Stevens, D. Zamarin, I. A. Wilson, A. Garcia-Sastre, T. M. Tumpey, C. F. Basler, J. K. Taubenberger, and P. Palese. 2005. A single amino acid sub-stitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. *J Virol* 79:11533-6.

Greninger AL., et al. A metagenomic analysis of pandemic influenza A (2009 H1N1) infection in patient from North America. *Plos One* 5(10) e13381

Hatada, E., S. Saito, and R. Fukuda. 1999. Mutant influenza viruses with a defective NS1 protein cannot block the activation of PKR in infected cells. *J Virol* 73:2425-33.

Hatta M, Hatta Y, Kim JH et al, Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. *PLoS Pathog* 2007; 3 (10):1374-9.

Hayden, F. G., W. J. Hall, R. G. Douglas, Jr., and D. M. Speers. 1979. Amantadine aerosols in normal volunteers: pharmacology and safety testing. *Antimicrob Agents Chemother* 16:644-50.

Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y, et al, (2005) Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol* 3(9): e300.

Ito T, Couceiro JN, Kelm S et al, Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *J Virol* 1998; 72 (9):7367-73.

Jackson D, Hossain MJ, Hickman D et al, A new influenza virus virulence determinant: the NS1 protein four C-terminal residues modulate pathogenicity. *Proc Natl Acad Sci U S A* 2008; 105 (11):4381-6.

Jiao P, Tian G, Li Y et al, A single-amino-acid sub-stitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. *J Virol* 2008; 82 (3):1146-54.

Jung K, Song DS. Evidence of the co-circulation of influenza H1N1, H1N2 and H3N2 viruses in the pig population of Korea. *Vet Rec* 2007; 161 (3):104-5.

- Katoh K, Asimenos G, Toh H. (2009) Multiple Alignment of DNA Sequences with MAFFT. In 'Bioinformatics for DNA Sequence Analysis' edited by D. Posada, Methods in Molecular Biology 537:39-64
- Kawaoka Y, Webster RG. Sequence requirements for cleavage activation of influenza virus hemagglutinin expressed in mammalian cells. Proc Natl Acad Sci U S A 1988; 85.
- Kawaoka, Ed. Caister Influenza Virology: Current Topics. Academic Press, Wymondham, UK, 2006
- Knipe, David M., and Peter M. Howley. Fields Virology. 5th ed. Philadelphia: Lippincott Williams & Wilkins.
- Kuroda, M., Katano, H., Nakajima, N. 2010 Characterization of quasispecies of pandemic 2009 influenza A virus (A/H1N1/2009) by de novo sequencing using a next-generation DNA sequencer. Plos one 5(4) e10256
- Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18(11): 1851–1858
- Li Z, Jiang Y, Jiao P et al, The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. J Virol 2006; 80 (22):11115-23.
- Lin T, Wang G, Li A, et al. The hemagglutinin structure of an avian. H1N1 influenza A virus. Virology, 2009, 392: 73–81.
- Lycett, S. J., M. J. Ward, F. I. Lewis, A. F. Poon, S. L. Kosakovsky Pond, and A. J. Leigh Brown. 2009. Detection of mammalian virulence determinants in highly pathogenic avian influenza H5N1 viruses: multivariate analysis of published data. J Virol.
- Ma W, Lager KM, Vincent AL, Janke BH, Gramer MR, Richt JA. The Role of swine in the Generation of Novel Influenza Viruses. Zoonoses Public Health. 2009 May 20.
- MapView: visualization of short reads alignment on a desktop computer. Bioinformatics, 2009, 25 (12) : 1554-1555
- Massin P, van der Werf S, Naffakh N. Residue 627 of PB2 is a determinant of cold sensitivity in RNA replication of avian influenza viruses. J Virol 2001; 75 (11):5398-404.

Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk H-D. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc Natl Acad Sci U S A*. 2004;101:4620–4624.

McAuley JL, Hornung F, Boyd KL et al, Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe* 2007; 2 (4):240-9.

Mochalova L, Gambaryan A, Romanova J, Tuzikov A, Chinarev A, et al. Receptor-binding properties of modern human influenza viruses primarily isolated in Vero and MDCK cells and chicken embryonated eggs. *Virology*. 2003;313:473–480.

Moscona, A. 2004. Oseltamivir-resistant influenza? *Lancet* 364:733-4.

Naeve, C. W., V. S. Hinshaw, and R. G. Webster. 1984. Mutations in the hemagglutinin receptor-binding site can change the biological properties of an influenza virus. *J Virol* 51:567-9.

Nakamura, S., Yang, ChS., Sakon, N et al. 2009 Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *Plos One* 4(1) e4219

Nelson M.I., Simonsen L., Viboud C., Miller M.A., Holmes E.C. The origin and global emergence of adamantane resistant A/H3N2 influenza viruses (2009) *Virology*, 388 (2), pp. 270-278.

Nelson MI, Edelman L, Spiro DJ, Boyne AR, Bera J, et al, 2008 Molecular Epidemiology of A/H3N2 and A/H1N1 Influenza Virus during a Single Epidemic Season in the United States. *PLoS Pathog* 4(8): e1000133. doi:10.1371/journal.ppat.1000133

Neumann G, Noda T, Kawaoka Y. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* 2009; 459 (7249):931-9.

Nobusawa E., Sato K. 2006 Comparison of the mutation rates of human influenza A and B viruses. *J. Virol.* 80, 3675–3678

OMS 2009, posting date. Influenza (Seasonal) N211. OMS. [Online.]

OMS; Hoja de Datos de Influenza N°211, Revisada en 2003-2009 (<http://www.who.int/mediacentre/factsheets/fs211/en/>, [http://www.who.int/csr/don/2009\\_04\\_30\\_a/en/index.html](http://www.who.int/csr/don/2009_04_30_a/en/index.html)).

