



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

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

**INSTITUTO DE BIOTECNOLOGÍA**

**DESARROLLO DE UN MICROARREGLO PARA LA  
DETERMINACIÓN DEL ORIGEN FILOGENÉTICO DE LOS  
SEGMENTOS DEL GENOMA DEL VIRUS DE INFLUENZA A**

**T E S I S**

QUE PARA OPTAR POR EL GRADO DE

**MAESTRO EN CIENCIAS**

P R E S E N T A

LCG. LUIS FELIPE PAULÍN PAZ

DIRECTOR DE TESIS:

Dr. CARLOS FEDERICO ARIAS ORTIZ  
INSTITUTO DE BIOTECNOLOGÍA

INTEGRANTES DE COMITÉ TUTOR:

Dr. LORENZO SEGOVIA, INSTITUTO DE BIOTECNOLOGÍA  
Dra. SUSANA CASTRO OBREGÓN, INSTITUTO DE  
BIOTECNOLOGÍA

**CUERNAVACA, MORELOS**

**MARZO 2013**

El presente trabajo de tesis fue realizado en el Departamento de Genética del Desarrollo y Fisiología Molecular del Instituto de Biotecnología de la Universidad Nacional Autónoma de México, campus Morelos, bajo la dirección del Dr. Carlos Federico Arias Ortiz, en el laboratorio de los doctores Carlos Arias y Susana López. Se contó con el apoyo DGAPA PAPIIT, UNAM IN213410.

Durante la realización del mismo el alumno contó con el apoyo económico de CONACyT

Comité tutorial:

**Dr. Carlos Arias Ortiz**

**Dr. Lorenzo Segovia**

**Dra. Susana Castro**

Jurado de examen:

**Dr. Juan Enrique Morett Sánchez**

**Dr. Mario Enrique Zurita Ortega**

**Dra. Rosa María Gutiérrez Ríos**

**Dra. Laura Alicia Palomares Aguilera**

**Dr. Pablo Vinuesa Fleischmann**

## **DEDICATORIAS**

Les dedico esta tesis a mi familia y amigos que siempre me apoyaron en cada momento.

A mi madre y a mi padre por el apoyo incondicional en cada uno de los días desde el inicio de mis estudios de licenciatura y que siguen apoyándome en todo momento.

A mi hermano Chucho por soportarme todos esos días de cuando llegaba de malas y agobiado y por esos ratos de relajación en las retas de Street Fighter.

A mis amigos por todos los momentos que pasamos en clases y el laboratorio y claro en los momentos de recreación, díganse fiestas y convivios.

A mi novia Benelli por levantarme el ánimo desde el primer día que nos reencontramos, y que sigue haciéndolo cada que me ve estresado.

Al Dr. Carlos Arias por darme la oportunidad de trabajar bajo su tutela, no importando el poco tiempo que tenía disponible, siempre tenía un momento para mí.

A mis maestros, asesores, jefes, tutores, sinodales y compañeros de laboratorio, alumnos y doctores, por sus valiosos comentarios que hicieron este trabajo un éxito.

## **AGRADECIMIENTOS**

Al Dr. Carlos Arias y a la Dra. Susana López por permitirme trabajar en su laboratorio, por los comentarios y críticas que ayudaron a la realización del proyecto.

Al Dr. Pavel Isa por su tiempo y ayuda en el laboratorio y estar pendiente de cómo avanzaba el proyecto

A María de los Dolores Soto del Río por haberme ayudado en la estructuración del proyecto con sus comentarios, y por enseñarme como trabajar en el laboratorio.

A todos los integrantes del grupo Arias/López. A mis compañeros de mesa: unos se iban y otros llegaban, por los momentos en, claro está la mejor mesa del laboratorio.

A los integrantes de mi Comité Tutorial, Lorenzo, Susana, gracias por sus críticas, consejos a lo largo de estos más de dos años.

A Jerome por su ayuda y disponibilidad en el manejo del cluster.

A la Unidad de Docencia por su disposición en todo momento.

A los miembros del jurado de examen por sus valiosas observaciones.

# ÍNDICE GENERAL

I. Introducción	1
1. Virus de Influenza A	1
1.1. Estructura	1
1.2 Evolución	3
1.3 Influenza y pandemias	5
1.4 Subtipos y hospederos	7
1.5 Genes y función	8
1.6 Ciclo de Replicación	11
2. Microarreglos	13
2.1 Generalidades	13
2.2 Trabajos previos	13
II. Antecedentes	17
1. Metodologías de detección y vigilancia epidemiológica en México	17
2. Microarreglo de detección de virus asociados a enfermedades respiratorias	18
III. Hipótesis	19
IV. Objetivos	19
V. Materiales y Métodos	20
1. Obtención de las secuencias	20
2. Generación de las bases de datos	20
3. Reducción de complejidad en el análisis	21
4. Análisis filogenéticos	21
5. Definición de clados	21
6. Generación de oligonucleótidos de longitud 70	22
7. Selección de oligonucleótidos	23
8. Ensayos in silico	25
9. Conversión de clado a origen	26
10. Preparación de oligonucleótidos	26
11. Impresión de microarreglos	27
12. Post-procesamiento de los microarreglos	27
13. Cepas de influenza A utilizadas para validar el microarreglo	28
14. Extracción de RNA y amplificación del material genético presente en las muestras	31
15. Purificación y marcaje del DNA	35
16. Hibridación	36
17. Lavados y lectura de los microarreglos	36
18. Análisis de datos	37
VI. Resultados	39
1. Desarrollo del microarreglo	39
1.1 Obtención de las secuencias	39

1.2 Reducción de complejidad en el análisis	40
1.3 Generación de oligonucleótidos de longitud 70	40
1.4 Definición de Clados	42
1.5 Selección de oligonucleótidos	44
1.6 Nomenclatura de los oligonucleótidos seleccionados	45
1.7 Ensayos in silico	46
1.7.1 Ensayos in silico con secuencias de referencia del virus de influenza A pertenecientes a distintos subtipos	46
1.7.2 Ensayos in silico con genomas de influenza A/H1N1	46
1.7.3 Ensayos in silico con genomas de influenza A/H3N2	48
1.7.4 Ensayos in silico con genomas completos de influenza A aviar, porcina y humana	48
2. Estadística y análisis de resultados	49
3. Ensayos de Validación	50
3.1 Hibridación con cepas de referencia	50
3.2 Hibridación de cepas de origen porcino	52
3.3 Hibridación de una cepa influenza A pandémica 2009	53
3.4 Hibridación con cepas de origen aviar	54
3.5 Hibridación con cepas de origen humano	56
4. Pruebas de sensibilidad del microarreglos	58
5. Asignación de un origen animal a cada clado	59
VII. Discusión	62
1. Definición de clados	62
2. Ensayos in silico	63
3. Validación experimental del microarreglo	64
3.1 Cepas de referencia	64
3.2 Muestras clínicas y de campo	65
3.2.1 Muestras de origen porcino	65
3.2.2 Muestras de origen humano	66
3.2.3 Muestras de origen aviar	66
3.2.4 Comparación con el método de secuenciación masiva	67
4. Asignación de origen a los clados	67
VIII. Conclusión y Perspectivas	69
IX. Bibliografía	71
X. Anexos	75

## ÍNDICE DE TABLAS

Tabla 1	29
Tabla 2	30
Tabla 3	30
Tabla 4	30
Tabla 5	39
Tabla 6	39
Tabla 7	40
Tabla 8	41
Tabla 9	44
Tabla 10	47
Tabla 11	49
Tabla 12	50
Tabla 13	52
Tabla 14	53
Tabla 15	54
Tabla 16	55
Tabla 17	57
Tabla 18	58
Tabla 19	58
Tabla anexa 1	75
Tabla anexa 2	80
Tabla anexa 3	83
Tabla anexa 4a	85
Tabla anexa 4b	87
Tabla anexa 4c	90
Tabla anexa 5	92
Tabla anexa 6	93
Tabla anexa 7	94
Tabla anexa 8	100
Tabla anexa 9	100

## ÍNDICE DE FIGURAS

Figura 1A	2
Figura 1B	2
Figura 1C	3
Figura 2	4
Figura 3	7
Figura 4	8
Figura 5	12
Figura 6	14
Figura 7	22
Figura 8	24
Figura 9	40
Figura 10	43

## RESUMEN

Se desarrolló una metodología para la detección y clasificación del virus de influenza A basado en microarreglos de DNA. El diseño actual identifica y clasifica los ocho segmentos de los que se compone el genoma de influenza A, y para el caso de los genes de Hemaglutinina (HA) y Neuraminidasa (NA) para cada subtipo, en distintos clados basados en análisis filogenéticos. El microarreglo fue probado y validado con un conjunto de cepas de referencia las cuales representan a los 16 subtipos de HA y los 9 de NA, además que comprende a los hospederos principales del virus: humanos, cerdos y aves. Adicionalmente se realizaron hibridaciones con muestras clínicas de provenientes de humanos, cerdos y aves los cuales arrojaron resultados muy favorables. Para el caso de las cepas humanas fue posible identificar las diferencias entre la cepa de influenza A/H1N1 estacional 2008 y la cepa de influenza A/H1N1 pandémica 2009. En las muestras porcinas se identificaron dos virus sin reporte previo en México, los cuales son triple rearreglantes porcinos A/H1N1 y A/H3N2. En el caso de las muestras de origen aviar nuestro diseño logró identificar diversos subtipos, de los cuales no existen reportes previos de los virus que circulan en México. Con nuestro diseño podemos identificar si una cepa ha sido previamente reportada. A su vez es posible identificar nuevas combinaciones de segmentos producto de eventos de rearreglo génico, hasta un total de 1 260 000 posibles combinaciones. Este microarreglo puede ser utilizado en la detección y clasificación de influenza A en estudios clínicos, así como en ecología de influenza A en aves silvestres.

## **ABSTRACT**

We developed a methodology for the detection and classification of the influenza A virus based on DNA microarrays. The current design identifies and classifies each of the eight segments of the influenza A genome, and for the case of Hemagglutinin (HA) and Neuraminidase (NA) for each subtype, into different clades based on a phylogenetic analysis. The microarray was tested with different reference strains that represent the 16 subtypes of HA and the 9 subtypes of NA, of the three principal influenza A hosts: human, swine and birds. Additionally human, swine and avian clinic and field samples were hybridized giving very promising results. In the case of human samples we could differentiate between the pre pandemic and the pandemic A/H1N1 flu strains. In swine we identify triple reassortant A/H1N1 and A/H3N2 not previously reported in Mexican swine population. In wild birds there was no previous report of what is circulating, and we could identify several subtypes. With our design we can identify if a strain was previously reported, and in which year and geographic region was first and last reported. It can also identify new segment combination, due to reassortment between 1 260 000 possible combinations. This microarray could be used in influenza A clinic studies, and in studies of the influenza A ecology in wild birds.

# I. Introducción

El virus de influenza pertenece a la familia *Orthomyxoviridae* y se clasifica en influenza A, B y C (Lamb & Krug, 2007). El virus de influenza A infecta naturalmente a una gran variedad de especies de aves silvestres y domésticas. Este virus también infecta a un gran número de mamíferos, incluyendo al ser humano. El virus de influenza B infecta exclusivamente al ser humano, y el virus influenza C infecta al hombre y algunos otros animales.

## 1. Virus de Influenza A

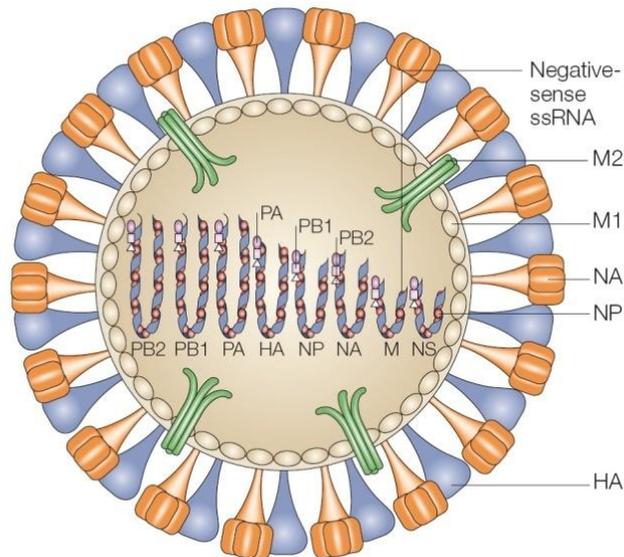
### 1.1. Estructura

El virus de influenza A es un virus de genoma segmentado de RNA de cadena sencilla con polaridad negativa. Su genoma consta de 13.5 kilobases, el cual está dividido en 8 segmentos que codifican para 11 proteínas: HA, NA, NP, NS1, NS2, M1, M2, PA, PB2, PB1 y PB1-F2 (Figura 1A) (Bouvier & Palese, 2008). Los 8 segmentos que conforman el genoma viral tienen regiones 5' y 3' no traducidas (5' UTR y 3' UTR) altamente conservadas (Zhou *et al.* 2009). Para el caso de la 5' UTR son 13 nucleótidos conservados y para la 3' UTR 12 nucleótidos (Figura 1B).

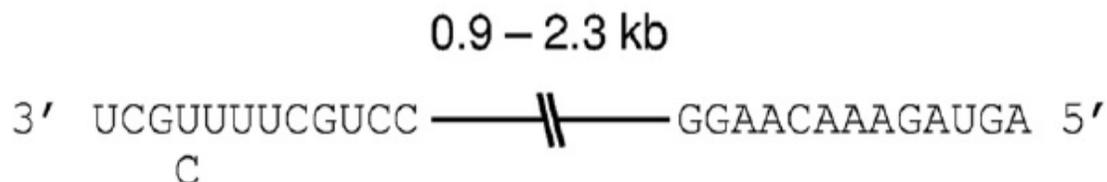
Los segmentos 1, 3, 4, 5 y 6 codifican sólo para una proteína. Los segmentos 7 y 8 codifican para 2 productos, uno colinear y otro que es formado mediante un proceso de splicing, mientras que el segmento 2 codifica para 2 productos que se traducen a partir de dos marcos de lectura diferentes (Figura 1C).

Las partículas virales tienen una envoltura lipídica la cual proviene de la membrana plasmática de la célula huésped. En ésta se encuentran insertadas aproximadamente 500 espículas formadas por los principales antígenos del virus, la hemaglutinina (HA) y la neuraminidasa (NA), los cuales se proyectan hacia el exterior. También se encuentra insertada en la envoltura lipídica, aunque en menor proporción, la proteína M2, la cual forma un poro en la membrana. Bajo la envoltura lipídica se encuentra la proteína de matriz (M1), la cual encierra a las ribonucleoproteínas (RNP). Estos complejos están

formados por las proteínas de la nucleocápside (NP) y el complejo de RNA polimerasa dependiente de RNA constituido por la polimerasa ácida (PA), la polimerasa básica 1 (PB1) y la polimerasa básica 2 (PB2) asociada a los 8 segmentos del genoma del virus (Figura 1A) (Inglis *et al.*, 1976; Lamb & Choppin, 1976).



**Figura 1A.** Diagrama esquemático del virus de Influenza A. Las dos glicoproteínas de superficie, hemaglutinina (HA) y neuraminidasa (NA), y el canal iónico M2 se encuentran embebidos en la membrana viral, la cual proviene de la membrana plasmática de la célula huésped. El complejo de ribonucleoproteínas se compone de los segmentos de RNA viral asociados con la nucleoproteína (NP) y las tres proteínas de la polimerasa (PB2, PB1 y PA). La proteína de la matriz (M1) está asociada tanto a las ribonucleoproteínas como a la membrana viral. (Figura 1 de Horimoto & Kawaoka. Influenza: lessons from past pandemics, warnings from current incidents. *Nature Rev. Microbiol.* 3, 591–600, 2005).



**Figura 1B.** Nucleótidos conservados en la 5' y 3' UTR del genoma de influenza A. Estos nucleótidos son conservados en los 8 segmentos (Figura 1 de Zhou *et al.*, *Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses.* *Journal of Virology*, Oct. 2009).

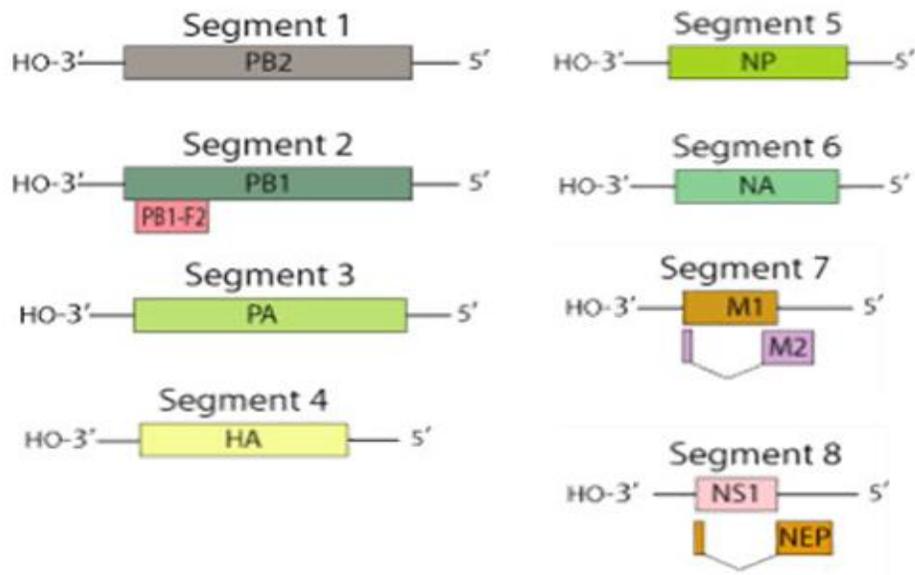


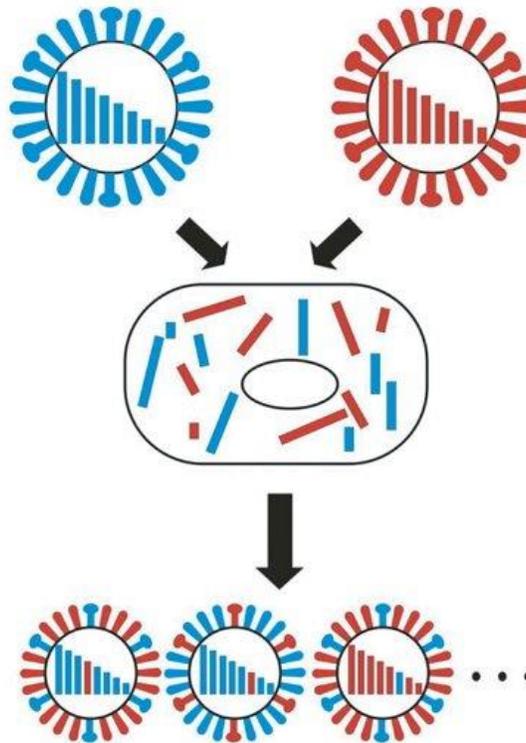
Figura 1C. 8 segmentos que conforman el genoma del virus de influenza A, los cuales codifican para 11 proteínas. Los segmentos varían en tamaño desde 2341 a 890 nucleótidos, dando un total de 13.5Kb. ([viralzone.expasy.org/all\\_by\\_species/130.html](http://viralzone.expasy.org/all_by_species/130.html)).