Palese, A. Garcia-Sastre, J. M. Katz, and T. M. Tumpey. 2008. Single gene reassortants identify a critical role for PB1, HA, and NA in the high virulence of the 1918 pandemic influenza virus. *Proc Natl Acad Sci U S A* 105:3064-9.

Palese, P., and M. L. Shaw. 2007. Orthomyxoviridae: The Viruses and Their Replication, p. 1647-1690. In D. Knipe and P. Howley (ed.), *Fields Virology*, vol. II. Lippincott Williams & Wilkins, Philadelphia.

Pappas, C., P. V. Aguilar, C. F. Basler, A. Solorzano, H. Zeng, L. A. Perrone,

Pensaert M, Ottis K, Vandeputte J et al, Evidence for the natural transmission of influenza A virus from wild ducts to swine and its potential importance for man. *Bull World Health Organ* 1981; 59 (1):75-8.

Posada D and Crandall KA 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14 (9): 817-818.

Prevention and control of influenza: recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm Rep* 48:1-28.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Ramakrishnan MA, TU ZJ, Sing S. et al, 2009 The feasibility of using high resolution genome sequencing of influenza A viruses to detect mixed infections and quasispecies. *Plos One*, 4(9) e7105

Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from <http://beast.bio.ed.ac.uk/Tracer>

Rambaut A, Pybus O, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. *Nature*. 2008;453:615–19.

Rimmelzwaan GF, Shaw MW, Skepner E, Stohr K, Tashiro M, Fouchier RA, Smith DJ. The global circulation of seasonal influenza A (H3N2) viruses. *Science*. 2008 Apr 18;320 (5874):340-6.

Rogers GN, Paulson JC. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* 1983; 127 (2):361-73.

Rogers, G. N., and B. L. D'Souza. 1989. Receptor binding properties of human and animal H1 influenza virus isolates. *Virology* 173:317-22.

Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, Gust ID, Hampson AW, Hay AJ, Hurt AC, de Jong JC, Kelso A, Klimov AI, Kageyama T, Komadina N, Lapedes AS, Lin YP, Mosterin A, Obuchi M, Odagiri T, Osterhaus AD,

Seo SH, Hoffmann E, Webster RG. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nat Med* 2002; 8 (9):950-4.

Shinde V, Bridges CB, Uyeki TM, et al. Triple-reassortant swine influenza A (H1) in humans in the United States, 2005–2009. *N Engl J Med* 2009.

Shope R, Lewis P. swine influenza:experimental transmission and pathology. *J Exp Med* 1931; 54.

Simonsen L, Viboud C, Grenfell BT, Dushoff J, Jennings L, Smit M, Macken C, Hata M, Gog J, Miller MA, Holmes EC. The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. *Mol Biol Evol.* 2007;24:1811–1820.

Skehel, J. J., A. J. Hay, and J. A. Armstrong. 1978. On the mechanism of inhibition of influenza virus replication by amantadine hydrochloride. *J Gen Virol* 38:97-110.

Smith GJ, Bahl J, Vijaykrishna D, Zhang J, Poon LL, Chen H, Webster RG, Peiris JS, Guan Y, Dating the emergence of pandemic influenza viruses. *Proc Natl Acad Sci U S A.* 2009 Jul 14;106(28):11709-12.

Stegmann T, White JM, Helenius A. Intermediates in influenza induced membrane fusion. *EMBO J.* 1990 Dec;9(13):4231-41.

Stephenson I, Democratis J. Influenza: current threat from avian influenza. *Br Med Bull.* 2006 May 1;75-76:63-80.

Stevens J, Corper AL, Basler CF et al, Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* 2004; 303 (5665):1866-70.

Stevens, J., O. Blixt, L. Glaser, J. K. Taubenberger, P. Palese, J. C. Paulson, and I. A. Wilson. 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *J Mol Biol* 355:1143-55.

Strauss J, Strauss E. Minus-strand RNA viruses. In, translator and editor Viruses and human disease. Second edn. Vol. 4. Oxford, UK: Elsevier; 2008; p. 137-75.

Strauss, J., and E. Strauss. 2008. Minus-strand RNA Viruses., p. 137-175, Viruses and Human sidease, 2 ed. Elsevier, oxford.

Taubenberger JK, Reid AH, Lourens RM et al, Characterization of the 1918 influenza virus polymerase genes. Nature 2005; 437 (7060):889-93.

Tran TH, Nguyen TL, Nguyen TD et al, Avian influenza A (H5N1) in 10 patients in Vietnam. N Engl J Med 2004; 350 (12):1179-88.

Trifonov V, Khiabanian H, Rabadan R. Geographic dependence, surveillance, and origins of the 2009 influenza A (H1N1) virus. N Engl J Med 2009; 361 (2):115-9.

Tumpey, T. M., T. R. Maines, N. Van Hoeven, L. Glaser, A. Solorzano, C. Pappas, N. J. Cox, D. E. Swayne, P. Palese, J. M. Katz, and A. Garcia-Sastre. 2007. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. Science 315:655-9.

Vincent AL, Lager KM, Ma W et al, Evaluation of hemagglutinin sub-type 1 swine influenza viruses from the United States. Vet Microbiol 2006; 118 (3-4):212-22.

Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL. Microarray-based detection and genotyping of viral pathogens. Proc Natl Acad Sci U S A. 2002 Nov 26;99(24):15687-92. Epub 2002 Nov 12.

Webby RJ, Swenson SL, Krauss SL et al, Evolution of swine H3N2 influenza viruses in the United States. J Virol 2000; 74 (18):8243-51.

Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses, Microbiol Rev. 1992 Mar;56(1):152-79

Wells DL, Hopfensperger DJ, ArdenNH, et al. swine influenza virus infections: transmission from ill pigs to humans at a Wisconsin agricultural fair and subsequent probable person-to-person transmission. JAMA 1991;265:478-81.

Wise HM, Foeglein A, Sun J, Dalton RM, Patel S, Howard W, Anderson EC, Barclay WS, Digard P. A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. J Virol. 2009;83:8021–8031



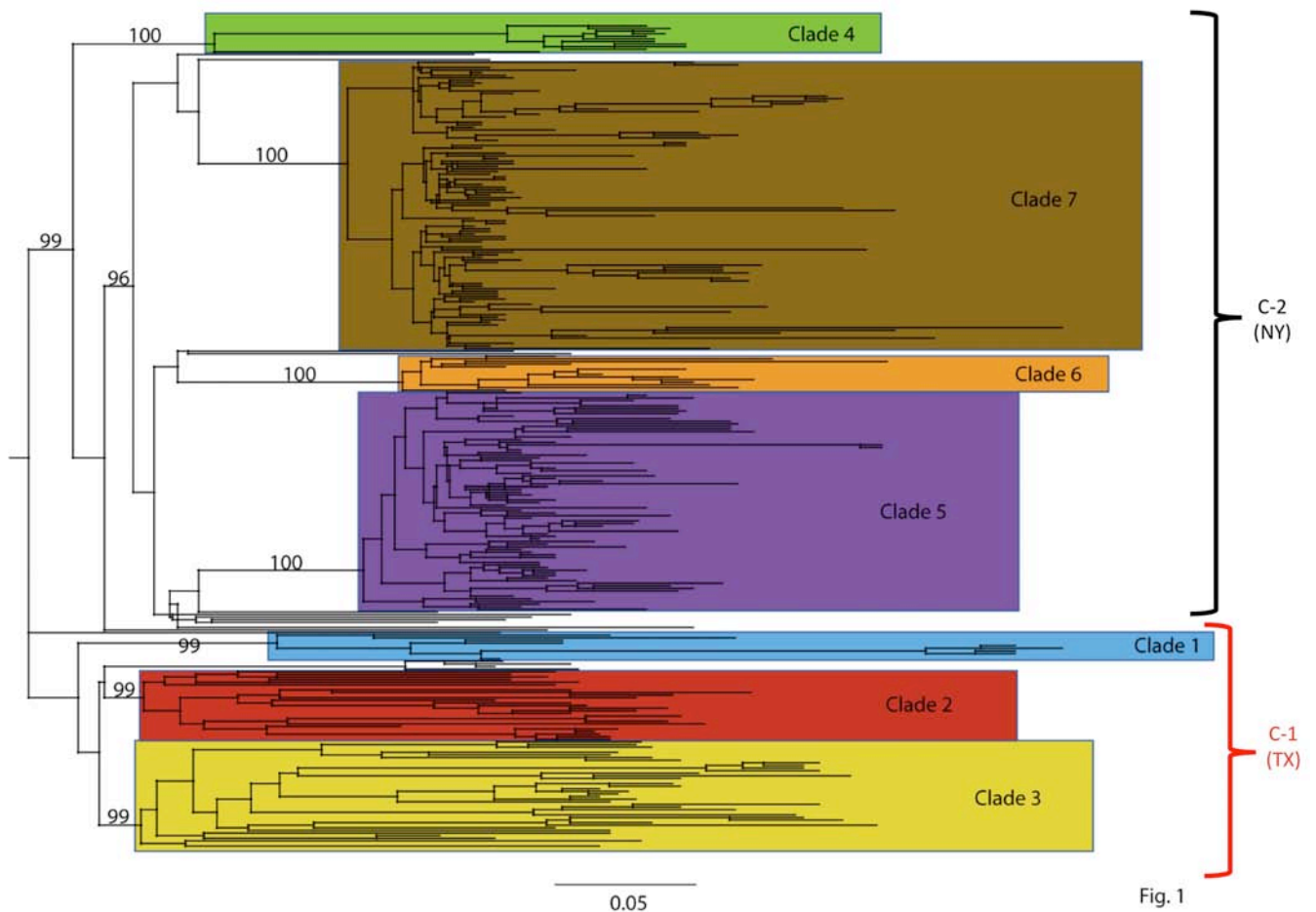
Wright, P., G. Neumann, and Y. Kawaoka. 2007. Orthomyxoviridae, p. 1691-1740. In D. Knipe and P. Howley (ed.), *Fields Virology*, vol. II. Lippincott Williams &Wilkins, Philadelphia.

Yongfeng H., fan Y., Jie D. et al. Direct pathogen detection from swab samples using a new high-throughput sequencing technology. *Clin Microbiol Infect* 2010

Zamarin D, Garcia-Sastre A, Xiao X et al, Influenza virus PB1-F2 protein induces cell death through mitochondrial ANT3 and VDAC1. *PLoS Pathog* 2005; 1 (1):e4.