## 1.2 Evolución

Dos son los procesos principales que dirigen la evolución y adaptación del virus de influenza A. Estos son la deriva génica y el intercambio o rearreglo génico.

La deriva génica es un proceso por el cual las secuencias de nucleótidos cambian gradualmente debido a mutaciones aleatorias originadas por la RNA polimerasa (Ohta, 2005). Estas mutaciones son seleccionadas rápidamente a favor o en contra dependiendo de la presión selectiva ejercida por el hospedero. La acumulación de estas mutaciones puede conducir a cambios de aminoácidos que provoquen, por ejemplo, una menor afinidad de los anticuerpos por las proteínas de superficie del virus. Esto resulta en que las cepas de virus originadas por la acumulación de mutaciones sean neutralizadas sólo parcialmente por los anticuerpos inducidos en infecciones previas. La baja fidelidad de la RNA polimerasa del virus de influenza A y la falta de actividad de corrección de la misma llevan a una tasa de mutación aproximada de 1 nucleótido cambiado por cada evento de

replicación de genoma del virus, es decir  $7.3 \times 10^{-5}$  mutaciones por nucleótido replicado (Drake, 1993).



**Figura 2. Rearreglo génico en el virus de Influenza A. Los colores del virus representan de que virus parental proviene cada segmento. Para los casos de los segmentos 4 y 6, en forma descendente, cambia el color de las espículas que representan HA y NA. (<http://emssolutionsinc.wordpress.com/2010/03/08/h1n1-swine-flu-human-infection-with-a-pandemic-h1n1-and-seasonal-h1n1-flu-reported-from-canada-in-three-workers/>)**

El intercambio o rearreglo de genes se debe a la naturaleza segmentada del genoma viral. Al haber una coinfección con dos o más cepas de influenza los segmentos de estos virus pueden reorganizarse aleatoriamente, generando virus con nuevas combinaciones de genes (Figura 2). Este fue el caso de las pandemias de 1957, 1968 y 2009, las cuales fueron ocasionadas por virus rearreglantes. Cuando existe algún rearreglo en los genes que codifican para las proteínas de superficie HA y NA el virus tiene la capacidad de modificar su rango de hospedero, y puede ocurrir un evento de zoonosis en donde un virus que infectaba por ejemplo a un ave, adquiere la capacidad de infectar humanos. Al igual que el caso anterior, un cambio en las proteínas de superficie conlleva a la ausencia de

protección inmunológica de la población. La combinación de ambas situaciones puede causar un grave problema de salud pública (Taubenberger & Kash, 2010).

### **1.3 Influenza y pandemias**

Se estima que anualmente se infecta aproximadamente el 15% de la población mundial con los virus de influenza estacional, que comprende influenza A/H1N1 A/H3N2 e influenza B, causando de 3 a 5 millones de casos de enfermedad respiratoria severa y alrededor de 250 000 a 500 000 muertes. Estas cifras aumentan considerablemente cuando surgen cepas pandémicas, en donde se puede infectar hasta el 30% de la población mundial (WHO, 2009). De enero de 2012 a enero de 2013 se han reportado 250 183 casos confirmados de influenza A y B (WHO, 2013). Hasta antes de la pandemia de 2009, los virus estacionales circulantes eran A/H1N1, A/H3N2 e influenza B. En el 2010 y 2011 han continuado circulado los virus A/H3N2 e influenza B, sin embargo, el virus pandémico de origen porcino A/H1N1 pdm2009 ha reemplazado al virus estacional humano A/H1N1.

Durante el siglo pasado se presentaron tres grandes pandemias ocasionadas por el virus de influenza A, la gripe española en 1918 (influenza A/H1N1), la gripe asiática en 1957 (influenza A/H2N2) y la gripe de Hong Kong en 1968 (influenza A/H3N2). En 2009 se suscitó la primera pandemia de siglo XXI, la cual fue causada por un virus rearrreglante de influenza A/H1N1 de origen porcino. Las tres últimas pandemias han sido causadas por virus rearrreglantes.

La primera, en 1918, se originó aparentemente por la transmisión a humanos de una cepa aviar de influenza A/H1N1 (Taubenberger *et al*, 2005). Durante esta pandemia se infectó aproximadamente el 30% de la población mundial y está registrada como la más grave, habiendo ocasionado más de 50 millones de muertes en 1918 y 1919 (Taubenberger & Morens, 2006).

En 1957 surgió una cepa de influenza A/H2N2 por un proceso de rearreglo de un virus de influenza A/H1N1 descendiente del pandémico de 1918 con un virus de influenza aviar A/H2N2. En el rearreglo el virus H1N1 adquirió 3 segmentos del virus H2N2, que

codificaban para la hemaglutinina, la neuraminidasa y la polimerasa básica 1 (Figura 3). Al obtener los genes de HA y NA de un subtipo diferente se perdió la inmunidad de la población inducida por infecciones previas con influenza H1N1. Esta pandemia no fue tan grave. Al igual que el virus de la gripe española de 1918, después de pasados algunos años este virus se volvió una cepa estacional.

Al igual que en la pandemia de 1957, en 1968 surgió un virus de influenza A rearreglante. La nueva cepa H3N2 fue producto de un rearreglo del virus H2N2 de 1957 con un virus aviar subtipo H3 (Figura 3). En este proceso el virus adquirió los genes HA y PB1. El virus se diseminó rápidamente a nivel mundial, aunque su letalidad fue baja. Esta cepa se convirtió en una cepa estacional después de unos años, la cual sigue circulando hasta nuestros días.

En 1977 resurgió la cepa H1N1 después de no haber circulado por cerca de 20 años, aparentemente por la liberación accidental de un laboratorio. Esta cepa se convirtió en estacional, la cual circuló junto con el virus A/H3N2 (Taubenberger & Kash, 2010).

En 2009 surgió una cepa nueva de influenza A/H1N1, la cual fue originada por un rearreglo de dos cepas porcinas no relacionadas, una cepa H1N1 porcina euroasiática similar a cepas aviares y una cepa porcina triple rearreglante. Análisis de los segmentos del genoma del virus pandémico A/H1N1 2009 demuestran que éstos provenían originalmente de 5 distintos tipos de virus: influenza porcina clásica de Norteamérica, influenza aviar de Norteamérica, influenza humana, y dos genes de virus porcinos euroasiáticos (Figura 4) (Arias *et al.*, 2009; Smith *et al.*, 2009; Taubenberger & Kash, 2010).

Es importante mencionar que después de cada una de las pandemias la cepa estacional que circulaba previamente ha dejado de hacerlo y ha sido remplazada por la cepa pandémica, la cual con el paso del tiempo se convierte en estacional (Figura 3). Esto mismo sucedió parcialmente en 2009, ya que la cepa A/H1N1 pdm remplazó a la cepa A/H1N1 estacional que circulaba hasta ese año, sin embargo no desplazó al virus A/H3N2 el cual sigue circulando.

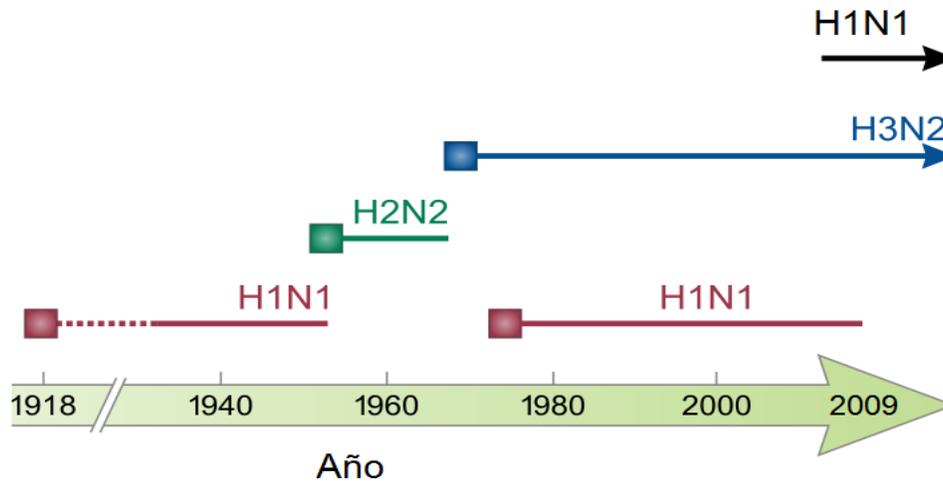


Figura 3. Principales cepas circulantes del virus de Influenza A en humano desde 1918. Los cuadros representan la aparición de una nueva cepa, la cual causo una pandemia. La línea punteada indica incertidumbre en la identificación de dicha cepa. (Modificado de [http://en.wikipedia.org/wiki/File:Influenza\\_subtypes.svg](http://en.wikipedia.org/wiki/File:Influenza_subtypes.svg)).

#### 1.4 Subtipos y hospederos

Como se menciona anteriormente, el virus de influenza A se clasifica en subtipos con base en la antigenicidad de las proteínas de superficie del virus HA y NA, de los cuales existen 16 subtipos para la HA y 9 para la NA reportados hasta Marzo de 2012 cuando describió un nuevo subtipo de HA (H17) y un nuevo subtipo de NA (N10) (Tong et al, 2012). Aunque el virus tiene la capacidad de infectar tanto a mamíferos como a aves, destacan las aves acuáticas silvestres, las que pueden ser infectadas por todos los subtipos tanto de HA como de NA, y en la gran mayoría de los casos de manera asintomática, por lo que se les considera como el reservorio natural del virus.

La especificidad para que un virus infecte a aves o a mamíferos está dada por el reconocimiento de la HA hacia las moléculas de ácido N-acetilneuramínico (ácido siálico) de la superficie de las células. El ácido siálico se encuentra al final de cadenas de glicoconjugados, usualmente unido a una galactosa en enlaces  $\alpha 2,3$  ó  $\alpha 2,6$ . El tipo de unión de la galactosa con el ácido siálico es el que marca la diferencia de reconocimiento por la HA. Los virus que infectan a humano tienen preferencia por ácidos siálicos con unión  $\alpha 2,6$ , los cuales se encuentran en las células epiteliales del tracto respiratorio

superior, mientras que los virus que infectan a aves tienen preferencia por el enlace  $\alpha 2,3$ , el cual se encuentra en el epitelio intestinal de las aves. Es importante mencionar que el cerdo tiene ácido siálico con ambos enlaces en el tracto respiratorio, por lo que éste puede ser infectado tanto por virus de origen humano como de origen aviar y por supuesto de origen porcino, por lo que puede funcionar, al ser infectado por virus diferentes, como un generador de virus de influenza rearreglantes (Arias *et al.*, 2009).

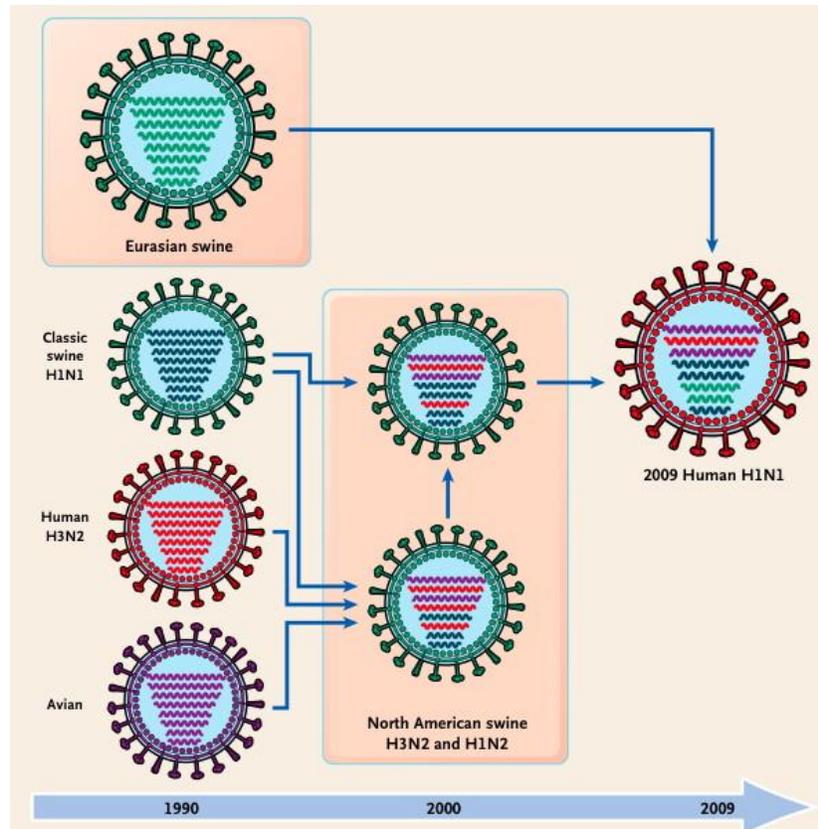


Figura 4. Origen del virus de influenza A/H1N1 pandémico. El virus A/H1N1 pandémico se originó por un evento de rearreglo génico entre un virus porcino A/H1N1 / A/H1N2 triple rearreglante y un virus porcino euroasiático A/HAN1 (subtipo de HA desconocido). Cada color de segmento representa un origen distinto. En la parte inferior se muestra el tiempo aproximado de aparición de cada cepa. (<http://www.virology.ws/2009/06/29/reassortment-of-the-influenza-virus-genome/>)

### 1.5 Genes y función

#### POLIMERASA VIRAL

Los tres segmentos más grandes del genoma del virus de influenza codifican para subunidades que forman el complejo de la polimerasa viral, la cual es una RNA polimerasa

dependiente de RNA. Las proteínas fueron nombradas por su comportamiento en geles de enfoque isoeléctrico, dos de ellas proteínas básicas (PB1 y PB2) y la tercera ácida (PA). El complejo de la polimerasa viral es el encargado de sintetizar los RNA mensajero, el RNA anti-genómico y el genoma viral. Esta polimerasa carece de actividad de corrección cuando incorpora un nucleótido equivocado, lo que genera altas tasas de mutación.

Las polimerasas virales están asociadas a la virulencia. Se han reportadas mutaciones en PB1 y PA que aumentan la actividad de la polimerasa (Arias *et al.*, 2009). PB1-F2 es un marco de lectura alternativo de PB1, el cual está caracterizado como un factor pro-apoptótico. PB1-F2 interactúa con proteínas mitocondriales y causa la liberación del citocromo c al citoplasma, lo cual es una señal de muerte celular. PB2 se ha asociado con la capacidad del virus de replicarse mejor a diversas temperaturas. Aquellos que se replican mejor a temperaturas bajas (33°C) tienen una virulencia media en humanos, mientras que aquellos que pueden replicarse mejor a 37°C tienen la capacidad de replicarse más eficientemente en el tracto respiratorio inferior de mamíferos, lo cual se asocia con una mayor virulencia.

#### *PROTEÍNAS NO ESTRUCTURALES*

Las proteínas no estructurales se sintetizan en células infectadas por el virus pero no se incorporan a las partículas virales. El segmento 8 del genoma de influenza A codifica para dos proteínas no estructurales, NS1 y NS2. NS1 está codificada por un transcrito co-lineal, mientras que NS2 es codificada por un mRNA que sufre un proceso de splicing. NS1 tiene diversas funciones dentro de las que destacan la unión al extremo 5' no traducido de los RNA mensajeros virales para favorecer que éstos sean traducidos, inhibir el splicing de pre-mRNAs celulares e inhibir la respuesta celular por interferón. NS1 se expresa abundantemente en células infectadas. Esta proteína interviene en el control de la respuesta inmune celular. NS1 se une a RNA de doble cadena (dsRNA, formado por los intermediarios de replicación del genoma viral) que se encuentra en el núcleo durante la replicación de virus. De esta forma evita la respuesta mediada por interferón, previniendo su activación vía PKR, la cual es una cinasa dependiente de dsRNA (Arias *et al.*, 2009). NS2 se catalogaba inicialmente como no estructural, aunque recientemente se encontró que

se encuentra asociada a M1 en los viriones. NS2 tiene la función de exportar las ribonucleoproteínas (RNPs) del núcleo debido a que tiene una señal de exportación nuclear, por lo que se le conoce también como NEP, del inglés “nuclear export protein”.

#### *PROTEINA DE MATRIZ*

El segmento 7 del genoma de influenza A codifica también para dos proteínas. Este fue el primer ejemplo de un virus de RNA sin intermediario de DNA en su replicación, en el que se describía un proceso de splicing. El primer producto proteico es la proteína M1, un producto co-linear. M1 es la proteína de la matriz, la cual se encuentra dentro de la envoltura lipídica.

El segundo producto proteico, M2, se traduce a partir de un mRNA que sufre splicing. M2 es una proteína integral de la membrana viral, la cual tiene actividad de canal iónico altamente específico para protones. Se han reportado diversas mutaciones que contribuyen a la resistencia a inhibidores del canal, como la drogas de adamantano (Hay AJ, 1992).

#### *NUCLEOCAPSIDE*

La proteína de la nucleocápside, NP, es la proteína estructural que se encuentra en mayor abundancia en los viriones. NP interactúa con los segmentos que componen el genoma viral formando ribonucleoproteínas (RNP). NP es transportada activamente al núcleo por medio de señales de localización nuclear.

#### *HEMAGLUTININA (HA)*

La hemaglutinina es la proteína de superficie más abundante en el virus de influenza A. HA es la responsable del anclaje de virus a la célula vía receptores que contienen ácido siálico. Como se mencionó anteriormente la HA confiere la especificidad de hospedero en función de su afinidad por enlaces del ácido siálico-galactosa  $\alpha 2,6$  y  $\alpha 2,3$ . Dependiendo del tipo de células que infecte, la HA presente en la progenie viral se puede encontrar tanto en la forma no procesada, H<sub>0</sub>, como en la procesada que consiste en dos cadenas, H<sub>1</sub> y H<sub>2</sub>, unidas por enlaces disulfuro. El procesamiento es un corte proteolítico, el cual es un requisito para que el virus sea infeccioso y es un factor determinante en la virulencia. HA

es responsable de la fusión de la membrana viral con la membrana del endosoma, lo que provoca la liberación de la partícula viral al citoplasma durante el proceso de entrada del virus a la célula hospedera, proceso que es dependiente de pH ácido. HA es el principal antígeno contra el cual están dirigidos los anticuerpos neutralizantes.

#### *NEURAMINIDASA (NA)*

La neuraminidasa es la segunda proteína más abundante en la superficie del virus de influenza A y a su vez el segundo antígeno principal. La función de la NA es remover el ácido siálico de las glicoproteínas, lo que promueve la liberación de las partículas durante su proceso de salida de la célula, y previene la agregación de los virus (Lamb & Krug, 2007). Se han reportado mutaciones en esta proteína que contribuyen a la resistencia de drogas antivirales e inhibidores de neuraminidasa.

### ***1.6 Ciclo de Replicación***

El ciclo de replicación del virus de influenza A comienza cuando el virus se une a los residuos de ácido siálico de glicoproteínas de la superficie celular, a través de la HA (Figura 5). Como se mencionó anteriormente, la especificidad por el huésped reside en el enlace del ácido siálico-galactosa, siendo éste un punto determinante en la restricción de hospedero del virus de influenza y los eventos de zoonosis. Una vez internalizado el virus por endocitosis mediada por clatrina, el pH bajo de los endosomas permite un cambio conformacional de la HA que induce la fusión de las membranas viral y endosomal, lo que resulta en la pérdida de la envoltura lipídica del virus y en su salida al citoplasma celular. La acidificación del interior del virus, vía M2, permite que las RNP se liberen de M1 y sean transportadas al núcleo (Helenius, 1992).

En el núcleo ocurren los procesos de replicación y transcripción del genoma a partir del mismo templado y con la misma enzima. La síntesis de los mensajeros virales es dependiente de un elemento nuclear, la estructura de cap de guanina localizada en los extremos 5' de los mensajeros celulares, los cuales se ocupan como cebadores para la síntesis de los mRNA virales. Por otro lado, la replicación es libre de cebadores y se lleva a

cabo en dos pasos: la síntesis del RNA templado o antigenoma y la síntesis del genoma viral a partir del antigenoma. Una vez replicado el RNA viral, éste se empaqueta en un proceso selectivo en donde cada partícula viral debe tener el genoma completo compuesto de los 8 segmentos. Sin los 8 segmentos las partículas virales no son infecciosas.

El virus sale de la célula por un proceso de gemación en el cual adquiere una membrana lipídica derivada de la membrana plasmática de la célula. La proteína M1 se acumula en la parte interna de la membrana plasmática. Cuando la gemación se completa las espículas de HA de los virus salientes se encuentran unidas al ácido siálico de las glicoproteínas de la membrana celular. Las partículas virales se mantienen unidas a la membrana plasmática hasta que la proteína viral NA rompe esta unión y éstas logran liberarse de la membrana, lo cual permite la trasmisión del virus (Samji, 2009; Bouviera & Palesea, 2008).

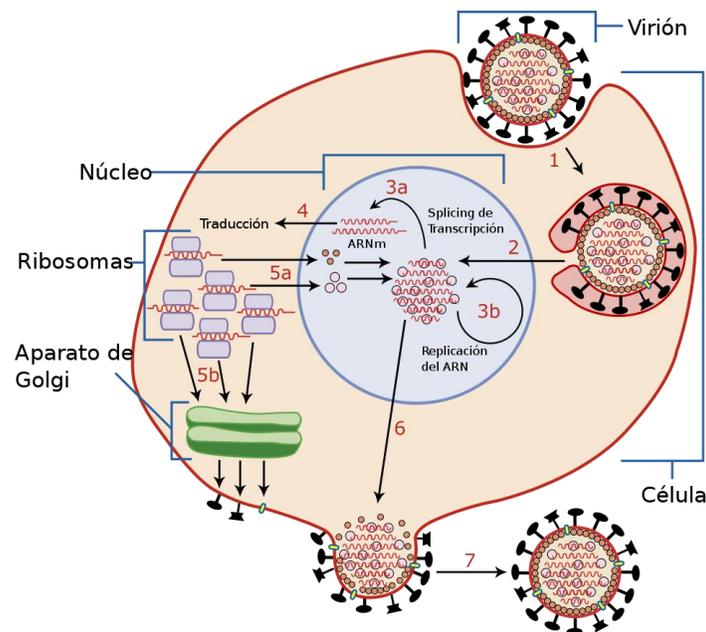


Figura 5. Ciclo infeccioso del virus de Influenza A. 1) El virión se une a la célula por interacción de HA con ácido siálico y entra a la célula por endocitosis mediada por receptor, 2) mediante la acidificación del endosoma HA se corta en H1-H2 lo que permite la liberación de las RNP, 3) en el núcleo se lleva a cabo la replicación y transcripción, 4) la traducción de los mRNA se lleva a cabo en citoplasma 5) ensamblaje de las RNP en el núcleo y procesamiento postraduccional de HA, NA y M2 en el complejo de Golgi que son exportadas a la membrana celular, 6)ensamble de la nueva partícula viral cerca de la membrana celular, 7) gemación de partículas virales, NA corta ácido siálico lo que permite la liberación de la partícula ([http://commons.wikimedia.org/wiki/File:Virus\\_Replication\\_large\\_es.svg](http://commons.wikimedia.org/wiki/File:Virus_Replication_large_es.svg))

## **2. Microarreglos**

### **2.1 Generalidades**

Los microarreglos de DNA son una metodología utilizada en biología molecular basada en hibridación de secuencias de ácidos nucleicos. Los microarreglos consisten en una base sólida y lisa, la cual puede ser de vidrio o membranas de nylon, en la que se encuentran Unidos arreglos de miles de oligonucleótidos con secuencias específicas en puntos discretos (*spots*). La hibridación de secuencias por complementariedad se detecta generalmente mediante fluorescencia, en donde las secuencias problema son marcadas con fluoróforos (Figura 6) (Miller & Tang, 2009).

Una de las características importantes de esta metodología es el manejo de un gran número de secuencias, lo que permite tener un espacio de búsqueda mucho mayor comparado con el PCR-multiplex. Esto genera una ventaja, ya que en un solo ensayo se puede caracterizar una mezcla compleja de ácidos nucleicos en búsqueda de miles de genes.

Esta metodología se utilizó inicialmente para realizar perfiles de expresión génica, sin embargo, actualmente existen una gran cantidad de aplicaciones, siendo una de ellas la detección de patógenos (Chen *et al.* 2011).

### **2.2 Trabajos previos**

Varios trabajos han reportado el uso de microarreglos como metodología de detección de patógenos virales (Li *et al.*, 2001; Wang *et al.*, 2002; Sengupta *et al.*, 2003; Townsend *et al.* 2006; Lenz *et al.*, 2008; Gall *et al.*, 2009; Li *et al.* 2009; Gardner *et al.*, 2010; Gauthier *et al.*, 2010; Sip *et al.*, 2010). De los trabajos enfocados a influenza, cada uno ha ocupado diseños experimentales distintos, sin embargo muchos han tratado de demostrar el uso de los microarreglos como una opción para detectar y subtipificar a este virus.

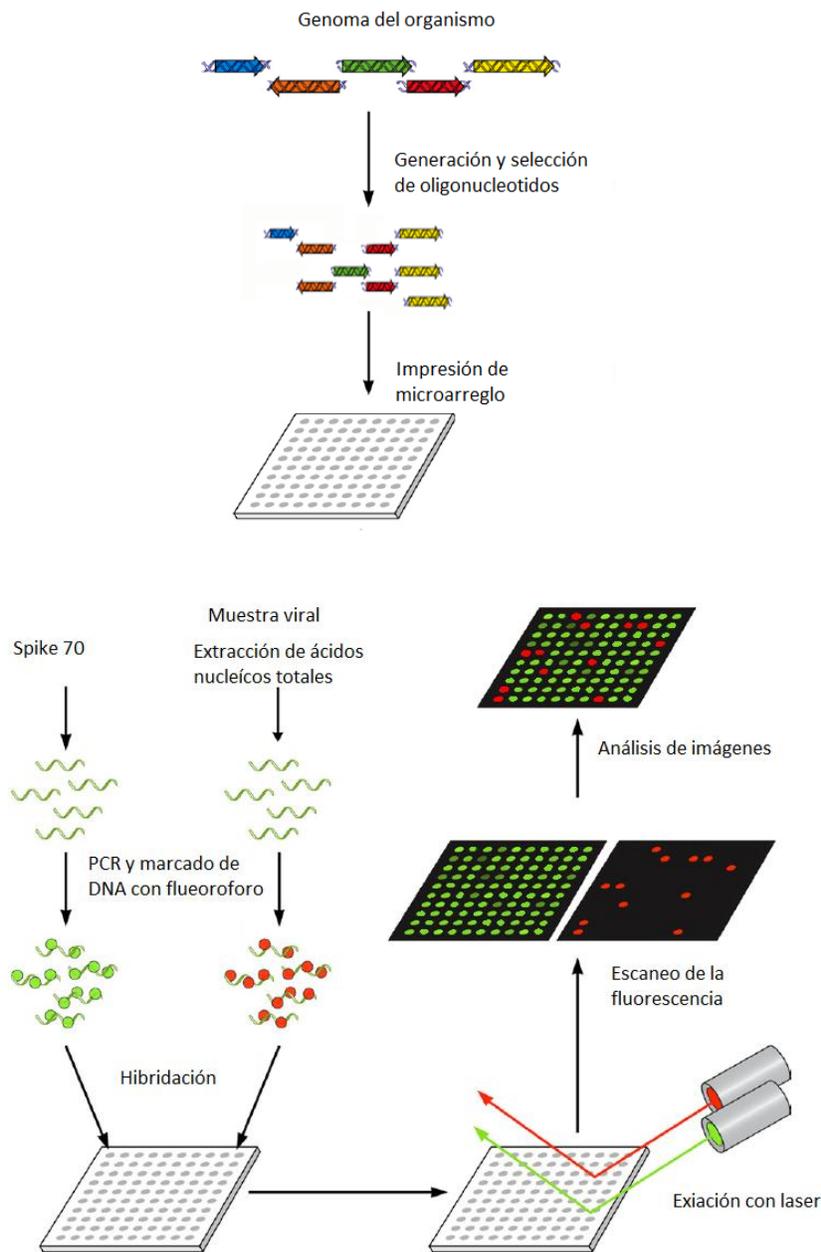


Figura 6. Estrategia utilizada para el desarrollo y validación del microarreglos de detección para influenza A. *Superior:* A partir de una colección de genomas de influenza A se generaron y seleccionaron oligonucleótidos de longitud 70 de acuerdo a diversos parámetros (Materiales y Métodos), los cuales fueron impresos en laminillas recubiertas de poli L-lisina. En cada punto (spot) se encuentra un oligonucleótido acompañado de una secuencia sin homólogos reportados en las bases de datos de longitud 70 (Spike 70). *Inferior:* Para la validación se hibridaron un conjunto de muestras de distintos hospederos. El material genético de la muestra es marcado con un fluoróforo. Adicionalmente a la muestra a hibridar se añade el complemento perfecto del Sipke 70, Probe 70, el cual sirve de control de hibridación y localización de cada uno de los spots. (Modificado de <http://cmr.asm.org/content/22/4/611/F2.large.jpg>)

En 2001 Li *et al.* diseñaron un microarreglo utilizando secuencias de 500 nucleótidos de longitud, producto de PCR, las cuales inmovilizaron en una placa de vidrio. En este estudio se lograron caracterizar 3 virus de influenza desconocidos. Su metodología permitía un desapareamiento de hasta el 14% y para el estudio utilizaron muestras aisladas previas a 1977.

En 2002 Wang *et al.* diseñaron un microarreglo para la detección y genotipificación de diversos patógenos virales utilizando un diseño de oligonucleótidos de DNA de 70 nucleótidos. En este trabajo usaron una estrategia de amplificación por PCR aleatorio antes de la hibridación del microarreglo, con la que pudieron detectar diversos virus respiratorios humanos sin la necesidad de cebadores específicos para cada virus.

En 2003 Sengupta *et al.* realizaron un microarreglo de detección y subtipificación de influenza A. Para ello escogieron un conjunto de oligonucleótidos específicos para el virus de influenza A de 18 a 28 nucleótidos de longitud la base de datos VirOligo (Onodera and Melcher, 2002). Los autores lograron discernir entre subtipos de hemaglutinina y neuraminidasa, además de identificar cada uno de los ocho segmentos que componen el genoma del virus de influenza A.

En 2006 Townsend *et al.* implementaron un microarreglo para la vigilancia epidemiológica del virus de influenza A. En él incluyeron los subtipos H1N1, H3N2 y H5N1, además de algunas secuencias de influenza B. El trabajo discute la rapidez con que se obtienen los resultados comparando con las metodologías tradicionales de identificación del virus de influenza.

En 2009 Gall *et al.* diseñaron un microarreglo para la detección y subtipificación de virus de influenza aviar. En este trabajo utilizaron secuencias que cubrían la variabilidad de los 16 subtipos de hemaglutinina.

En 2009 Li *et al.* diseñaron un microarreglo compuesto de 46 oligonucleótidos de entre 19 y 27 nucleótidos de longitud, los cuales detectaban 5 subtipos: H1N1, H1N2, H3N2, H5N1 y H9N2. En este trabajo los autores mencionan que el análisis de las muestras se realiza en 7 horas.

En 2010 Gardner *et al.* desarrollaron un microarreglo de detección microbial que abarcaba todos los virus conocidos, fagos, bacterias y plásmidos. En este trabajo tomaron en cuenta la variación de secuencia que se encuentra en los microorganismos, además de un análisis estadístico para poder analizar muestras complejas que contengan más de un organismo.

## II. Antecedentes

### 1. Metodologías de detección y vigilancia epidemiológica en México

Las metodologías clásicas para la detección de una cepa de influenza requieren el aislamiento del virus, cultivo y caracterización por ensayos inmunoquímicos. Para ello es necesario esperar de 4 a 7 días de cultivo previo a la prueba antigénica. En algunos casos también se realiza la prueba de detección por inmunofluorescencia directamente en las muestras respiratorias de los pacientes. Estas pruebas se realizan contra proteínas de la superficie del virus, por lo que sólo se conoce el subtipo al que pertenece el virus y se desconoce la variabilidad de los otros segmentos restantes del mismo (Townsend *et al.* 2006).

Actualmente, con la implementación de tecnologías de biología molecular en las prácticas rutinarias de laboratorio, se ha utilizado PCR multiplex para la detección del virus utilizando cebadores específicos contra el genoma. Esta metodología es mucho más rápida que las pruebas inmunoquímicas, sin embargo sólo se pueden buscar del orden de una decena de patógenos en paralelo, por lo que la capacidad de detectar en cada muestra varios patógenos en intervalos de tiempo cortos, es indudablemente una ventaja de los microarreglos.

En México la detección de cepas de influenza circulantes se limita a la indagación de unos cuantos subtipos de hemaglutinina, entre otros, los subtipos de hemaglutinina H1 y H3 por estar presentes en las cepas estacionales que infectan a humanos, y el subtipo H5 por estar presente en una cepa aviar de alto riesgo (H5N1). Este tipo de vigilancia epidemiológica es muy limitada ya que no se monitorean la mayoría de los subtipos de hemaglutinina y neuraminidasa, ni la diversidad de los segmentos restantes del genoma del virus, de los cuales en varios se ha reportado que están asociados a diversos aspectos de la virulencia del mismo. Las cepas de influenza cuyo subtipo no se logra caracterizar en nuestro país por el Instituto de Diagnóstico y Referencia Epidemiológicos (InDRE), son enviadas para su caracterización al centro de referencia de la OMS en el Centro de Control y Prevención de Enfermedades (CDC) en Atlanta, EUA.

## **2. Microarreglo de detección de virus asociados a enfermedades respiratorias**

En nuestro laboratorio, como su proyecto de maestría, MD Soto implementó un microarreglo para la detección de patógenos virales asociados a enfermedades respiratorias, así como la subtipificación del virus de influenza A. Este proyecto permitió adquirir experiencia, dando muy buenos resultados en el uso de microarreglos para detección de virus y representa una base teórica importante para el desarrollo del presente proyecto.

En paralelo, en nuestro laboratorio se está desarrollando un segundo microarreglo de detección viral. En este caso es un microarreglo para la detección de virus gastrointestinales, a cargo de MA Martínez.

### **III. Hipótesis**

Es posible diseñar e implementar una metodología basada en microarreglos que identifique el origen filogenético de cada uno de los ocho segmentos que componen el genoma del virus de influenza tipo A, el cual puede ser utilizado en vigilancia epidemiológica.

### **IV. Objetivos**

#### **Objetivo General**

Diseñar, implementar y validar una metodología para la detección y caracterización de los ocho segmentos del virus de Influenza tipo A mediante un microarreglo que permita identificar posibles rearrreglos en el genoma del virus.

#### **Objetivos Particulares**

- Diseñar un microarreglo que permita identificar el origen filogenético de cada uno de los genes del virus de influenza tipo A.
- Imprimir el microarreglo con los oligonucleótidos seleccionados.
- Validar el microarreglo con cepas de virus aisladas de aves, cerdos y humanos.

## **V. Materiales y Métodos**

### **1. Obtención de las secuencias**

Las secuencias de influenza A se obtuvieron de la base de datos de influenza de NCBI (Bao *et al.* 2008). Se descargaron todas las secuencias de nucleótidos existentes de influenza A hasta Julio de 2010, cuyo origen de colecta fuese distinto a humano, aviar o porcino. Sólo se tomaron en cuenta secuencias de genes completos y se colapsaron las secuencias idénticas tomando únicamente la de mayor antigüedad. Para secuencias de origen humano, porcino y aviar, se obtuvieron la secuencias utilizadas por Smith *et al* en 2009. Para no sesgar el análisis a todas las secuencias se les cambió el nombre por un identificador único. Este identificador únicamente tiene el nombre del segmento y un número consecutivo, sin embargo se conservó el nombre original en asociado al identificador único.

También se obtuvieron secuencias de genomas completos de virus de influenza A/H1N1 y A/H3N2 hasta abril de 2011, así como genomas completos de cepas de referencia donadas por el Dr. Webster del Departamento de Enfermedades Infecciosas del Hospital Pediátrico de Saint Jude en Memphis, Tennessee.

### **2. Generación de las bases de datos**

Todas las secuencias de nucleótidos obtenidas fueron formateadas para ser utilizadas por el algoritmo BLAST. Se realizaron distintas bases de datos para ser analizados por dicho algoritmo. Las bases de datos ocupadas son: base de datos con todas las secuencias disponibles de influenza A (FLUDB), base de datos de genomas de influenza A/H1N1 (FLUH1N1DB), base de datos de genomas de influenza A/H3N2 (FLUH3N2DB), base de datos de hemaglutinina y neuraminidasa (FLUHANADB) y base de datos con genomas de las cepas de referencia del hospital St. Jude (FLUSTJUDE).

### **3. Reducción de complejidad en el análisis**

Se utilizó el algoritmo CD-HIT (Li & Godzik, 2006) para realizar grupos por identidad en secuencia de nucleótidos con el fin de eliminar la redundancia que se encuentra en las bases de datos y reducir el tiempo de análisis. Se realizaron los grupos por segmento y, en el caso de la hemaglutinina y neuraminidasa por subtipo. Se utilizaron los parámetros por defecto del algoritmo con un corte de agrupamiento de 90% en identidad de secuencia de nucleótidos. Al descargar sólo secuencias completas en la base de datos de NCBI, la diferencia de tamaño entre las secuencias para cada segmento es mínima.

### **4. Análisis filogenéticos**

Se infirieron filogenias para cada segmento del genoma viral, y para el caso de HA y NA por subtipo. Los alineamientos múltiples se realizaron con el algoritmo MUSCLE utilizando los parámetros por defecto e indicando que se trata de secuencias de nucleótidos (Edgar, 2004). Para la inferencia de los árboles se utilizó el algoritmo Phyml (Guindon *et al.* 2010) por el método de máxima verosimilitud, bajo el modelo de sustitución Hasegawa–Kishino–Yano (HKY85). Adicional se utilizó una distribución Gamma para modelar la heterogeneidad de tasas de sustitución entre sitios (HKY + G).

### **5. Definición de clados**

Para cada uno de los segmentos del genoma de influenza A se definieron clados por identidad en secuencia y posición en el árbol filogenético. En el caso de la HA y NA éstos se definieron dentro de cada subtipo. La definición inicial de los clados se realizó a 90% de identidad en nucleótidos, los cuales fueron definidos por el algoritmo CD-HIT. Posteriormente cada uno de los grupos definidos fueron mapeados en los árboles filogenéticos. Si los grupos no concordaban con la topología de la filogenia, se bajaba el porcentaje de identidad en 1 unidad y se repetían las comparaciones con los árboles (Figura 7).