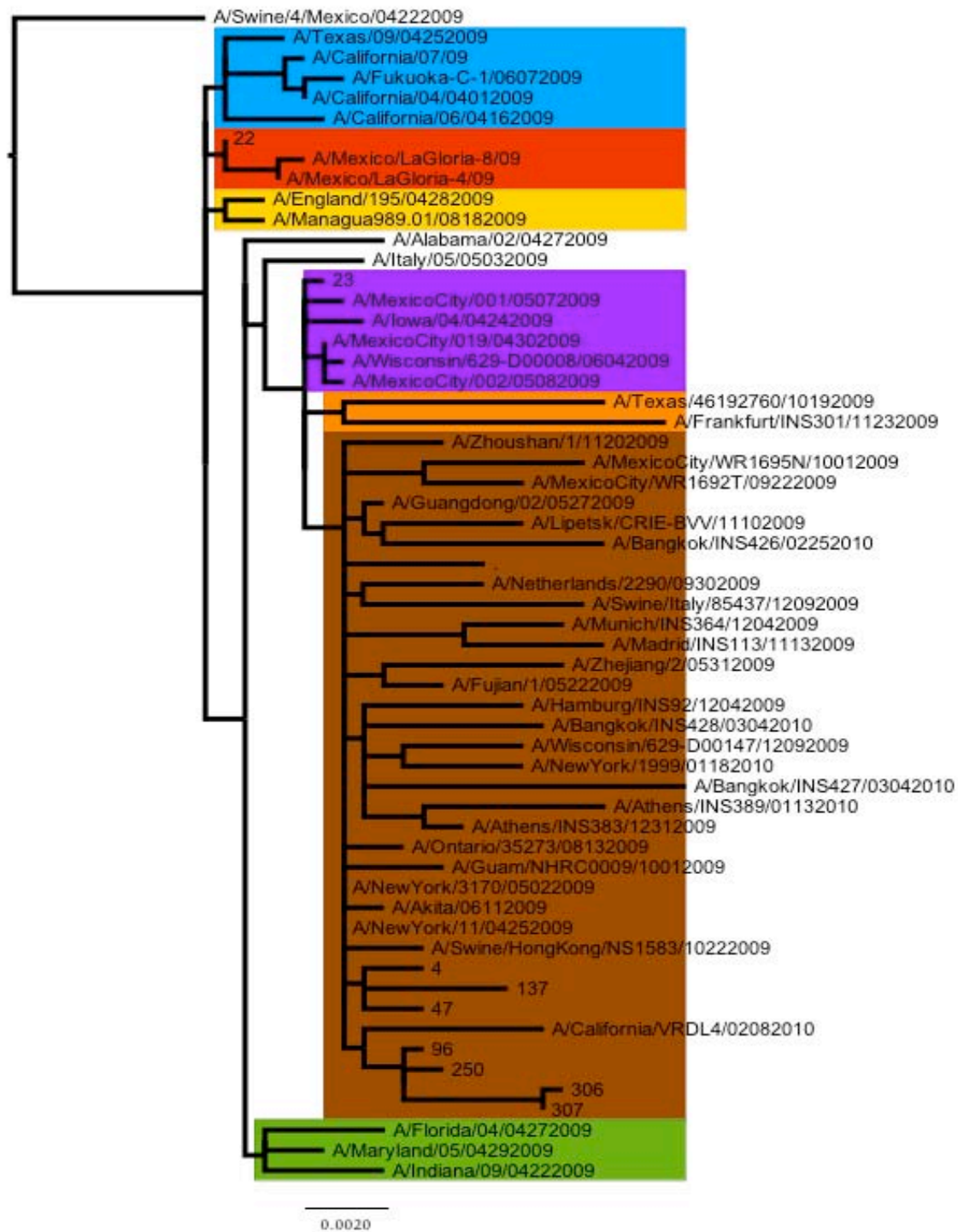
Zürcher T, Yates PJ, Daly J, Sahasrabudhe A, Walters M, Dash L, Tisdale M, McKimm-Breschkin JL. Mutations conferring zanamivir resistance in human influenza virus N2 neuraminidases compromise virus fitness and are not stably maintained in vitro, *J Antimicrob Chemother*. 2006 Oct;58(4):723-32. Epub 2006 AuG.

## ANEXO



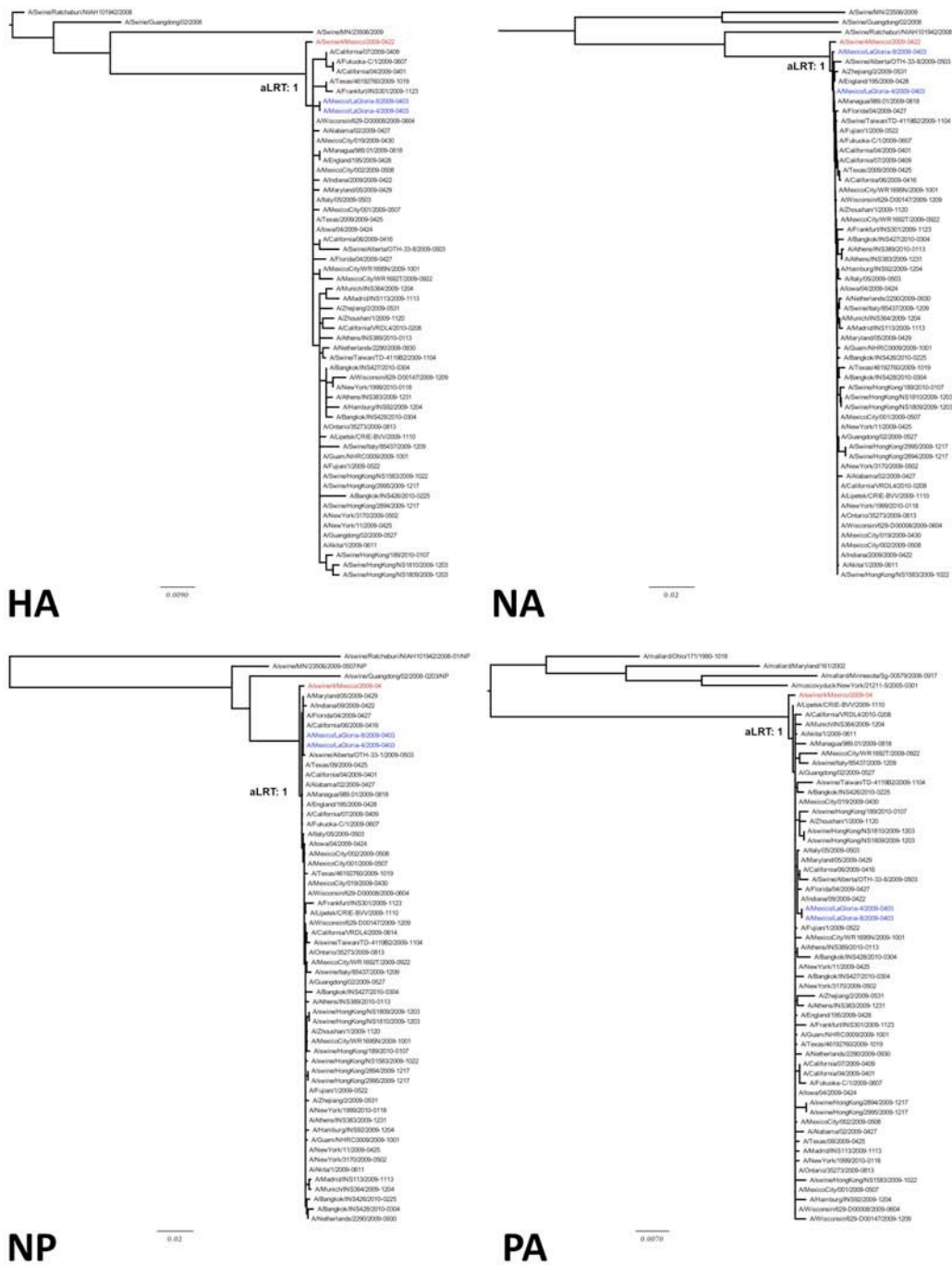
**Figura 1. Clusters y Sub-Clados de los Virus A/H1N1/2009 Pandémicos**

Árbol filogenético obtenido por análisis Bayesiano (MCC) de los ocho genes virales concatenados, donde se observa la agrupación de distintos virus A/H1N1/2009 pandémicos en siete diferentes sub-clados. Los sub-clados dentro del Cluster 1: sub-clado 1-Azul Claro, sub-clado 2-rojo, sub-clado 3-amarillo, sub-clado 4-verde. Los sub-clados dentro del Cluster 2 o New York: sub-clado 5-morado, sub-clado 6-naranja, sub-clado 7-café. El Cluster 1 está marcado en rojo, mientras que el Cluster 2 en negro. (Modificado de Nelson et al, 2009).



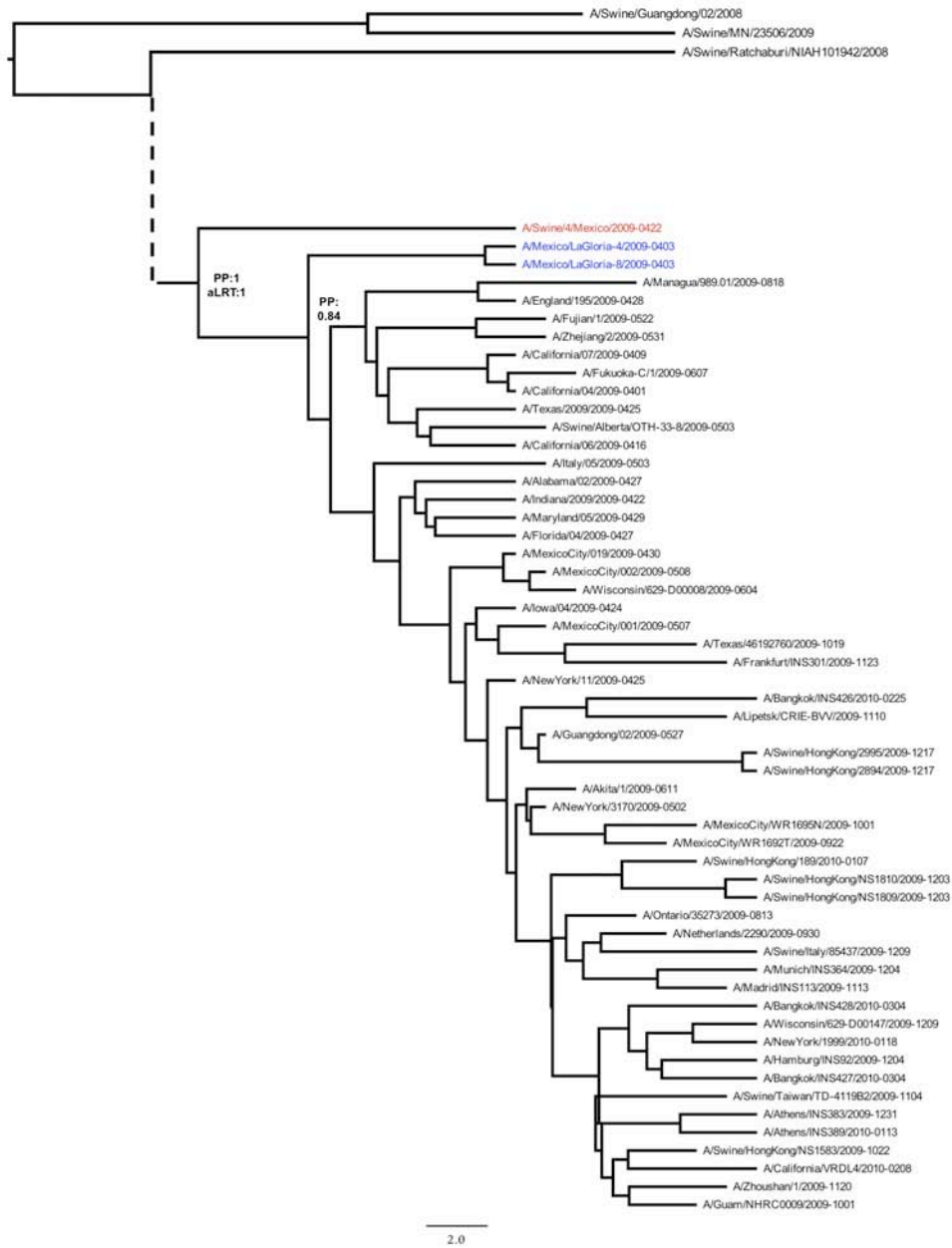
**Figura 2. Los Virus A/H1N1/2009 Pandémicos Secuenciados agrupan en distintos Sub-Clados.**