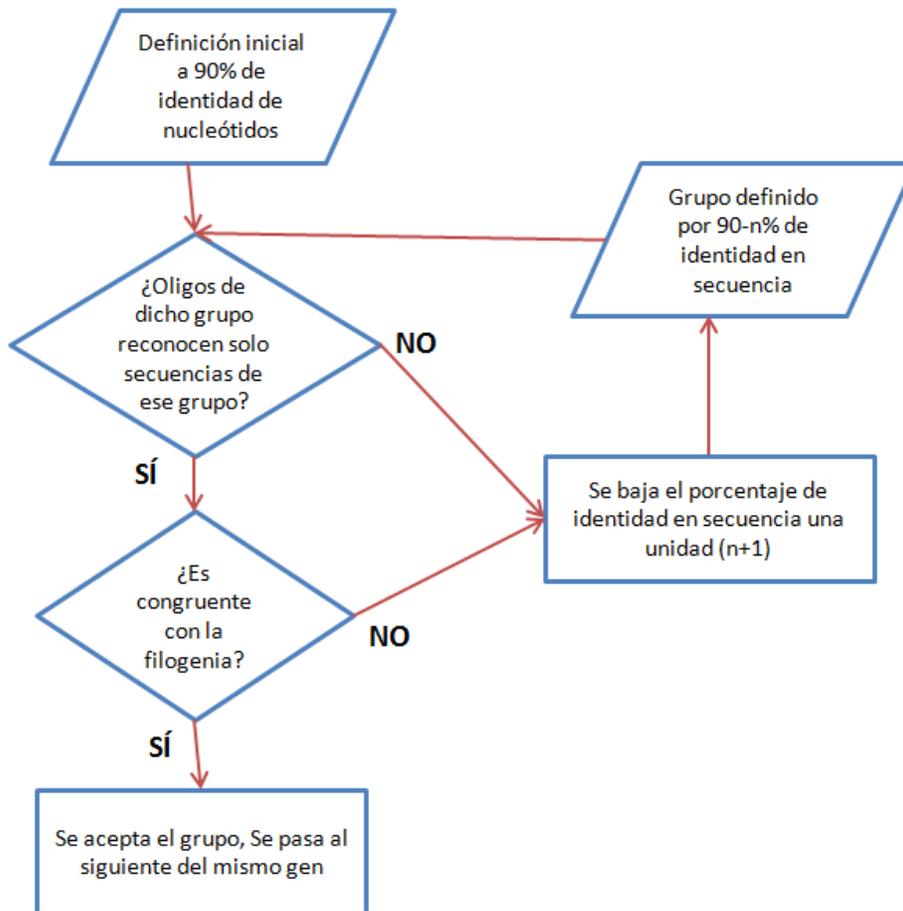


Figura 7. Estrategia para el desarrollo de un microarreglos de detección.

### 6. Generación de oligonucleótidos de longitud 70

Del análisis de agrupamiento se tomó una secuencia al azar por cada grupo. De cada una de las secuencias se extrajeron fragmentos de 70 nucleótidos con un corrimiento de 5 bases. Es decir, se tomó de la posición 1 a la 70 para el primer oligonucleótido, de la 6 a la 76 para el segundo y así sucesivamente hasta cubrir el largo de la secuencia. Cada oligonucleótido conservó el nombre de la secuencia de donde fue obtenido, así como las posiciones de inicio y fin.

## **7. Selección de oligonucleótidos**

Los oligonucleótidos generados se clasificaron de acuerdo a su especificidad. Se consideraron específicos aquellos que únicamente reconocieran secuencias dentro de su clado, mas no de clados distintos al propio (Figura 7).

Para la clasificación se realizaron alineamientos de los oligonucleótidos generados con la base de datos con que inició el análisis mediante el algoritmo BLAST (blastn, por tratarse de secuencias de nucleótidos). Del resultado de los alineamientos se tomó en cuenta el número de nucleótidos alineados, el valor  $e$  del algoritmo BLAST y se calculó el  $\Delta G$  teórico de hibridación por el método Nearest Neighbor (Gray, 1997) mediante un script en Perl escrito por María de los Dolores Soto del Río.

El método de Nearest Neighbor realiza el cálculo de la  $\Delta G$  teórica de hibridación calculando la  $\Delta S$  y  $\Delta H$  de cada interacción. El método es sensible a la posición del nucleótido, es decir, da valores diferentes si éste se encuentra en los extremos o en la parte central de la secuencia, y al tipo de desapareamientos.

Con estos parámetros se definieron dos grupos de oligonucleótidos:

- Oligonucleótidos específicos: Aquellos que reconocen secuencias dentro de su clado con un  $\Delta G$  teórico de hibridación menor o igual a  $-75$  kcal/mol, un mínimo de 60 nucleótidos alineados y un valor de  $e$  menor a  $10^{-15}$ , y que reconocen secuencias fuera de su clado con un  $\Delta G$  teórico de hibridación mayor a  $-55$  kcal/mol, que tengan menos de 55 nucleótidos alineados y un valor de  $e$  mayor a  $10^{-12}$  (Figura 8).
- Oligonucleótidos inespecíficos: Aquellos que no cumplan los valores de corte anteriores.

Esta clasificación permitió obtener oligonucleótidos que reconocen secuencias sólo dentro de su clado.

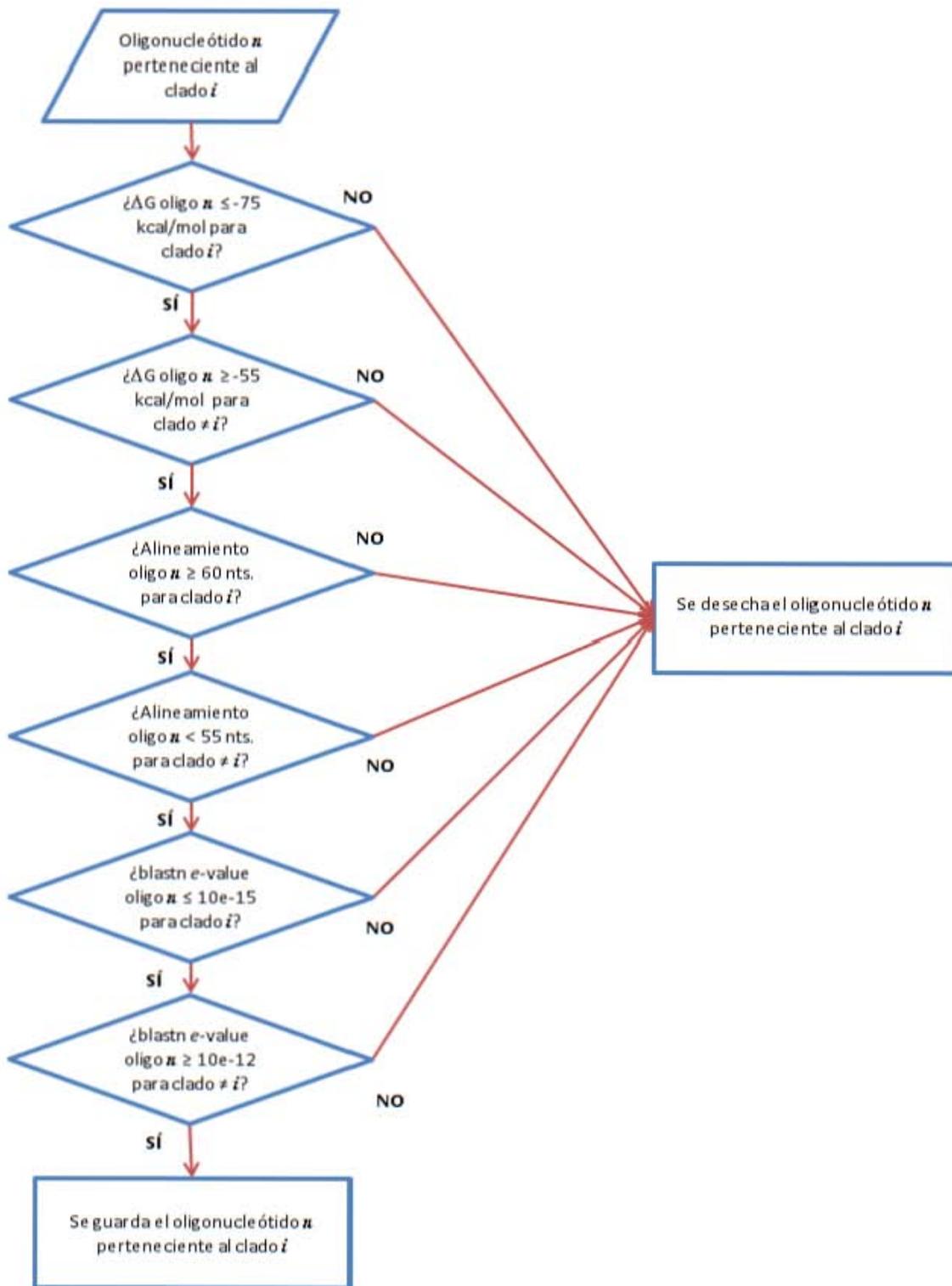


Figura 8. Diagrama de flujo que ejemplifica el proceso de selección de los oligonucleótidos específicos y no específicos.

Una vez obtenido el grupo de oligonucleótidos específicos, se escogió el número mínimo de oligonucleótidos con los cuales se reconociera a todas las secuencias de su clado al menos 3 veces. Asimismo, se escogieron aquellos oligonucleótidos que tuvieran el  $\Delta G$  teórico de hibridación menor y que cubrieran diferentes regiones del segmento.

Al conjunto de oligonucleótidos seleccionados se les dio un nombre único de acuerdo al segmento que reconocen, el clado al que pertenecen y al número de oligonucleótido que representan para dicho clado.

### **8. Ensayos *in silico***

Se realizaron alineamientos de los oligonucleótidos que serían impresos en el microarreglo con diversos conjuntos de secuencias para observar el comportamiento de los oligonucleótidos con secuencias conocidas. Los alineamientos fueron realizados con el algoritmo BLAST (blastn). Se tomaron los 2 parámetros para tomar como positiva una hibridación: 60 de los 70 nucleótidos apareados, y una  $\Delta G \leq -70$  kcal/mol, es decir más laxa que en la selección. Los conjuntos utilizados fueron:

- a) Genomas de cepas de referencia de influenza A del hospital St. Jude, la cual contiene los 16 subtipos de hemaglutinina y los 9 de neuraminidasa. Estos resultados sirvieron para comparar los resultados de hibridación de las cepas de referencia en la validación del microarreglo.
- b) Genomas de influenza A/H1N1 provenientes de ave, cerdo y humano hasta mayo de 2011 y genomas de influenza A/H3N2 provenientes de ave, cerdo, perro, gato, visón y humano hasta mayo de 2011. Con este ensayo se observó la variabilidad de los virus circulantes en humano (A/H1N1 y A/H3N2) y se comparó con la existente en otros hospederos que infecta.
- c) Genes completos de influenza A con origen de colecta de todos los orígenes hasta mayo de 2012, con el que se realizó el análisis de asignación de un origen a cada uno de los clados definidos.
- d) Genomas de influenza A de todos los orígenes hasta marzo de 2012, exceptuando las secuencias de influenza A/H17N10, rearrreglantes y ambientales. Con estos resultados se definieron los genomas únicos en 5 hospederos: humano, aves, cerdos, caballos y perros.

## **9. Conversión de clado a origen**

Para poder asignar el origen animal de cada uno de los clados creados para cada segmento del genoma del virus de influenza A se utilizó el ensayo *in silico* contra una colección de genes completos de influenza A de todos los orígenes. Para cada clado se analizó que origen tenían las secuencias que los oligonucleótidos reconocían con los parámetros antes descritos. Posteriormente se determinó la proporción de secuencias que pertenecían a un origen humano, porcino, aviar, equino, canino o de otro origen, el cual era representado por secuencias que no caen en ninguno de los anteriores. A partir de este análisis, si el 90% de un clado estaba formado por secuencias de un solo origen, éste se le asignaba al clado. Por ejemplo, si el clado 0 del segmento 1 tenía 46 / 50 secuencias de origen aviar y 4 / 50 de origen porcino, el clado 0 se definía como aviar.

Adicionalmente se analizaron los ensayos *in silico* realizados con genomas de influenza provenientes de humano, cerdos, aves, caballos y perros, con el fin de identificar que virus han circulado en los diferentes hospederos, cuyo patrón de combinación de genes es único, basados en la clasificación de clados descritos anteriormente, por ejemplo **1\_C0 2\_C1 3\_C3 4\_H1\_C1 5\_C4 6\_N1\_C1 7\_C0 8\_C1**. A estos patrones se les denominó "genomas únicos" los cuales son la combinación de clados y subtipos que componen un genoma de influenza A. Dicha combinación aparece una vez y se repite por cierto número de años, y en ocasiones llega a no observarse por largos periodos. Con los genomas únicos podemos observar si una determinada combinación de clados para cada segmento ya había sido reportada, en qué año y en qué hospedero, lo cual nos permite trazar una línea de origen.

## **10. Preparación de oligonucleótidos**

Los oligonucleótidos se sintetizaron químicamente por la compañía Invitrogen-Illumina, en San Diego California, dispuestos en placas de 384 pozos a una concentración de 40mM. Estos se resuspendieron en 3x SSC (0.45 M cloruro de sodio, 0.045 M citrato de sodio, pH 7) a una concentración de 50 µM.

A cada uno de los oligonucleótidos en las placas de 384 pozos se les añadió un oligonucleótido control denominado "Spike-70", sintetizado por la misma compañía, cuya secuencia es 5' -**ACCTCGCTAACCTCTGTATTGCTTGCCGGACGCGAGACAAACC TGAACATTGAGAGTCACCCTCGTTGTT**-3', el cual fue impreso en una proporción de 1:50 con respecto al oligonucleótido presente en cada uno de los pozos de la placa (Urisman *et al*, 2005).

### **11. Impresión de microarreglos**

La impresión de los microarreglos se llevó a cabo en la Unidad de Microarreglos del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México, sobre portaobjetos cubiertos con poli-L-lisina. Los oligonucleótidos se unen al vidrio mediante interacciones electrostáticas. Cada una de las sondas está representada cuatro veces en el microarreglo, lo que se utiliza en el análisis para dar un mayor peso estadístico al resultado.

En el mismo lugar donde se imprime la sonda se encuentra una secuencia llamada Spike 70 la cual es una secuencia de DNA de longitud 70 la cual no tiene homólogos reportados en las bases de datos públicas. Este sirve como control de localización e hibridación a la hora de realizar los experimentos.

### **12. Post-procesamiento de los microarreglos**

Los oligonucleótidos impresos en el portaobjetos de vidrio se someten a un proceso de preparación previo a la hibridación.

Primero se marca con un lápiz de punta diamante la parte trasera del vidrio en la zona en que está impreso el microarreglo. Esto debido a que en los pasos subsecuentes los microarreglos no podrán ser vistos y es necesario conocer la zona en la que están impresos. Posteriormente, para unir los oligonucleótidos de manera covalente al vidrio, se entrecruzan mediante la exposición a luz ultravioleta de 60mJ durante 20 minutos.

Posteriormente las laminillas se sumergen en una solución SSC 3X y SDS 0.2% durante 5 min a 65°C. Se lavan dos veces con agua para eliminar el exceso de SDS y posteriormente se lavan con etanol al 95% y se secan por centrifugación por 5 min a 50 x *g*. A continuación las laminillas se sumergen en una solución con 5.5 g de anhídrido succínico (Sigma) disueltos en 15 ml de borato de sodio 1M a pH 8.0 y 355 ml de 1-metil-2-pirrolidinona (Sigma) por 15 min. Lo anterior con la finalidad de bloquear los grupos amino expuestos de la poli-L-lisina del vidrio, a través de la unión del anhídrido succínico a las cargas positivas libres. Finalmente los microarreglos se lavan con agua, y etanol al 95%. Las laminillas se secan por centrifugación durante 5 min a 50 x *g*. Estos pasos deben hacerse lo más rápido posible para evitar que las laminillas se sequen antes de la centrifugación y queden manchados. Es importante mencionar que el paso de bloqueo debe realizarse en una campana de extracción debido a que la pirrolidinona es un compuesto tóxico. Una vez procesadas las laminillas son guardadas en un desecador.

### ***13. Cepas de influenza A utilizadas para validar el microarreglo***

Los primeros experimentos de hibridación se realizaron con 17 de las 33 cepas de referencia de influenza A (Tabla 1), que representan los 16 subtipos de HA y los 9 de NA donadas por el Doctor R.G. Webster, del Departamento de Enfermedades Infecciosas del Hospital Pediátrico de Saint Jude en Memphis, Tennessee. Las muestras que fueron donadas contienen RNA genómico por lo que es posible detectar todos los segmentos del genoma de influenza A.

También se analizaron virus suministrados por el Dr. Iván Sánchez, del Departamento de Producción Porcina de la Facultad de Medicina Veterinaria y Zootecnia (FMVZ) de la UNAM. Estos virus de influenza A fueron aislados y crecidos en embrión de pollo por el Dr. Sánchez a partir de biopsias pulmonares de cerdos con secreción nasal, tos y/o estornudo y lesión pulmonar. Las muestras se colectaron en rastros de 17 estados del país de noviembre de 2009 a febrero de 2011 (Tabla 2). Las muestras colectadas se probaron por RT-PCR en tiempo real en la FMVZ. Aquéllas que resultaron positivas para influenza A las analizaron en el laboratorio del Dr. Sánchez por RT-PCR punto final para subtipificarlas,

buscando los subtipos H1N1 y H3N2 porcinos, que son los que circulan usualmente de manera estacional en la población porcina. En nuestro laboratorio se hizo RT-PCR dirigido a amplificar una región del gen M de influenza A para comprobar la presencia del virus en las muestras. Para esto se utilizaron los oligonucleótidos (sentido 5' – **CTTCTAACCGAGGTCGAAACG**–3' y antisentido 5' –**CATGCAACTGGCAAGTGCACC**–3' ) utilizando el kit SuperScript One-Step RT-PCR Platinum Taq III de Invitrogen.

Adicionalmente se analizaron 20 muestras provenientes de aves del estado de Sonora (Tabla 3), una muestra de influenza A/H1N1 aislada en el laboratorio a partir de un enfermo durante la pandemia de 2009 y 10 muestras de pacientes humanos (Tabla 4).