Árbol filogenético de 58 secuencias de HA y NA concatenadas de virus pandémicos A/H1N1/2009 del mundo, incluidas las muestras secuenciados en este trabajo. El árbol se obtuvo bajo Máxima Verosimilitud (ML, GRT+I+G). Se observa (con el mismo código de color del árbol de la Figura 7), la distribución de los virus en los siete diferentes sub-clados.



**Figura 3. A/swine/4/Mexico/09 es Basal a Otros Virus Pandémicos en el Análisis Filogenético de los Genes HA, NA, NP y PA .**

Árboles filogenéticos de cincuenta y cinco secuencias virales del linaje pandémico 2009-2010, incluidos el aislado porcino A/swine/4/09 y los virus de La Gloria. Los árboles están enraizados con virus H1N1 porcinos triple rearreglantes (HA, NA, NP) o con virus de origen aviar (PA). Cada árbol se obtuvo bajo Máxima Verosimilitud (ML, HKY85 + G, aLRT como valor de soporte para nodos divergentes). Los valores de soporte para los nodos divergentes entre la cepa A/swine/4/Mexico/09 y el resto de los virus pandémicos son los siguientes: HA, aLRT = 1,0; NA aLRT = 1,0; aLRT NP = 1.0; aLRT PA = 1.0.

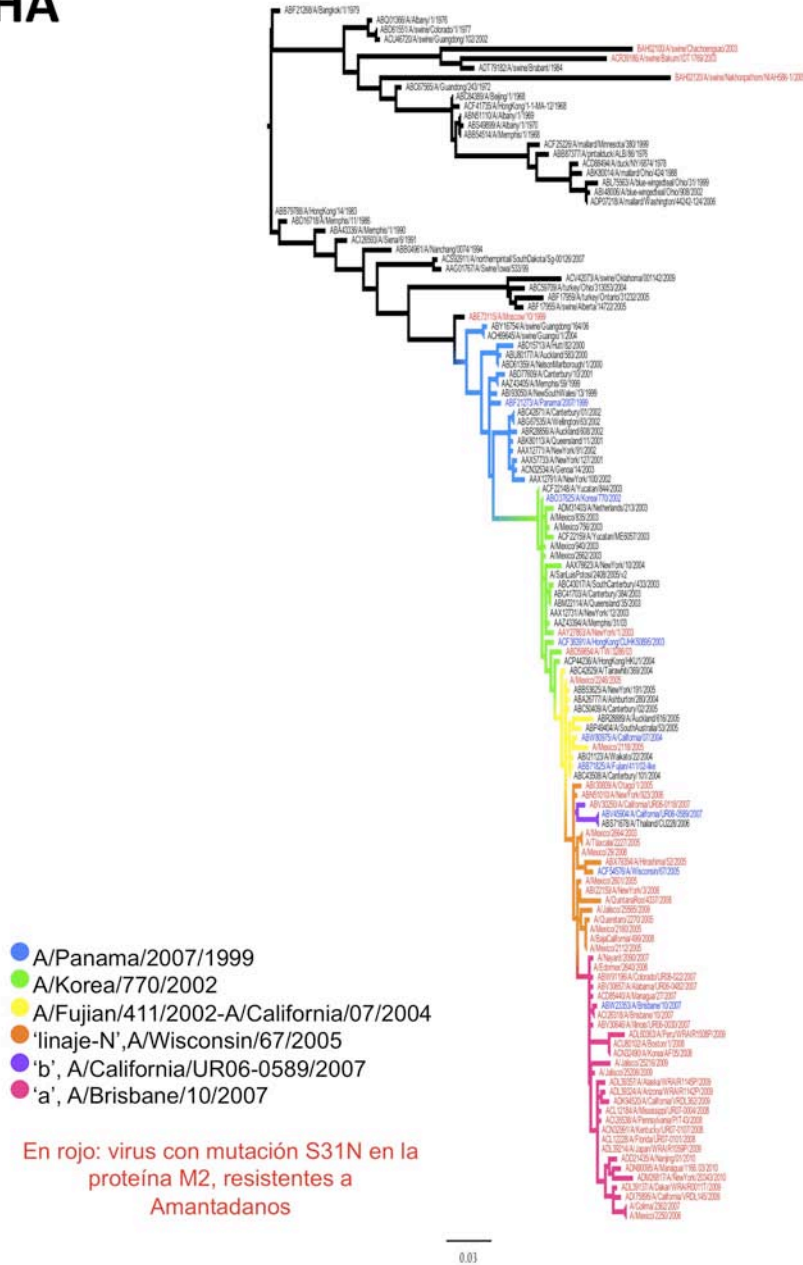


## HA-NA

**Figura 4. A/swine/4/Mexico/09 es basal a otros Virus Pandémicos en el Análisis Filogenético De Los Genes HA y NA Concatenados**

Árbol filogenético de máxima credibilidad (MCC) de 55 secuencias virales pandémicas de 2009-2010 (BEAST, modelo SRD06 bajo UCLD). Los virus de La Gloria y el aislado porcino A/swine/4/09 se incluyeron en el análisis. El árbol está enraizado con virus H1N1 triple rearreglantes de origen porcino (clado marcado con línea punteada). Una topología alterna fue obtenida bajo el criterio de Máxima Verosimilitud (ML, HKY85 + G, aLRT como el soporte). El valor de soporte para los nodos claves está incluido en probabilidad posterior (PP:1) y en aLRT:1.