**Tabla 1. Cepas de referencia utilizadas para la validación del microarreglo. Se utilizaron 17 de las 33 cepas de referencia, las cuales representan los 16 subtipos de HA y los 9 de NA**

Muestra	Lugar de colecta	Subtipo	Nombre de la cepa
3	Fort Monmouth	H1N1	Influenza A virus (A/Fort_Monmouth/1/1947(H1N1))
4	New Jersey	H1N1	Influenza A virus (A/New_Jersey/1976(H1N1))
5	Singapore	H2N2	Influenza A virus (A/Singapore/1-MA12C/1957(H2N2))
8	Miami	H3N8	Influenza A virus (A/equine/Miami/1/1963(H3N8))
10	Czechoslovakia	H4N6	Influenza A virus (A/duck/Czechoslovakia/1956(H4N6))
12	Australia	H6N5	Influenza A virus (A/shearwater/Australia/1/1972(H6N5))
15	New Jersey	H7N3	Influenza A virus (A/ruddy turnstone/NJ/65/1985(H7N3))
16	Ontario	H8N4	Influenza A virus (A/turkey/Ontario/6118/1968(H8N4))
17	Wisconsin	H9N2	Influenza A virus (A/turkey/Wisconsin/1/1966(H9N2))
18	Germany	H10N7	Influenza A virus (A/chicken/Germany/n/1949(H10N7))
19	Memphis	H11N9	Influenza A virus (A/duck/Memphis/546/1974(H11N9))
20	Alberta	H12N5	Influenza A virus (A/duck/Alberta/60/1976(H12N5))
21	Maryland	H13N6	Influenza A virus (A/gull/Maryland/704/1977(H13N6))
22	Astrakhan	H14N5	Influenza A virus (A/mallard/Astrakhan/263/1982(H14N5))
23	Western Australia	H15N9	Influenza A virus (A/wedge-tailed_shearwater/W_Australia/2576/1979(H15N9))
24	Delaware	H16N3	Influenza A virus (A/shorebird/Delaware/168/06(H16N3))
27	South Africa	H5N3	Influenza A virus (A/tern/South_Africa/61(H5N3))

**Tabla 2. Muestras de virus porcinos utilizadas para la validación del microarreglo, 10 de éstas fueron secuenciadas y comparados los resultados.**

Muestra	Fecha de colecta	Lugar de colecta	Subtipo	Nombre de la cepa
FMVZ 04		Tamaulipas	-	-
FMVZ 19	2009	Estado de México	H1N1	A/swine/Mexico/Mex19/2009/(H1N1)
FMVZ 29	2010	Veracruz	H1N1	A/swine/Mexico/Ver29/2010/(H1N1)
FMVZ 31	2010	Veracruz	H1N1	A/swine/Mexico/Ver31/2010/(H1N1)
FMVZ 32	2010	Querétaro	H1N1	A/swine/Mexico/Qro32/2010/(H1N1)
FMVZ 35	2010	Querétaro	H1N1	A/swine/Mexico/Qro35/2010/(H1N1)
FMVZ 37	2010	Veracruz	H1N1	A/swine/Mexico/Ver37/2010/(H1N1)
FMVZ 40	2010	Michoacán	H3N2	A/swine/Mexico/Mich40/2010/(H3N2)
FMVZ 50	2010	Estado de México	H3N2	A/swine/Mexico/Mex50/2010/(H3N2)
FMVZ 51	2010	Estado de México	H3N2	A/swine/Mexico/Mex51/2010/(H3N2)
FMVZ 52	2010	Estado de México	H1N1	A/swine/Mexico/Mex52/2010/(H1N1)

**Tabla 3. Muestras de virus aviares utilizadas para la validación del microarreglo. Todas fueron secuenciadas (Datos no analizados).**

Muestra	Fecha de colecta	Lugar de colecta	Subtipo	Especie
150C	2007	Cd. Obregón, Sonora	-	<i>Anas strepera</i>
769C	2008	Cd. Obregón, Sonora	H9N2	<i>Anas americana</i>
796C	2008	Cd. Obregón, Sonora	-	<i>Anas clypeata</i>
797C	2008	Cd. Obregón, Sonora	H5N3	<i>Anas clypeata</i>
827C	2009	Cd. Obregón, Sonora	-	<i>Anas crecca</i>
829C	2009	Cd. Obregón, Sonora	H6N5	<i>Anas crecca</i>
840C	2009	Huatabampo, Sonora	-	<i>Anas crecca</i>
1132C	2009	Huatabampo, Sonora	H10N3	<i>Anas crecca</i>

**Tabla 4. Muestras de virus humanos utilizadas para la validación del microarreglo. Se tienen virus pandémicos y estacionales**

Muestra	Fecha de colecta	Lugar de colecta	Subtipo
ibt2	2009	Distrito Federal	H1N1
2662	2003	Distrito Federal	H3N2
2664	2003	Distrito Federal	H3N2
25216	2009	Jalisco	H3N2
2408	2005	San Luis Potosí	-
1	-	Veracruz	H3N2
33	-	Hospital General de México	H3N2
121	-	Hospital Civil de Guadalajara	H3N2
127	-	Hospital Civil de Guadalajara	-
128	-	Hospital Civil de Guadalajara	-
A/H1N1 estacional	-	-	H1N1

#### **14. Extracción de RNA y amplificación del material genético presente en las muestras**

La extracción de RNA viral a partir de lisado de células infectadas con virus se hizo utilizando el kit PureLink™ RNA Mini Kit de Invitrogen, de acuerdo al protocolo de los fabricantes, partiendo de 200 µl de muestra inicial y eluyendo el RNA en 32µl de agua libre de RNAsas. Una vez extraído el material genético, éste se amplificó a través de 3 rondas de PCR.

##### *a. Ronda A*

En este paso se realiza la transcripción reversa del RNA total presente en la muestra y la síntesis de DNA de doble cadena, utilizando el **cebador A (5'-CACTGGAGGATANNNNNNNNN-3')** en donde N es cualquier nucleótido A, T, C o G. Se utilizó el kit SuperScript III Reverse Transcriptase de Invitrogen con las siguientes cantidades de reactivos por muestra:

**2.0 µl de 5x RT buffer**

**0.5 µl de 100mM DTT**

**1.0 µl de 12.5 mM dNTPs**

**1.0 µl de H<sub>2</sub>O libre de RNAsas**

**0.5 µl de la enzima transcriptasa reversa SuperScript III RT (200 unidades/µl).**

Del RNA extraído se toman 5 µl, se mezcla con 1 µl del cebador A, el cual se encuentra a una concentración de 40 pmol/µl. Se hierve por 5 min y posteriormente se añaden 5 µl de la mezcla previamente descrita. Se incuba una hora a 42°C, seguido de 10 min a 94°C.

Para la síntesis de DNA de doble cadena se utiliza la enzima Sequenasa 2.0 (DNA polimerasa 1 del fago T7 que carece de la actividad 3-5' exonucleasa, y tiene una alta capacidad procesiva) de Affymetrix/USB. El protocolo para la síntesis de DNA doble cadena fue el siguiente:

94°C	2 min	} 2 Ciclos
10°C	10 min	
10°C	2 min	
15°C	2 min	
20°C	2 min	
25°C	2 min	
30°C	2 min	
37°C	8 min	
94°C	10 min	
10°C	10 min	

En el primer ciclo se añaden a cada muestra 5 µl de la mezcla de reactivos I [3.8 µl de H<sub>2</sub>O, 1.0µl de 5x Sequenase buffer, 0.15 µl de enzima Sequenasa (13 unidades/µl)], y en el segundo ciclo se añaden 0.6 µl de la mezcla de reactivos II [0.45 3.8 µl de H<sub>2</sub>O, 1 µl de 5x Sequenase buffer, 0.15 µl de enzima Sequenasa (13 unidades/µl)].

*b. Ronda B*

Esta es la primera ronda de amplificación, en la cual se utilizan 5µl del producto de la Ronda A como templado. En esta ronda se utiliza el **cebador B (5'-GTTCCCACTGGAGG ATA-3')**, el cual es complementario con la parte conocida del cebador A y tiene 6 bases más en el extremo 5'. Para la PCR se utiliza el kit KlenTaq LA de Sigma (mezcla de una mutante de la Taq polimerasa con una delección en el extremo N-terminal, por lo que no tiene actividad de exonucleasa, y una pequeña cantidad de una DNA polimerasa termoestable con actividad de 3'-exonucleasa). La mezcla de los reactivos para una muestra es la siguiente:

- 5 µl de 10X buffer para KlenTaq LA
- 1 µl de 12.5 mM dNTPs
- 1 µl del cebador B a una concentración de 100pmol/µl
- 37 µl de H<sub>2</sub>O libre de RNAsas
- 1 µl de KlenTaq LA (5 unidades/µl).

El protocolo de amplificación es:

94°C	2 min	}	25 Ciclos
94°C	30 seg		
50°C	1 min		
72°C	1 min		
72°C	10 min		
10°C	Mantenido		

*c. Ronda C*

Para la segunda ronda de amplificación se toma como templado 5µl del producto de la Ronda B. En esta ronda se utiliza la misma enzima de la ronda B, sin embargo varía la concentración del cebador B, que para este caso es de 40 pmol/µl y los dNTPs utilizados, los cuales son una mezcla de los 4 desoxirribonucleótidos más 5-(3-aminoalil)-2'-deoxiuridinda 5'-trifosfato (aa-dUTP) de TriLink Biotechnologies, el cual es un derivado del nucleótido uracilo con un grupo amino reactivo. El aa-dUTP/dTTP se usa en una proporción 7:3 aa-dUTP/dTTP.

El protocolo de amplificación es:

94°C	2 min	}	20 Ciclos
94°C	30 seg		
50°C	1 min		
72°C	1 min		
72°C	10 min		
10°C	Mantenido		

Finalmente se realiza una electroforesis en gel de agarosa al 1% para separar los productos de PCR y confirmar la presencia de los amplicones, los cuales se observan como barridos de DNA con un tamaño mayor a 300 pares de bases.

*d. Amplificación de genoma completo (WGA)*

Para algunas muestras, se realizó un paso de amplificación previo a la ronda A. En este paso se toman de 2 a 5 µl de RNA como templado para la reacción. Para este paso se utilizan 2 cebadores (Zhou *et al.* 2009) a una concentración de 100 pmol los cuales son complementarios a los extremos 5' y 3' del virus de influenza A que tienen 12 y 13 nucleótidos conservados (**cebador MBT uni 12 (5'-ACGCGTGATCAGCAAAAGCAGG-3')**, **cebador MBT uni 13 (5'-ACGCGTGATCAGTAGAAACAAGG-3')**). Se utilizó el kit SuperScript III One Step RT-PCT Platinum Hi-Fi de Invitrogen con las siguientes cantidades de reactivos por muestra:

**12.0 µl de 2x RT buffer**

**1.0 µl del cebador MBT uni 12 100pmol**

**1.0 µl del cebador MBT uni 13 100pmol**

**7.0 µl de H<sub>2</sub>O libre de RNAsas**

**0.5 µl de la enzima transcriptasa reversa SuperScript III RT Platinum Taq HiFi (200 unidades/µl).**

El protocolo de amplificación es:

42°C	60 min	}	x1
94°C	2 min		
94°C	30 seg	}	x5
45°C	30 seg		
68°C	3 min		
94°C	30 seg	}	x31
57°C	30 seg		
68°C	3 min		
68°C	8 min	}	x1
4°C	Mantenido		

### **15. Purificación y marcaje del DNA**

Los fragmentos de DNA amplificado se recuperan por medio de columnas de afinidad (Zymo columns DNA Clean & Concentrator™-5, Zymo Research) de acuerdo al protocolo del fabricante. El DNA se eluye en 11µl de agua. Seguido de esto se realiza el acoplamiento del fluoróforo a la muestra. Inicialmente los colorantes Cy3 (532 nm-verde) y Cy5 (635 nm-rojo) se resuspenden en 50 µl de dimetil sulfóxido (DMSO) (el fluoróforo no usado se alicuota y se guarda a -70°C). Se preparan dos mezclas de reacción, la primera con los 11µl del producto purificado de la ronda C se mezclan con 1.5 µl de amortiguador de carbonatos de sodio 1M (pH9) y 5µl del colorante Cy5 (Amersham) resuspendido. La segunda es una mezcla de 1.5 µl del amortiguador de carbonatos de sodio 1M (pH9), 5 µl de Cy3 (Amersham), 7µl de agua y 2 µl de la sonda control (Probe 70). Esta sonda contiene 5 bases modificadas con un grupo amino unido al carbono 6 [AmC6-dT] para el posterior acoplamiento del fluoróforo. Este Probe 70 (5' -**AACAACGAGGG [AmC6-dT] GACTCTCAA [AmC6-T] GTTCAGGTTTGTC [AmC6dT] CGCGTCCGGCAAGCAA [AmC6-dT] ACAGAGGT [AmC6-dT] AGCGAGGT-3'**), es complementario al Spike70 (Urisman *et al* 2005) y está a una concentración de 1 pg/µl.

Las muestras se incuban por una hora en oscuridad a temperatura ambiente para permitir que el grupo amino reactivo de los nucleótidos amino-alil presentes en las cadenas amplificadas del DNA reaccione con el N-hidroxisuccinimidil éster de los fluoróforos. Las muestras se someten a una segunda ronda de purificación para eliminar el fluoróforo no acoplado y concentrar la cantidad de DNA marcado, utilizando las columnas de Zymo. El DNA producto de las amplificaciones se eluye en 18 µl de agua y la sonda Probe 70 marcada se eluye en 30 µl. Posteriormente se realiza una cuantificación de la cantidad de fluoróforo incorporado por µg de DNA utilizando un lector de absorbancia NanoDrop (ND-1000 Thermo Scientific).

## **16 .Hibridación**

Para la hibridación de las muestras éstas se preparan mezclando 12.5µl de agua, 2.5 µl de la preparación de Probe70 marcado con Cy5 y 17 µl de cada muestra marcada con Cy3. Posteriormente se le añade a cada muestra 6 µl de SSC 20x pH 7, 1 µl de 1M HEPES (ácido 2-[4-(2-hidroxietil)-1-pireraciniil-etanosulfónico) pH 8 y 1 µl de SDS al 10%.

Se hierven las muestras por 2 min, se centrifugan rápidamente y se dejan enfriar antes de colocarlas sobre el microarreglo. Se colocan las laminillas post procesadas en cámara de hibridación (Die Tech 5005) y se pone un cubreobjetos M Series Lifterslip (Thermo Scientific) de manera que cubra completamente el área donde se encuentra impreso el microarreglo. Entre cada cubreobjetos y laminilla se introduce por capilaridad una “película” 30 µl de la mezcla de productos de PCR y Probe 70 marcados con los fluoróforos. Se vierten 20 µl de SSC 3X sobre cada cubreobjetos para ayudar a evitar la desecación de la muestra y el movimiento del cubreobjetos. La cámara de hibridación se cierra herméticamente y se coloca en un baño a 65°C de 8 a 16 h.

## **17 .Lavados y lectura de los microarreglos**

Los microarreglos se sacan cuidadosamente de la cámara de hibridación y se introducen en un recipiente con 2X SSC y 0.03% SDS a 65°C. Se retira el cubreobjetos mientras el portaobjetos permanece sumergido en la solución. Posteriormente las laminillas se transfieren a una rejilla y se sumerge vigorosamente en esta misma solución durante 1 min. Seguido de esto se realizan dos lavados con una solución SSC 1X y finalmente un lavado con una solución de SSC 0.2X. Los microarreglos se secan por centrifugación durante 5 min a 50 x g. Una vez secos, estos se leen en un lector GenePix 4000B utilizando en general condiciones de lectura con una ganancia del fotomultiplicador de 500 para Cy3 con un poder de láser de 100%, y una ganancia del fotomultiplicador de 750 para Cy5 con un poder de láser de 100%.

## 18. Análisis de datos

Las imágenes en formato TIFF producto del escaneo se analizaron mediante el software GenePix 6.0.2 que viene con el equipo lector. El programa hace un mapeo de la posición de cada oligonucleótido, relaciona cada punto con su secuencia respectiva con ayuda de un archivo de referencia (.gal) y genera un archivo de salida (.gpr) que contiene toda la información obtenida por el lector, como intensidad y ruido de cada canal, nombre de los oligonucleótidos y medidas de normalización, por nombrar algunos.

A partir del archivo .gpr se calcula el valor normalizado de la intensidad de cada punto para el canal marcado con Cy3 con la siguiente fórmula:

$$\text{Valor normalizado} = \left( \frac{\text{media de la señal Cy3 del punto } xy}{\text{media de la señal Cy3 de todos los puntos}} \right) - \left( \frac{\text{media del ruido Cy3 del punto } xy}{\text{media del ruido Cy3 de todos los puntos}} \right)$$

Todos los resultados que se reportan en este trabajo están basados en este cálculo normalizado para cada uno de los oligonucleótidos impresos en el microarreglo.

El principal objetivo del análisis de los datos es la identificación de aquellos oligonucleótidos que tengan una señal significativamente alta con respecto al resto, para poder saber qué material genético estaba presente en la muestra analizada.

Para dar un soporte a los resultados se utilizó una distribución hipergeométrica, la cual nos da un valor de probabilidad de acuerdo al número de oligonucleótidos vistos para un grupo dado mediante la siguiente fórmula:

$$P(Y = k) = \frac{\binom{D}{k} \binom{N-D}{n-k}}{\binom{N}{n}}$$

Donde  $P$  es la probabilidad de obtener  $k$  elementos de la categoría  $D$ . Para el caso concreto del microarreglo:

- $D$  es el número de oligonucleótidos totales para un clado dado
- $k$  el número de oligonucleótidos encontrados para un clado dado
- $N$  el número de oligonucleótidos impresos en el microarreglo
- $n$  el número de casos tomados.

Para determinar el número de casos tomados (oligonucleótidos que se consideraran para el análisis  $n$ ), se tomó el valor que corresponde al 5% del valor de intensidad normalizada. Todos aquellos oligonucleótidos que estén por encima de este valor se toman en cuenta, por lo que  $n$  no está definida *a priori*. Este análisis fue realizado en el cluster del Instituto de Biotecnología de la UNAM con el lenguaje y entorno de programación para análisis estadísticos R ([r-project.org](http://r-project.org)).

## VI. Resultados

### 1. Desarrollo del microarreglo

#### 1.1 Obtención de las secuencias

De la base de datos de influenza de NCBI se descargaron 1492 secuencias completas pertenecientes a virus aislados de 18 hospederos (camello, perro, gato, civeta, caballo, hurón, oso hormiguero, leopardo, visón, rata almizclera, pica, pica del altiplano, perro mapache, foca, garduña, tigre y ballena). Estas secuencias se filtraron para quitar las secuencias idénticas. Adicionalmente se obtuvieron 6413 secuencias de origen humano, porcino y aviar las cuales fueron utilizadas por Smith *et al* en 2009 (Tabla 5, Tabla 6 y Tabla 7).

Tabla 5. Numero de secuencias totales por cada segmento del genoma de influenza A.

Segmento / Gen	No. de secuencias
1 Polimerasa PB2	926
2 Polimerasa PB1	897
3 Polimerasa PA	910
4 Hemaglutinina (HA)	1066
5 Nucleoproteína	906
6 Neuraminidasa (NA)	1043
7 Proteína de la matriz - Canal M2	1001
8 NS1 - NS2	1146

Tabla 6. Numero de secuencias totales para cada subtipo de hemaglutinina.