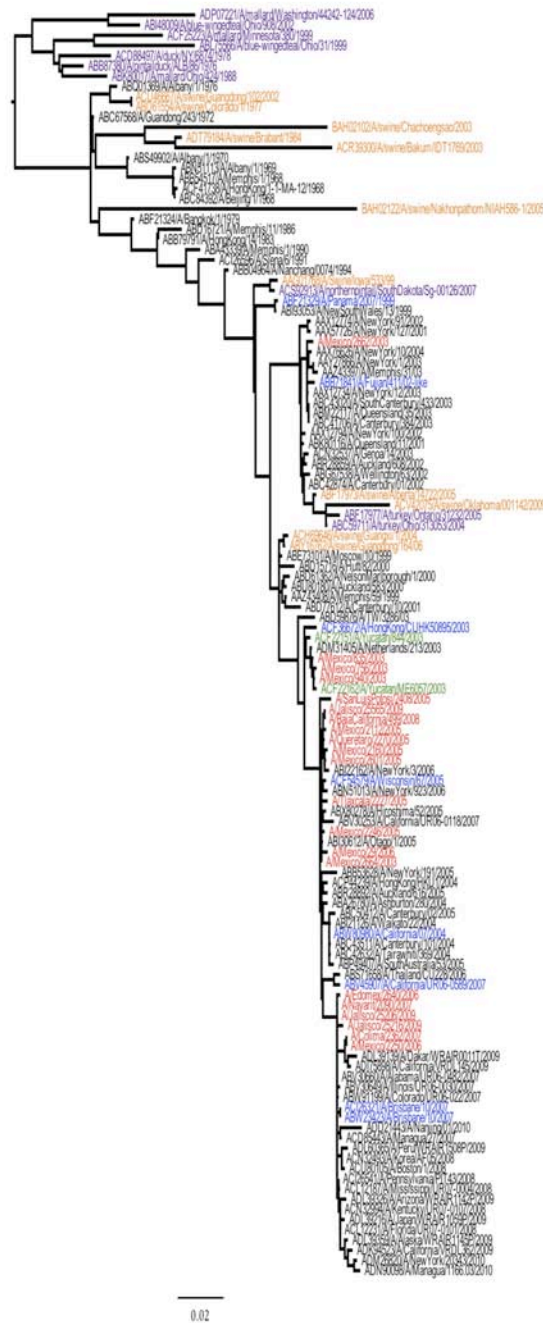
# HA



**Figura 5. Evolución de los Virus H3N2 basado en el Análisis de HA.**

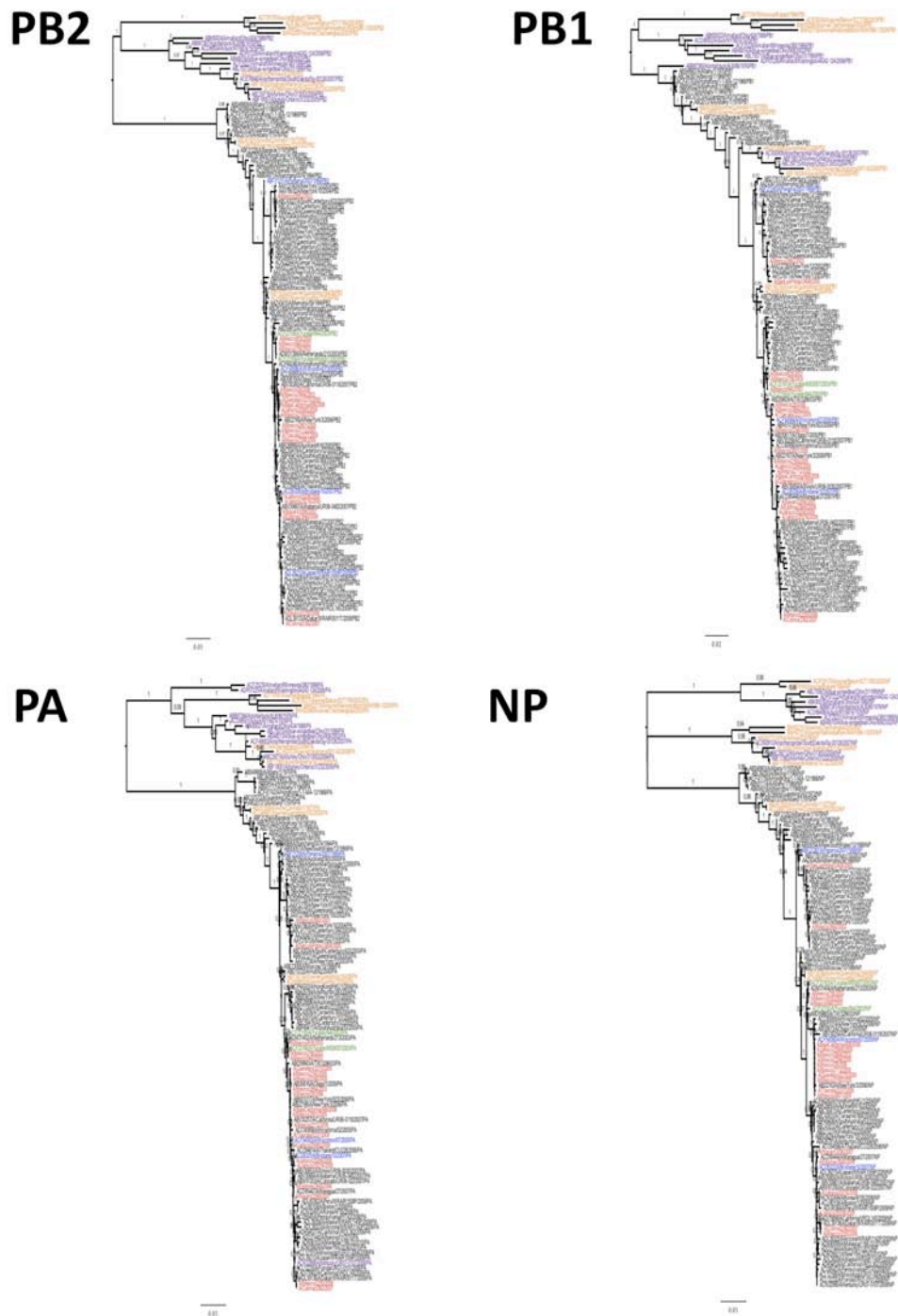
Árbol filogenético de 125 secuencias de virus tipo H3N2 construido bajo el criterio de Máxima Verosimilitud (ML, GTR+I+G); las 21 cepas mexicanas secuenciadas en este trabajo fueron incluidas. Cada clado está representado con un color distinto, y las cepas de referencia que definen a cada clado están marcadas en azul. Las secuencias en negro corresponden a virus sensibles a amantados (S en la posición 31 de M2), mientras que las secuencias en rojo corresponden a virus resistentes a amantados (N31 en M2). El árbol está enraizado por punto medio para mayor claridad.