Subtipo HA	No. de secuencias	Subtipo HA	No. de secuencias	Subtipo HA	No. de secuencias
H1	206	H7	77	H12	8
H2	41	H8	5	H13	3
H3	379	H9	61	H14	2
H4	18	H10	8	H15	7
H5	189	H11	9	H16	8
H6	45				

**Tabla 7. Numero de secuencias totales para cada subtipo de neuraminidasa.**

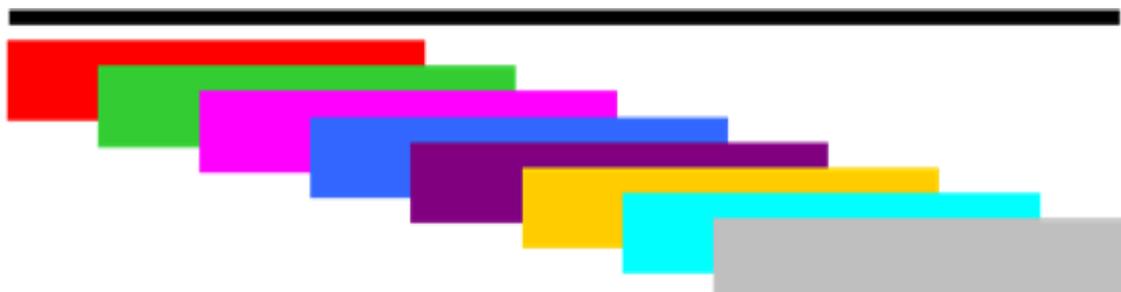
Subtipo NA	No. de secuencias	Subtipo NA	No. de secuencias	Subtipo NA	No. de secuencias
N1	360	N4	10	N7	28
N2	351	N5	14	N8	207
N3	40	N6	19	N9	14

### **1.2 Reducción de complejidad en el análisis**

Se realizaron grupos a 90% de identidad con el algoritmo CD-HIT. Los grupos se realizaron por segmento, y en el caso de HA y NA por subtipo (Tabla 8). Cada uno de los grupos contiene diferente número de secuencias.

### **1.3 Generación de oligonucleótidos de longitud 70**

Se generaron setentámeros (oligonucleótidos de longitud 70) a partir de las secuencias que tenían una identidad menor a 90% (Tabla 8). Estas se analizaron por ventanas de 70 nucleótidos con un desfase de 5 nucleótidos (Figura 9). Cada uno de los oligonucleótidos generados se guardaba con el nombre original de la secuencia, la posición de inicio y final de la secuencia de donde se obtuvo. (142\_H1N2\_Swine\_swine\_cloppenburg\_idt4777\_2005\_1698\_bp\_501\_570)



**Figura 9. Esquema que ejemplifica la generación de oligonucleótidos de longitud 70.**

Tabla 8. Número de secuencias iniciales, secuencias a 90% de identidad y setentámeros generados por segmento del genoma de influenza A. Para el caso de hemaglutinina y neuraminidasa se muestra para cada uno de los subtipos.

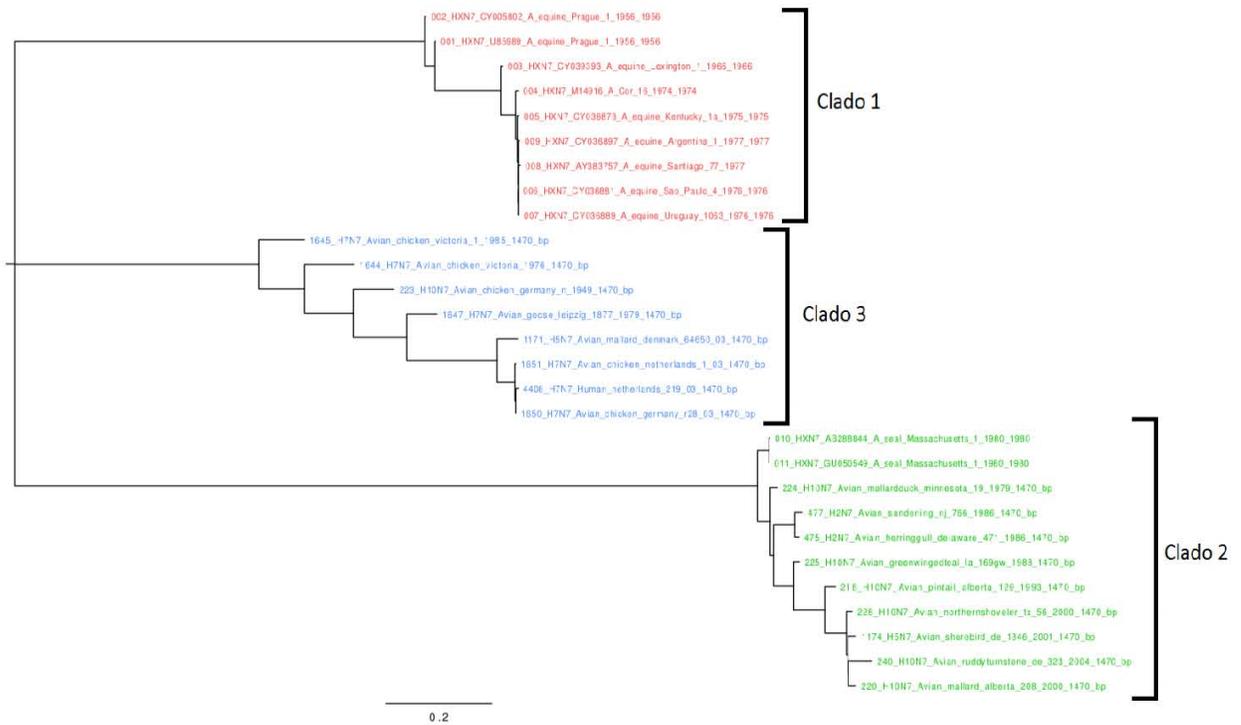
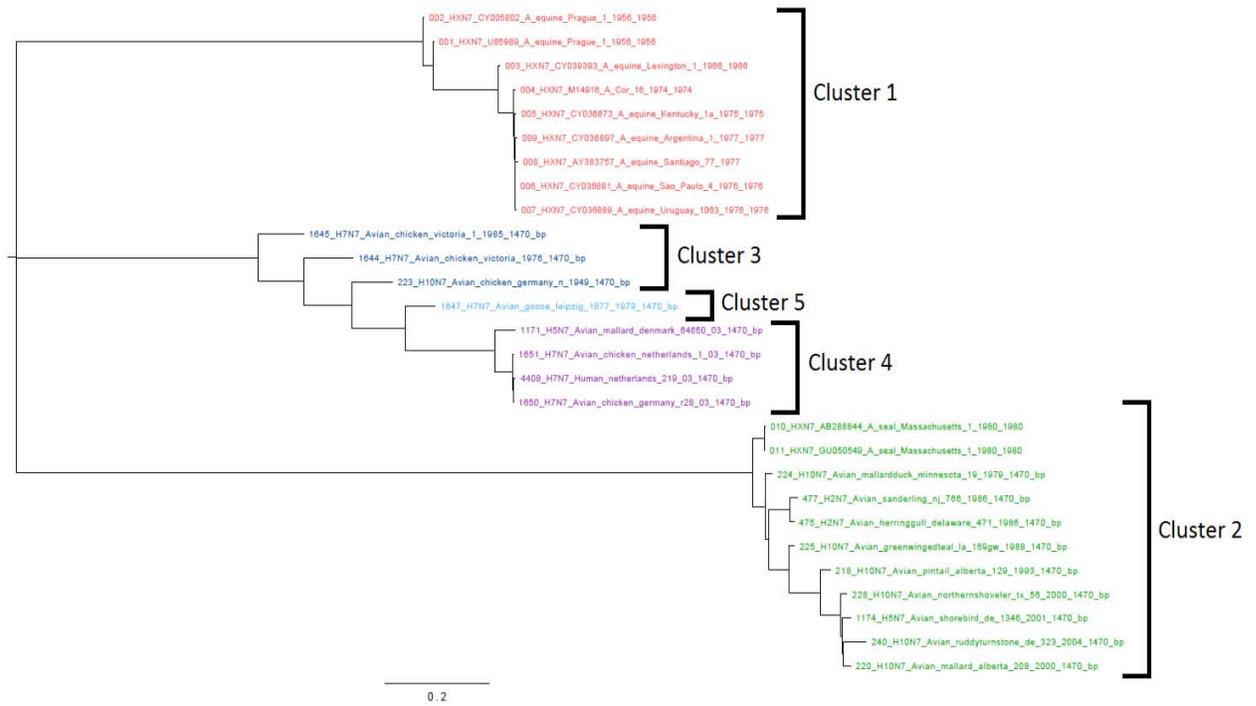
GEN	SECUENCIAS INICIALES	SECUENCIAS A 90% IDENTIDAD DE NUCLEOTIDOS	OLIGOS GENERADOS
PB2	926	21	9502
PB1	879	13	5880
PA	910	19	8435
NP	906	15	4447
M	1001	10	1890
NS	1146	19	3014

GEN	SECUENCIAS INICIALES	SECUENCIAS A 90% IDENTIDAD DE NUCLEOTIDOS	OLIGOS GENERADOS
H1	206	13	4324
H2	41	6	1974
H3	379	13	4373
H4	18	6	1998
H5	189	6	2006
H6	45	9	2988
H7	77	7	2333
H8	5	1	332
H9	61	6	1968
H10	8	2	658
H11	9	3	993
H12	8	2	662
H13	3	2	678
H14	2	1	342
H15	7	1	345
H16	8	4	1362

GEN	SECUENCIAS INICIALES	SECUENCIAS A 90% IDENTIDAD DE NUCLEOTIDOS	OLIGOS GENERADOS
N1	352	11	3145
N2	351	25	6684
N3	40	6	1618
N4	10	4	1087
N5	14	2	543
N6	19	4	1096
N7	28	5	1358
N8	207	9	2461
N9	14	2	557

#### **1.4 Definición de Clados**

Se definieron clados dentro de cada segmento mediante identidad de secuencia y posición dentro de los árboles filogenéticos. Primero se definieron los clados como las secuencias de los grupos a 90% de identidad. Con este corte no se pudieron definir clados en todos los segmentos, debido a que había un alto número de setentámeros no específicos, es decir, setentámeros que mapeaban en más de un clado, lo que ocasionaba que no todas las secuencias fuesen reconocidas en un único clado. En estos casos se tomó un porcentaje de identidad menor para definir los clados. Este proceso de selección fue iterativo hasta que se lograran seleccionar oligonucleótidos específicos que identificaran a todas las secuencias de su clado al menos 3 veces. En la tabla 9 se muestra el número de clados por segmento y el porcentaje de identidad al que se definió. Las figura 10 muestra un ejemplo de los clados mapeados en los árboles filogenéticos, en este caso para la neuraminidasa de subtipo 7.



**Figura 10.** Árboles filogenéticos pertenecientes al gen de neuraminidasa subtipo 7. Cada color indica un clado definido. En la parte superior se muestran los clusters de secuencias a 90% de identidad mapeados en el árbol. En la parte inferior se muestran los clados mapeados después del ajuste iterativo de reducción del porcentaje de identidad, para este caso 89%. Se hace notar que los clusters 3, 4 y 5 se agruparon en el clado 3.

### 1.5 Selección de oligonucleótidos

Los oligonucleótidos se clasificaron de acuerdo a su especificidad en el reconocimiento de secuencias dentro de su clado. La definición de clados fue un punto crítico para la selección debido a que cuanto más conservado es un gen, la diversidad en el universo de secuencias es menor y una definición de clado a 90% de identidad no fue posible. Se seleccionó el menor número de oligonucleótidos que reconociera todas las secuencias de su clado al menos tres veces. Una vez seleccionados los oligonucleótidos éstos se calificaron de acuerdo al número de secuencias que reconocían y al  $\Delta G$  teórico de hibridación (Tabla 9).

**Tabla 9. Número de clados por segmento del genoma de influenza A. Para el caso de HA y NA los clados por subtipo.**

GEN	PORCENTAJE DE IDENTIDAD	NÚMERO DE CLADOS	OLIGOS SELECCIONADOS
PB2	86%	2	13
PB1	85%	5	66
PA	81%	2	17
NP	84%	5	57
M	85%	4	40
NS	84%	3	54

GEN	PORCENTAJE DE IDENTIDAD	NÚMERO DE CLADOS	OLIGOS SELECCIONADOS
N1	85%	4	28
N2	82%	2	27
N3	87%	3	13
N4	89%	3	10
N5	90%	2	6
N6	87%	2	7
N7	87%	3	10
N8	86%	4	16
N9	90%	2	6

GEN	PORCENTAJE DE IDENTIDAD	NÚMERO DE CLADOS	OLIGOS SELECCIONADOS
H1	85%	4	21
H2	85%	2	10
H3	86%	5	24
H4	87%	4	14
H5	87%	2	11
H6	87%	3	12
H7	89%	4	15
H8	90%	1	4
H9	87%	3	18
H10	90%	2	7
H11	90%	2	6
H12	90%	2	6
H13	90%	2	6
H14	90%	1	4
H15	90%	1	4
H16	90%	4	12

### ***1.6 Nomenclatura de los oligonucleótidos seleccionados***

A cada uno de los oligonucleótidos seleccionados se les dio una nomenclatura para distinguir de qué segmento proviene y a que clado pertenece, con los cuales se distinguirá en cada uno de los resultados de hibridación el resultado para cada segmento. **FLUCHP\_S1\_PB2\_C1**, donde **FLUCHP** es el nombre del microarreglo y se encuentra en todos los oligonucleótidos. **S1\_PB2** denota el número del segmento y el nombre común del gen; para el caso de HA y NA se pone el subtipo al que pertenece (**S4\_H5**, **S6\_N3**) y por último **C1** denota al clado que pertenece. De esta forma **FLUCHP\_S4\_H5\_C0** sería HA subtipo 7 perteneciente al clado 0 y **FLUCHP\_S4\_N3\_C2** NA subtipo 3 perteneciente al clado 2.

## **1.7 Ensayos in silico**

Se realizaron alineamientos con el algoritmo BLAST (blastn) de los 544 setentámeros escogidos, contra diversas bases de datos. Los resultados de cada uno de los ensayos tuvieron un propósito en particular, siendo el caso de las cepas que se hibridaron las cuales tenían reportada la secuencia de nucleótidos la comparación con un segundo método, la secuenciación (1.7.1 Cepas de referencia). En otros casos se logró detectar rearrreglos génicos previamente reportados, así como observar el cambio de cepa circulante en humanos en los años 1957, 1968, 1977 y 2009.

### **1.7.1 Ensayos in silico con secuencias de referencia del virus de influenza A pertenecientes a distintos subtipos**

Los ensayos realizados con las 25 cepas de referencia del virus de influenza A donadas por el Dr. R. Webster, del Hospital St. Jude (Memphis, TN) sirvieron como base para la validación del microarreglo, ya que contábamos con un punto de comparación para cada uno de los segmentos. En todos los casos se reconoció correctamente el subtipo de hemaglutinina y neuraminidasa, además se identificó al clado que pertenecía cada uno de los 8 segmentos que componen el genoma viral. Estos resultados se utilizaron para comparar los resultados experimentales de las cepas de referencia.

### **1.7.2 Ensayos in silico con genomas de influenza A/H1N1**

Los 544 setentámeros se alinearon mediante el algoritmo BLAST contra una base de datos de 3901 genomas completos de influenza A/H1N1 provenientes de tres orígenes: humano, aviar y porcino. La mayor variabilidad se detectó en los genomas de virus aislados de cerdos y aves. De los genomas provenientes de cepas humanas se logró diferenciar aquellas que circularon antes y después de la pandemia de 2009. No hubo casos de reconocimiento inespecífico a nivel de clados en ninguno de los genomas analizados.

Se observaron casos interesantes en la composición de genes de algunos virus, dentro de las que destacan la cepa A/Iowa/CEID23/2005(H1N1) (Gray *et al.* 2007) y A/Saskatchewan/5131/2009(H1N1), A/Saskatchewan/5150/2009(H1N1) y A/Saskatchewan/5151/2009(H1N1) (Bastien *et al.*, 2010). La primera se observó un evento de zoonosis y fue caracterizada como una triple rearreglante humana/porcina/aviar que se transmitió de cerdo a humano. Las otras tres cepas son casos notorios de rearreglo genómico entre un virus triple rearreglante porcino H1N1 y un virus humano H1N1 estacional (A/Shanghai/LWS1/2009/2009/01/03(H1N1) en los genes de PB1, PA, HA y NA (Tabla 10).

**Tabla 10. Rearreglos detectados en virus de influenza A/H1N1 mediante ensayos *in silico* (naranja). Todos han sido casos reportados previamente en la literatura. Los resultados muestran el [nombre del gen]-[clado al que pertenecen], cada columna representando a un segmento/gen.**

La cepa A\_Iowa\_CEID23\_2005\_2005 es un triple rearreglante humano/porcino/aviar siendo la cepa estacional (ejemplificada por A/Auckland/619/2005/2005/07/01/(H1N1)) la circulante en ese tiempo.

Para el caso de las cepas A/Saskatchewan/5131/2009/06/19(H1N1), A/Saskatchewan/5350/2009/06/18(H1N1) y A/Saskatchewan/5351/2009/06/18(H1N1) son casos de rearreglo entre una cepa humana estacional (ejemplificada por A/Shanghai/LWS1/2009/01/03(H1N1)) y una cepa triple rearreglante porcino (ejemplificada por A/swine/MN/23506/2009(H1N1))

Los colores ejemplifican el cambio de clado para un segmento en específico y son meramente demostrativos.

Rearreglo detectado en 2005									
A/Auckland/619/2005/07/01(H1N1)	PB2-C2	PB1-C0	PA-C0	H1-C0	NP-C4	N1-C1	M-C0	NS-C1	2005
A_Iowa_CEID23_2005_2005	PB2-C2	PB1-C3	PA-C3	H1-C1	NP-C4	N1-C2	M-C0	NS-C1	2005
Rearreglos detectados en 2009									
A/Saskatchewan/5131/2009/06/19(H1N1)	PB2-C2	PB1-C3	PA-C3	H1-C0	NP-C4	N1-C1	M-C0	NS-C1	2009
A/Saskatchewan/5350/2009/06/18(H1N1)	PB2-C2	PB1-C3	PA-C3	H1-C0	NP-C4	N1-C1	M-C0	NS-C1	2009
A/Saskatchewan/5351/2009/06/18(H1N1)	PB2-C2	PB1-C3	PA-C3	H1-C0	NP-C4	N1-C1	M-C0	NS-C1	2009
A/swine/MN/23506/2009(H1N1)	PB2-C2	PB1-C3	PA-C3	H1-C1	NP-C4	N1-C2	M-C0	NS-C1	2010
A/Shanghai/LWS1/2009/01/03(H1N1)	PB2-C2	PB1-C0	PA-C0	H1-C0	NP-C4	N1-C1	M-C0	NS-C1	2009

### ***1.7.3 Ensayos in silico con genomas de influenza A/H3N2***

Los 544 setentámeros escogidos también se alinearon contra una base de datos con 2110 genomas completos de influenza A/H3N2 provenientes de 6 orígenes: humano, aves, puerco, perro, gato y visón. Se detectó una mayor variabilidad en los genomas colectados de cerdos y aves. Los genomas de virus aislados de humano tuvieron una variabilidad mínima, sólo se detectaron diferencias en los genes PB1 y PA en casos aislados. No hubo casos de reconocimiento inespecífico a nivel de clados en ninguno de los genomas analizados.

### ***1.7.4 Ensayos in silico con genomas completos de influenza A aviar, porcina y humana***

Los 544 setentámeros escogidos se alinearon con una base de datos de genomas de influenza A provenientes de humano, cerdo y aves. Estos resultados fueron utilizados para comparar los resultados obtenidos de hibridaciones con muestras porcinas y humanas. Se observó la variabilidad reportada en bases de datos de combinaciones de subtipos HA y NA, siendo 91 de las 144 combinaciones posibles. De igual manera se observaron las combinaciones de clados para los 8 segmentos, siendo en este caso solo de 248 del 1 260 000 posibles, presentándose el mayor número en virus de origen aviar como era lo esperado. Estos resultados fueron utilizados posteriormente para el análisis de los resultados de hibridación (Véase **5. Asignación de un origen animal a cada clado**, sección de **genomas únicos**)

## 2. Estadística y análisis de resultados

Una vez leídos los microarreglos en el lector, este nos entrega una tabla con las intensidades de fluorescencia para cada uno de las sondas del microarreglo. Con estos datos se determina la intensidad normalizada para cada sonda y a partir de ésta se toma un punto de corte con el que se delimita cuántas sondas serán analizadas.

Una vez con que se tiene que sondas serán analizadas, para cada clado se calcula la probabilidad de que el resultado observado sea al azar de acuerdo a una distribución hipergeométrica (Tabla 11). Este valor nos ayuda a dar soporte al resultado observado y a discernir en casos en donde para un mismo segmento se observan dos clados. En este caso, si la diferencia entre la probabilidad del azar es menor a un logaritmo, se toman ambos resultados para su posterior análisis.

**Tabla 11. Ejemplo de resultado de microarreglo. A partir de esta tabla se elige para cada segmento los resultados verdaderos positivos (negritas) de los falsos positivos (itálicas) comparando la probabilidad al azar dada por la distribución hipergeométrica.**

SONDA/CLADO	OLIGOS_TOTALES	OLIGOS_VISTOS	D_HIPER
<i>FLUCHP_S1_PB2_C0</i>	3	1	3.03E-01
<i>FLUCHP_S1_PB2_C1</i>	25	1	1.03E-01
<b>FLUCHP_S1_PB2_C2</b>	<b>26</b>	<b>13</b>	<b>4.56E-06</b>
<i>FLUCHP_S2_PB1_C0</i>	16	2	2.90E-01
<i>FLUCHP_S2_PB1_C1</i>	4	1	3.50E-01
<i>FLUCHP_S2_PB1_C2</i>	3	1	3.03E-01
<b>FLUCHP_S2_PB1_C3</b>	<b>17</b>	<b>16</b>	<b>5.76E-14</b>
<i>FLUCHP_S3_PA_C0</i>	7	1	3.98E-01
<b>FLUCHP_S3_PA_C3</b>	<b>34</b>	<b>18</b>	<b>1.60E-08</b>
<i>FLUCHP_S3_PA_C4</i>	8	1	3.93E-01
<b>FLUCHP_S4_H7_C0</b>	<b>4</b>	<b>4</b>	<b>3.09E-04</b>
<i>FLUCHP_S4_H7_C2</i>	5	1	3.79E-01
<b>FLUCHP_S5_NP_C0</b>	<b>35</b>	<b>23</b>	<b>1.82E-13</b>
<b>FLUCHP_S6_N3_C0</b>	<b>4</b>	<b>3</b>	<b>8.21E-03</b>
<i>FLUCHP_S6_N3_C2</i>	6	1	3.94E-01
<i>FLUCHP_S6_N8_C0</i>	4	1	3.50E-01
<b>FLUCHP_S7_M_C0</b>	<b>9</b>	<b>4</b>	<b>1.94E-02</b>
<b>FLUCHP_S8_NS_C0</b>	<b>5</b>	<b>4</b>	<b>1.34E-03</b>
<i>FLUCHP_S8_NS_C1</i>	12	1	3.31E-01

### 3. Ensayos validación

#### 3.1 Hibridación con cepas de referencia

Se hibridaron 17 cepas de referencia de influenza A como primer paso para la validación experimental del microarreglo. Estas cepas pertenecen a un grupo de 33 cepas de referencia donadas por el Dr. R.G. Webster. Las 17 cepas cubren todos los subtipos de hemaglutinina (H1-H16) y neuraminidasa (N1-N9) reportados para el virus de influenza A hasta marzo de 2012, cuando se describió un nuevo subtipo de hemaglutinina (H17) y un nuevo subtipo de neuraminidasa (N10) (Tong *et al.* 2012). Aunque no para todas las hibridaciones se detectó únicamente un clado para cada segmento, estas hibridaciones prueban que los oligonucleótidos diseñados discriminan entre los distintos subtipos de hemaglutinina y neuraminidasa, así como entre los diversos clados que se definieron para cada uno de los 8 segmentos del genoma viral. Todos los resultados de estas hibridaciones fueron comparados con las hibridaciones teóricas de los oligonucleótidos con las secuencias reportadas para cada virus. Salvo un caso, los resultados de hibridación son totalmente congruentes con los resultados teóricos.

Tabla 12. Resultados de las cepas de referencia de influenza A (Todos en Tabla anexa 1). En “Oligos Totales” se muestra el número de oligonucleótidos para cada clado y en “Oligos Vistos” los detectados en el microarreglo. La probabilidad está dada por la distribución hipergeométrica. El resultado que aparece en color negro quiere decir que no se logró identificar en el microarreglo. Los resultados fueron comparados con los ensayos *in silico* que se muestran en la izquierda con fondo gris. Aquellos en el ensayo *in silico* que no se muestran es por falta de secuencias de las bases de datos.

De lado izquierdo en gris se muestra el resultado de hibridación *in silico*, mostrándose en rojo aquellos que no concuerden con los resultados de experimentos de hibridación. Si se muestra en negro quiere decir que no hay secuencia para ese segmento particular.

<i>in silico</i>	Muestra St. Jude 05 – A_Singapore_1_1957_H2N2	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	12	2.46E-05
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	15	2.55E-12
FLUCHP_S3_PA_C4	FLUCHP_S3_PA_C4	FLUCHP_S3_PA_C4	8	8	7.06E-08
	FLUCHP_S4_H2_C0	FLUCHP_S4_H2_C0	5	4	1.25E-03
FLUCHP_S5_NP_C4	FLUCHP_S5_NP_C4	FLUCHP_S5_NP_C4	17	10	8.30E-06
FLUCHP_S6_N2_C0	FLUCHP_S6_N2_C0	FLUCHP_S6_N2_C0	12	8	2.07E-05
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	3	8.18E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	5	1.12E-02

**Muestra St. Jude 16 – A\_turkey\_Ontario\_6118\_1968\_H8N4**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	22	4.15E-13
FLUCHP_S2_PB1_C0	FLUCHP_S2_PB1_C0	16	11	1.74E-05
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	21	2.69E-08
FLUCHP_S4_H8_C0	FLUCHP_S4_H8_C0	4	4	1.38E-03
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	32	3.49E-21
FLUCHP_S6_N4_C0	FLUCHP_S6_N4_C0	3	3	7.21E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	7	2.21E-04
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	9	3.85E-05

**Muestra St. Jude 23 – IAV\_A\_w-t\_shearwater\_WAustralia\_2576\_1979\_H15N9**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	13	3.65E-06
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	14	1.96E-10
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	28	1.93E-20
FLUCHP_S4_H15_C0	FLUCHP_S4_H15_C0	4	4	3.47E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	14	6.15E-05
FLUCHP_S6_N9_C1	FLUCHP_S6_N9_C1	3	3	2.58E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	3	9.05E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	8	3.00E-05

**Muestra St. Jude 27 – IAV\_A\_tern\_South\_Africa\_1961\_H5N3**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	13	5.66E-06
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	17	1.49E-15
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	14	6.49E-05
FLUCHP_S4_H5_C0	FLUCHP_S4_H1_C0	6	5	2.90E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	27	7.46E-18
FLUCHP_S6_N3_C2	FLUCHP_S6_N3_C2	6	5	2.90E-04
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	4	2.40E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	10	1.27E-07

### 3.2 Hibridación de cepas de origen porcino

Se hibridaron 10 cepas porcinas provistas por el Dr. Iván Sánchez de la Facultad de Medicina Veterinaria y Zootecnia de la UNAM, colectadas en los estados de México, Veracruz, Michoacán y Querétaro durante 2010, como un primer ensayo para analizar muestras de campo. Estos virus fueron adaptados a crecer en embrión de pollo y por lo tanto estaban disponibles a títulos mayores que en las muestras clínicas, aun así de dos de las muestras no se pudo detectar ningún segmento. En dos de los casos no se logró detectar el segmento 7 (M1/M2) (Tabla 13). De las 8 muestras analizadas 6 fueron subtipo H1N1 y 2 de subtipo H3N2. Todas las muestras que tenían el mismo subtipo tanto de HA como de NA se catalogaron en el mismo grupo, ya que el clado de los 6 segmentos restantes también eran iguales. En el caso de la muestra A/swine/Mexico/Ver31/2010(H1N1) para el segmento 6 se obtuvieron 2 resultados de clados distintos, de los cuales no se pudo distinguir cual era el correcto con los datos obtenidos.

**Tabla 13. Resultados de las cepas de origen porcino (Todos en Tabla anexa 2). En todos los casos solo el subtipo era conocido. En "Oligos Totales" se muestra el número de oligonucleótidos para cada clado y en "Oligos Vistos" los detectados en el microarreglo. La probabilidad está dada por la distribución hipergeométrica. El resultado que aparece en color negro quiere decir que no se logró identificar en el microarreglo.**

El resultado de cada hibridación se muestra de la siguiente manera FLUCHP\_S1\_PB2\_C1, donde FLUCHP es el nombre del microarreglo y se encuentra en todos los oligonucleótidos. S1\_PB2 denota el número del segmento y el nombre común del gen, y para el caso de HA y NA se pone el subtipo que es (S4\_H5, S6\_N3). Por último C1 denota al clado que pertenece.

AW	Muestra	A/swine/Mexico/Mex50/2010(H3N2)	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
			FLUCHP_S1_PB2_C2	26	5	0.054616
			FLUCHP_S2_PB1_C3	17	10	2.36E-07
			FLUCHP_S3_PA_C3	34	5	0.110437
			FLUCHP_S4_H3_CO	8	3	0.025829
			FLUCHP_S5_NP_C4	17	6	0.002212
			FLUCHP_S6_N2_CO	12	9	4.60E-08
			FLUCHP_S7_M_CO	9	2	0.153835
			FLUCHP_S8_NS_C1	12	6	0.000257

AL	Muestra	A/swine/Mexico/Mex52/2010(H1N1)	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
			FLUCHP_S1_PB2_C2	26	15	3.54E-05
			FLUCHP_S2_PB1_C3	17	15	4.63E-09
			FLUCHP_S3_PA_C3	34	27	1.71E-13
			FLUCHP_S4_H1_C1	7	7	1.84E-05
			FLUCHP_S5_NP_C4	17	12	1.32E-05
			FLUCHP_S6_N1_C2	9	9	7.83E-07
			FLUCHP_S7_M_C0	9	9	7.83E-07
			FLUCHP_S8_NS_C1	12	12	6.59E-09

Tabla 14. Resultados del ensayo *in silico* de genomas de influenza A porcinos. Se muestran los genomas más representativos (Todos en Tabla anexa 6). El nombre de los virus mostrados solo representa un ejemplo de cada uno, siendo de los más actuales para cada caso.

AE	A/swine/NC/239108/2010-2010/04-21	PB2_C2	PB1_C3	PA_C3	H1_C0	NP_C4	N2_C0	M_C0	NS_C1
AI	A/swine/Guangdong/09/2009-2009/09-02	PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
AJ	A/swine/Miyazaki/2006-2006	PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N2_C0	M_C0	NS_C1
AL	A/swine/MN/23506/2009-2009/05/07	PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
AN	A/swine/Italy/116114/2010-2010/06-01	PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N2_C0	M_C0	NS_C1
AQ	A/swine/Shandong/327/2009-2009	PB2_C1	PB1_C3	PA_C3	H1_C3	NP_C0	N1_C3	M_C0	NS_C1
AW	A/swine/Penn/057108/2010-2010/09-20	PB2_C2	PB1_C3	PA_C3	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
AX	A/swine/Sichuan/01/2006-2006	PB2_C2	PB1_C3	PA_C4	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
BA	A/swine/North-Sumatra/UT6034/2007-2007	PB2_C1	PB1_C3	PA_C3	H5_C0	NP_C0	N1_C3	M_C0	NS_C1

### 3.3 Hibridación de una cepa influenza A pandémica 2009

Con el propósito de validar la detección de diversos clados en virus de influenza que tuvieran el mismo subtipo, se hibridó un virus colectado en el laboratorio durante la pandemia 2009 (A/Mexico/ib2/2009/(H1N1)) y el resultado de éste se comparó con un virus influenza A/H1N1 tipo estacional, como se muestra en la Tabla 15. En este caso se puede distinguir el cambio de clado en 4 de los 8 segmentos. Para el análisis del origen de los segmentos que cambiaron es necesario relacionar a cada clado con un origen: porcino de Norteamérica, porcino Euroasiático, humano, aviar. Como se mencionó anteriormente hasta este momento los experimentos son ciegos para evitar el sesgo ocasionado por la

anotación de los genomas reportados en las bases de datos, ya que para cada virus se le asigna como hospedero el organismo del que fue aislado.

Tabla 15. Comparación de un virus de influenza pandémico 2009 (A/Mexico/ibt2/2009/(H1N1)) con una cepa H1N1 estacional. Las flechas en rojo muestran los clados que difieren para un mismo segmento.

Influenza A/H1N1 Estacional				A/Mexico/ibt2/2009/(H1N1)			
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	26	14	1.22E-006	FLUCHP_S1_PB2_C2	26	10	4.42E-005
→ FLUCHP_S2_PB1_C0	16	8	5.42E-004	FLUCHP_S2_PB1_C3	17	11	2.09E-008
→ FLUCHP_S3_PA_C4	8	4	1.50E-002	FLUCHP_S3_PA_C3	34	17	4.20E-010
→ FLUCHP_S4_H1_C0	6	6	0.000007	FLUCHP_S4_H1_C1	7	6	0.000004
FLUCHP_S5_NP_C4	17	12	1.23E-007	FLUCHP_S5_NP_C4	17	8	5.07E-005
→ FLUCHP_S6_N1_C1	6	4	4.34E-003	FLUCHP_S6_N1_C3	10	2	1.84E-001
FLUCHP_S7_M_C0	9	7	2.67E-005	FLUCHP_S7_M_C0	9	5	5.96E-004
FLUCHP_S8_NS_C1	12	11	3.18E-009	FLUCHP_S8_NS_C1	12	6	3.26E-004

### 3.4 Hibridación con cepas de origen aviar

Se realizaron hibridaciones con 15 muestras de campo colectadas entre 2007 y 2009 en el estado de Sonora, provistas por el Dr. Jesús Hernández López y la Dra. Maricela Montalvo Corral, del Centro de Investigación en Alimentación y Desarrollo, A.C. Para estos casos se extrajo el material genético total y se realizó la amplificación del genoma completo del virus. Sólo 4 de las 15 muestras analizadas dieron un resultado positivo para al menos 6 de los 8 segmentos (Tabla 16), y una dio positivo para 4 segmentos (no se muestra). En dos de las muestras (769, 701), el microarreglo dio señal positiva para más de un clado en el segmento 4 dentro del mismo subtipo.

En estas muestras, se encontró una mayor diversidad en la combinación de subtipos de hemaglutinina y neuraminidasa: muestra 769, H9N2; muestra 829, H6N5; muestra 701, H11N3y; muestra 797, H5Nx, ya que el segmento 6 (NA) no pudo ser identificado en este virus. Estas muestras fueron secuenciadas en la plataforma Illumina, sin embargo los resultados aún no están disponibles para el cotejo de los resultados obtenidos en el microarreglo.

Tabla 16. Resultados de las cepas de origen aviar. En “Oligos Totales” se muestra el número de oligonucleótidos para cada clado y en “Oligos Vistos” los detectados en el microarreglo. La probabilidad está dada por la distribución hipergeométrica. El resultado que aparece en color negro quiere decir que no se logró identificar en el microarreglo.

<b>Sonora 769</b>			
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	26	13	1.64E-05
FLUCHP_S2_PB1_C3	17	12	2.30E-07
FLUCHP_S3_PA_C3	34	6	1.65E-01
FLUCHP_S4_H9_C1	7	5	1.08E-03
FLUCHP_S4_H9_C2	7	4	1.05E-02
FLUCHP_S5_NP_C0	35	30	9.26E-22
FLUCHP_S6_N2_C1	15	9	5.88E-05
FLUCHP_S7_M_C0	9	8	1.60E-06
FLUCHP_S8_NS_C1	12	12	7.77E-11

<b>Sonora 797</b>			
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S2_PB1_C3	17	14	5.21E-13
FLUCHP_S3_PA_C3	34	8	6.81E-03
FLUCHP_S4_H5_C1	5	3	6.33E-03
FLUCHP_S5_NP_C0	35	19	2.58E-12
FLUCHP_S7_M_C0	9	8	3.07E-08
FLUCHP_S8_NS_C1	12	8	1.30E-06

<b>Sonora 829</b>			
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	26	11	1.87E-04
FLUCHP_S2_PB1_C3	17	14	1.49E-10
FLUCHP_S3_PA_C3	34	17	1.49E-07
FLUCHP_S4_H6_C0	6	6	5.47E-06
FLUCHP_S5_NP_C0	35	20	4.28E-10
FLUCHP_S6_N5_C0	3	3	2.43E-03
FLUCHP_S7_M_C0	9	9	1.14E-08
FLUCHP_S8_NS_C1	12	12	2.18E-11

<b>Sonora 701</b>			
<b>Nombre</b>	<b>Oligos Totales</b>	<b>Oligos Vistos</b>	<b>Probabilidad de Azar</b>
FLUCHP_S1_PB2_C2	26	21	6.57E-10
FLUCHP_S2_PB1_C3	17	16	8.69E-10
FLUCHP_S3_PA_C3	34	28	2.28E-13
FLUCHP_S4_H11_C0	3	3	1.34E-02
FLUCHP_S4_H11_C1	3	2	1.30E-01
FLUCHP_S5_NP_C0	35	23	7.13E-08
FLUCHP_S6_N3_C0	4	4	3.16E-03
FLUCHP_S7_M_C0	9	9	2.15E-06
FLUCHP_S8_NS_C1	12	12	2.57E-08

### ***3.5 Hibridación con cepas de origen humano***

Se realizaron hibridaciones con virus de influenza presentes en muestras clínicas (hisopados nasofaríngeos) de pacientes de diversas regiones del país con enfermedad respiratoria (Tabla 4, Materiales y Métodos). El RNA viral presente en estas muestras se amplificó por RT-PCR utilizando dos estrategias: la primera consiste en utilizar cebadores con nonúmeros al azar para la reacción de transcripción reversa, al igual que para el PCR; en la segunda para sintetizar el cDNA se utilizan cebadores específicos que reconocen las regiones 5' y 3' no traducida de los segmentos del virus, las cuales están altamente conservadas en todos los segmentos, lo que permite hacer cDNA específico del genoma completo del virus mediante transcripción reversa (WGA). Después de esta amplificación dirigida se realiza una reacción de amplificación que utiliza cebadores con nonúmeros al azar. Para todos los casos la segunda estrategia mejoró la detección de los segmentos 7 y 8, por lo que ésta se utilizó en la amplificación de todas las muestras analizadas con excepción de la A/Mexico/ibt2/2009/(H1N1). Todas las muestras analizadas dieron subtipo H3N2 con los mismos patrones de clados 2/3/4/H3-0/4/N2-0/0/1 (segmentos 1 a 8) con excepción de la muestra A/Mexico/ibt2/2009/H1N1 la cual pertenece a un aislado de cepa pandémica 2009.