NA



**Figura 6. Evolución de los Virus H3N2 basado en el Análisis de NA.**

Árbol filogenético de 125 secuencias de virus tipo H3N2 construido bajo el criterio de Máxima Verosimilitud (ML, GTR+I+G); las 21 cepas mexicanas secuenciadas en este trabajo fueron incluidas. Las cepas de referencia que definen a cada clado están marcadas en azul, las de origen aviar en morado y de origen porcino en naranja. Las secuencias en rojo corresponden las cepas mexicanas de este trabajo, mientras que cepas en verde corresponden a virus de referencia colectados en México en años previos (A/Yucatán/ME6057/2003, A/Yucatán/844/2003). El árbol está enraizado por punto medio para mayor claridad.

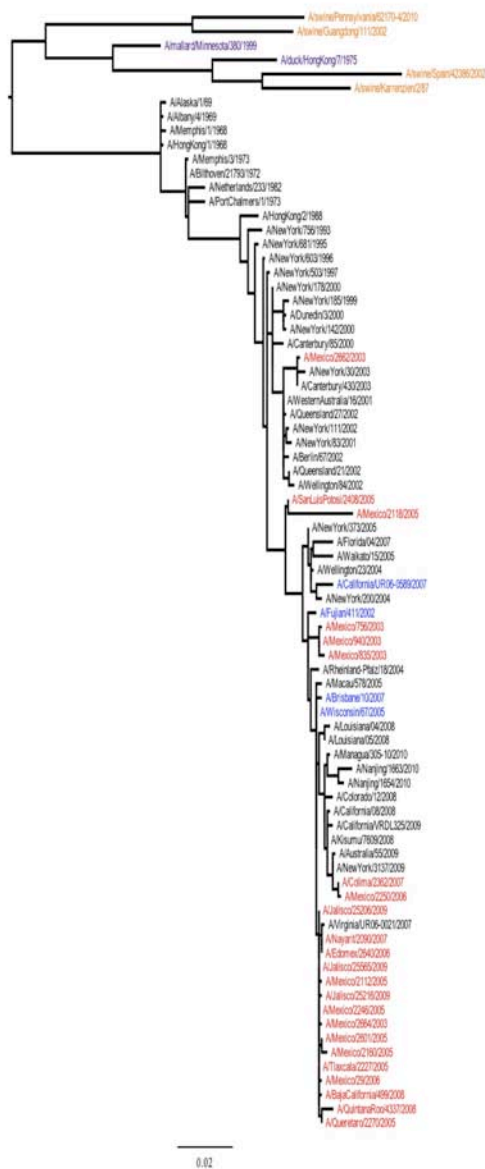


**Figura 7. Evolución de los Virus H3N2 basado en el Análisis de PB2, PB1, PA y NP**

Árbol filogenético de 125 secuencias de virus tipo H3N2 correspondientes a los genes PB2, PB1, PA y NP construido bajo el criterio de Máxima Verosimilitud (ML, GTR+I+G). Las 21 cepas mexicanas secuenciadas en este trabajo fueron incluidas. Las cepas de referencia que definen a cada clado están marcadas en azul, las de origen aviar en morado y de origen porcino en naranja. Las secuencias en rojo corresponden las cepas mexicanas secuenciadas en este trabajo, mientras que cepas en verde corresponden a virus de referencia colectados en México en años previos (A/Yucatán/ME6057/2003, A/Yucatán/844/2003). El árbol está enraizado por punto medio para mayor claridad.



# M1/M2



# NS1/NS2



**Figura 8. Evolución de los Virus H3N2 Basado en el Análisis de M1/M2 y NS1/NS2.**

Árbol filogenético de 80 secuencias (gen M1/M2 ) y 125 secuencias (gen NS1/NS2) de virus tipo H3N2, construido bajo el criterio de Máxima Verosimilitud (ML, GTR+I+G); las 21 cepas mexicanas secuenciadas en este trabajo fueron incluidas. Las cepas de referencia que definen a cada clado están marcadas en azul, las de origen aviar en morado y de origen porcino en naranja. Las secuencias en rojo corresponden las cepas mexicanas secuenciadas en este trabajo, mientras que cepas en verde corresponden a virus de referencia colectados en México en años previos (A/Yucatán/ME6057/2003, A/Yucatán/844/2003). El árbol está enraizado por punto medio para mayor claridad.