Tabla 17. Resultados de las cepas de origen humano (Todos en Tabla anexa 3). En “Oligos Totales” se muestra el número de oligonucleótidos para cada clado y en “Oligos Vistos” los detectados en el microarreglo. La probabilidad está dada por la distribución hipergeométrica. El resultado que aparece en color negro quiere decir que no se logró identificar en el microarreglo.

G	Muestra	A/Mexico/ibt2/(H1N1)	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
			FLUCHP_S1_PB2_C2	26	10	4.42E-05
			FLUCHP_S2_PB1_C3	17	11	2.09E-08
			FLUCHP_S3_PA_C3	34	17	4.20E-10
			FLUCHP_S4_H1_C1	7	6	3.92E-06
			FLUCHP_S5_NP_C4	17	8	5.07E-05
			FLUCHP_S6_N1_C3	10	2	1.84E-01
			FLUCHP_S7_M_C0	9	5	5.96E-04
			FLUCHP_S8_NS_C1	12	6	3.26E-04

K	Muestra	2662	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
			FLUCHP_S1_PB2_C2	26	8	2.02E-02
			FLUCHP_S2_PB1_C3	17	15	1.97E-11
			FLUCHP_S3_PA_C4	8	8	2.07E-07
			FLUCHP_S4_H3_C0	8	6	2.09E-04
			FLUCHP_S5_NP_C4	17	12	2.30E-07
			FLUCHP_S6_N2_C0	12	12	7.77E-11
			FLUCHP_S7_M_C0	9	7	3.85E-05
			FLUCHP_S8_NS_C1	12	9	4.07E-06

K	Muestra	Santos 121	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
			FLUCHP_S1_PB2_C2	26	15	3.28E-05
			FLUCHP_S2_PB1_C3	17	17	1.85E-12
			FLUCHP_S3_PA_C4	8	8	3.62E-06
			FLUCHP_S4_H3_C0	8	6	1.51E-03
			FLUCHP_S5_NP_C4	17	13	1.20E-06
			FLUCHP_S6_N2_C0	12	10	6.32E-06
			FLUCHP_S7_M_C0	9	8	2.60E-05
			FLUCHP_S8_NS_C1	12	12	6.11E-09

**Tabla 18. Resultados del ensayo *in silico* de genomas de influenza A humanos. Se muestran los genomas más representativos (Todos en Tabla anexa 5). El nombre de los virus mostrados solo representa un ejemplo de cada uno, siendo de los más actuales para cada caso.**

<b>A</b>	A/human/Wa/WRAIR1126P/2009-2009/01/09	PB2_C2	PB1_C0	PA_C4	H1_C0	NP_C4	N1_C1	M_C0	NS_C1
<b>E</b>	A/human/Maryland/12/1991-1991	PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
<b>G</b>	A/human/California/07/2009-2009/04/09	PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C3	M_C0	NS_C1
<b>H</b>	A/human/North-Carolina/1968-1968	PB2_C2	PB1_C3	PA_C4	H2_C0	NP_C4	N2_C0	M_C0	NS_C1
<b>K</b>	A/human/New-York/20343/2010-2010/06/16	PB2_C2	PB1_C3	PA_C4	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
<b>L</b>	A/human/Bangladesh/3233/2011-2011/03/09	PB2_C1	PB1_C3	PA_C3	H5_C0	NP_C0	N1_C3	M_C0	NS_C1

#### 4. Pruebas de sensibilidad del microarreglos

Para las pruebas de sensibilidad del microarreglo se hibridó una muestra en cuyo resultado se detectaran los 8 segmentos con un clado asignado a cada uno. De dicha muestra se realizaron diluciones seriadas hasta llegar a una dilución 1:10 000. Cada una de las diluciones se amplificó y preparó para la hibridación siguiendo el mismo protocolo. Hasta la dilución 1:1000 se pudo detectar la presencia de algunos segmentos. Para cada dilución la intensidad máxima de fluorescencia fue disminuyendo, hasta ser imperceptible en la dilución 1:10 000, en la que no se logró detectar ninguno de los segmentos.

**Tabla 19. Diluciones seriadas de la muestra FMVZ 50-2 512UH (H3N2). Para la dilución 1:10 000 no se logró detectar los segmentos. Se detecta un mayor número de segmentos en la dilución 1:100 que en la 1:10 ya que se normaliza cada microarreglo con la intensidad máxima del mismo.**

<b>Muestra FMVZ 50-2 512UH (H3N2)</b>			
<b>Resultado de hibridación</b>	<b>Oligos Totales</b>	<b>Oligos Vistos</b>	<b>Probabilidad de Azar</b>
FLUCHP_S1_PB2_C2	26	5	1.23E-01
FLUCHP_S2_PB1_C3	17	2	2.90E-01
FLUCHP_S3_PA_C2	5	2	1.05E-01
FLUCHP_S4_H3_C0	8	7	2.79E-06
FLUCHP_S5_NP_C4	17	8	3.69E-04
FLUCHP_S6_N2_C0	10	10	6.16E-10
FLUCHP_S7_M_C0	9	8	3.68E-07
FLUCHP_S8_NS_C1	12	11	7.46E-10

<b>Muestra FMVZ 50-2 512UH (H3N2) 1:10</b>			
<b>Resultado de hibridación</b>	<b>Oligos Totales</b>	<b>Oligos Vistos</b>	<b>Probabilidad de Azar</b>
FLUCHP_S2_PB1_C3	17	3	1.11E-02
FLUCHP_S4_H3_C0	8	4	3.92E-05
FLUCHP_S5_NP_C4	17	7	1.28E-07
FLUCHP_S6_N2_C0	10	4	1.12E-04
FLUCHP_S8_NS_C1	12	3	4.09E-03

<b>Muestra FMVZ 50-2 512UH (H3N2) 1:100</b>			
<b>Resultado de hibridación</b>	<b>Oligos Totales</b>	<b>Oligos Vistos</b>	<b>Probabilidad de Azar</b>
FLUCHP_S2_PB1_C3	17	5	3.17E-04
FLUCHP_S4_H3_C0	8	4	1.33E-04
FLUCHP_S5_NP_C4	17	7	1.23E-06
FLUCHP_S6_N2_C0	10	2	5.22E-02
FLUCHP_S7_M_C0	9	3	4.03E-03
FLUCHP_S8_NS_C1	12	5	4.83E-05

<b>Muestra FMVZ 50-2 512UH (H3N2) 1:1000</b>			
<b>Resultado de hibridación</b>	<b>Oligos Totales</b>	<b>Oligos Vistos</b>	<b>Probabilidad de Azar</b>
FLUCHP_S2_PB1_C3	17	3	7.53E-03
FLUCHP_S5_NP_C4	17	3	7.53E-03
FLUCHP_S6_N2_C0	10	2	2.40E-02
FLUCHP_S8_NS_C1	12	4	1.41E-04

## 5. Asignación de un origen animal a cada clado

Una vez obtenidos los resultados de hibridación era necesario asignar a cada segmento un origen de hospedero, es decir identificar de dónde provenían estas secuencias. Para esto se realizó un prueba de ensayo *in silico* con una colección de 70,708 secuencias de todos los orígenes animales reportados en la base de datos de Influenza de NCBI, exceptuando rearrreglantes, ambientales y los recientemente encontrados de murciélago (Zhou *et al.* 2012), ya que el diseño actual no toma en cuenta el virus H17N10.

Con esto cada conjunto de oligonucleótidos representantes de un clado reconocían a un número determinado de secuencias de un origen particular. Sin embargo, debido a la forma en que se asigna un hospedero al virus de influenza, en donde si un virus es aislado de un humano que se contagió de influenza aviar este es catalogado como humano, no tenemos valores absolutos para asignar un origen a cada clado, sino una combinación de dos o más orígenes, como se muestra en la Tabla anexa 1a.

Para realizar la asignación se consideró que la proporción de un origen particular que reconocían los oligonucleótidos de un clado dado fuera de 90% o mayor (Tabla anexa 1b). Con este corte se pudo asignar el origen del hospedero en 69 de los 88 clados que están definidos para los 8 segmentos del genoma de influenza A. Tomando cortes más relajados se obtiene una mayor cantidad de clados con origen asignado: con un corte de 85% restan 16 sin asignar, con 80% restan 12 y con 75% restan 10. Sin embargo tomar puntos de corte muy laxos nos daría una información errónea acerca de las secuencias, por lo que es necesario revisar otros parámetros como es qué tan parecidas son esas secuencias entre los distintos virus que infectan a las distintas especies animales, así como qué tanto una secuencia que inicialmente era de origen aviar pero lleva circulando 10 años en cerdo sigue perteneciendo a un origen u otro.

Con esta metodología para asignar los clados a una especie de hospedero en particular nos encontramos con dos problemas principalmente: 1) faltó por asignar un origen de hospedero a 19 de los 88 clados definidos y 2) algunos de los números de secuencias de virus aislados de humano están “inflados” ya que el humano es el organismo de mayor interés; además, debido a la creciente acumulación de secuencias de virus pandémicos a partir de 2009 existe una sobrerrepresentación de este genoma en las bases de datos. Para solucionar este problema y poder asignar un “origen de hospedero” a los clados encontrados para cada gen en una muestra, definimos los *genomas únicos* (Tablas anexa 5 – 9), representados por la combinación única de clados de cada uno de los segmentos del genoma de influenza presentes en un genoma particular. Este genoma único puede haber aparecido en un año en particular en una especie y haber seguido circulando o dejar de circular en la misma u otra especie. Tomemos el caso de virus de influenza A/H1N1

pdm2009, el cual circuló tanto en cerdos como en humanos. Este tiene el siguiente patrón genómico en humanos:

A/human/California/07/2009-2009/04/09



Y en cerdos:

A/swine/Taiwan/TD-2542/2010-2010/09/01



El patrón de clados que define a este virus es **|PB2\_C2|PB1\_C3|PA\_C3|H1\_C1|NP\_C4|N1\_C2|M\_C0|NS\_C1|**, el cual apareció por primera vez en 2009 y sigue circulando. Este patrón representaría a un genoma único o particular. Para el caso de virus colectados de humano encontramos 16 genomas únicos (Tabla anexa 5) en los datos de genomas completos de influenza A que hay en las bases de datos; para cerdo existen 33 (Tabla anexa 6) y para aves 248 (Tabla anexa 7).

Con el conjunto de genomas únicos de aves, cerdos y humanos se puede obtener el origen del virus observado en un resultado de hibridación, de tal forma que podemos conocer si esa combinación de segmentos ya se había reportado anteriormente, en qué hospedero, en qué año y en qué región del planeta.

En los resultados de hibridación realizados para validar este trabajo encontramos, de acuerdo a esta clasificación:

- Virus de cerdo triple rearreglante H1N1
- Virus de cerdo triple rearreglante H3N2
- Virus de humano H1N1 pdm 2009
- Virus de humano H3N2 estacional

De las muestras de origen aviar, en dos de ellas (769, 829) se reconocieron los 8 segmentos. Dichas combinaciones ya habían sido reportadas anteriormente en las bases de datos. En la muestra 769 se detectaron 2 clados para el segmento 4. Tomando las combinaciones con ambos clados, ambas combinaciones ya habían sido reportados en bases de datos.

## VII. Discusión

### 1. Definición de clados

No se pudieron definir clados para todos los genes utilizando una identidad de secuencia del 90%, debido a dos razones principales: la utilización del algoritmo CD-HIT para formar los grupos a 90% de identidad y la longitud de los oligonucleótidos. El algoritmo CD-HIT agrupa secuencias que tengan una identidad mayor a la que establecemos en los parámetros. El algoritmo elige una secuencia en la lista y compara con el resto, entonces agrupa todas aquellas que cumplan la condición de identidad y las elimina del análisis. Así la lista se reduce, y continúa tomando la siguiente secuencia de la nueva lista hasta que termina con todas las secuencias. Esto nos lleva al siguiente problema: consideremos que tenemos 3 secuencias, A, B y C. Las secuencias A y B son 90% idénticas, las secuencias B y C a su vez también son 90% idénticas, pero las secuencias A y C no lo son. El algoritmo agrupará las secuencias A-B y la secuencia C estará en un grupo distinto, sin embargo B-C era otro grupo posible, lo que causa problemas de inespecificidad de reconocimiento, ya que oligonucleótidos generados de C reconocerán a las secuencias B. La longitud de cada oligonucleótido es 70, con la cual es difícil tener grupos que se diferencien por 90% de identidad, más en genes conservados, tales como las polimerasas virales (segmentos 1, 2 y 3). Fueron pocos los segmentos que se pudieron agrupar a 90% de identidad (H10-H16 y N5), y de los que tuvieron un corte de identidad menor, ninguno llegó a ser menor a 80%, siendo el segmento 8 (NS) el más bajo, con 81%.

Una vez definidos los clados, se intentó aumentar la profundidad del análisis agregando subclados. Estos subclados subdividían a cada clado dependiendo de qué secuencias reconocía dentro del clado. Por ejemplo, de un total de 50 secuencias en un clado, un conjunto de estas (subclado 1) era reconocido por 5 oligonucleótidos y otro conjunto (subclado 2) era reconocido por 4 oligonucleótidos del mismo clado. Sin embargo no existía el problema si uno de los oligonucleótidos que reconocía secuencias del subclado 1, también reconocía del subclado 2, siempre que fueran pocos casos. Este análisis se realizó con ensayos *in silico* de hibridación mediante las  $\Delta G$  de hibridación. Se tomaban

como subclado aquellas que fueran reconocidas con un  $\Delta G$  teórico de hibridación menor a -95 kcal/mol. Una vez definidos se probaron con los experimentos de hibridación que se tenían. Ninguno de los subclados definidos en el análisis fue implementado, ya que los resultados no pudieron ser validados con los experimentos de hibridación. Sin embargo, gracias a este análisis se definieron clados nuevos, es decir, un clado se dividió en dos. Uno de estos casos fue en el segmento 3 (PA) en donde se separó el clado 0 que tenía secuencias porcinas y humanas en un clado porcino, clado 0, y un clado humano, denotado clado 4. Esto mismo sucedió con los segmentos 4 (H5, H11, H12, H13), 5 Y 6 (N5). Estos nuevos clados fueron probados por ensayos *in silico* y comparados con resultados de hibridación, los cuales fueron congruentes y por ello fueron implementados.

## **2. Ensayos *in silico***

Se observó una variabilidad mayor en mezcla de clados en los genomas de influenza A reportados de origen aviar en comparación con los reportados para cerdos y humanos, en los cuales sólo aparecen los subtipos de hemaglutinina 1, 2, 3, 5, 7 y 9 en humano y 1, 2, 3, 5, 9 y 11 en cerdo y los subtipo de neuraminidasa 1, 2, 3 y 7 en humano y 1, 2, 3, 6 y 8 en cerdo. Sin embargo los subtipos predominantes en humano (+90%) son H1, H2, H3, N1, N2 y N3, siendo los otros subtipos provenientes de eventos de zoonosis, los cuales fueron casos aislados; y para el caso de cerdo los subtipos predominantes son H1, H3, N1 y N2. Tanto en cerdos como humanos posibles eventos de rearreglo genómico pudieron ser identificados de forma frecuente en ensayos *in silico*. Para el caso de aves, la variabilidad es tal que no es posible identificar en los ensayos *in-silico* los rearreglos, ya que no se puede asignar con certeza cuales fueron las cepas paternas que dan origen a una cepa rearreglante.

La variabilidad de los genomas de los virus de influenza A/H1N1 y A/H3N2 de humano es muy baja (Tabla anexa 5). En el caso de influenza A/H3N2 la variabilidad es casi nula desde su origen en la pandemia de 1968. En el caso de influenza A/H1N1 se observó una clara identificación del cambio de cepa circulante en humanos después de la pandemia de

2009, así como rearrreglos entre una cepa triple rearrreglante porcina y una cepa A/H1N1 estacional.

En el caso de cerdos la variabilidad de combinación de genes de las cepas A/H1N1 y A/H3N2 es mayor que la de humano, encontrando aquí un número mayor de cepas (Tabla anexa 6), pudiéndose identificar cepas triple rearrreglantes, cepas pandémicas, además de rearrreglos que dieron como resultado virus triple rearrreglantes A/H1N2.

### **3. Validación experimental del microarreglo**

#### **3.1 Cepas de referencia**

Los resultados de experimentos de hibridación con cepas de referencia correlacionaron fuertemente con las predicciones computacionales, sin embargo, hay casos en donde la intensidad de la señal en los resultados del microarreglo no permitió clasificar a los 8 segmentos, en especial los segmentos 7 y 8, por la falta de sensibilidad del ensayo. Existen un caso, la muestra **St. Jude 08 – A/equine/Miami/1/1963(H3N8)** en la Tabla anexa 1, donde para el segmento 2 se determinan dos clados distintos: clado 1 (con 4 oligonucleótidos para identificarlo) y clado 2 (con 3 oligonucleótidos para identificarlo). Este resultado no se pudo definir ya que para ambos casos todos los oligonucleótidos son reconocidos. Este resultado se puede explicar debido a que ambos clados tienen secuencias de origen equino.

Para el caso particular del virus **A/tern/South\_Africa/1961/(H5N3)** (marcado con el número 27 en la Tabla 12) el análisis *in silico* del segmento 4, correspondiente a la hemaglutinina, la identificó como perteneciente al subtipo 5 clado 0, sin embargo la hibridación con el microarreglo indicó que era subtipo 1 clado 0. Esta discrepancia es bastante importante, ya que no fue un cambio en donde pudiese existir un entrecruzamiento de clados, sino un cambio de subtipo. Para verificar este caso se revisó el resultado experimental de un microarreglo de detección de patógenos virales que fue diseñado en el laboratorio, el cual también da como resultado hemaglutinina subtipo 1. Esto muy probablemente se deba a que la cepa que nos enviaron estaba marcada de manera equivocada. Sin embargo, sólo 1

de 8 segmentos fue el que da un resultado distinto al reportado, por lo que queda la duda de cuál es el subtipo real de la cepa enviada. Secuenciar esa muestra sería la mejor aproximación para encontrar la respuesta.

La comparación de las cepas de influenza A/H1N1 pandémica con la cepa estacional es un claro ejemplo de que el microarreglo puede discriminar a dos cepas del mismo subtipo que contengan segmentos de distinto origen, sin embargo, se requiere caracterizar un número mayor de cepas para confirmar esta conclusión.

### ***3.2 Muestras clínicas y de campo***

En la mayoría de los casos el método de amplificación con cebadores que tienen nonúmeros aleatorios no daba resultado de hibridación para los segmentos 6, 7 y 8. El problema se agravaba cuando la muestra clínica tenía muy poco material genético, por lo que en algunos casos no se obtenía resultado alguno. Para esto se realizó una amplificación del genoma completo (Whole Genome Amplification, WGA) del virus de influenza A con cebadores que son específicos para regiones conservadas en las regiones no traducidas del genoma de este virus. Con este paso de amplificación previo se aumenta la cantidad de material genético específico del virus de influenza A. Esta amplificación selectiva puede llevarse a cabo ya que el microarreglo diseñado y expuesto en este trabajo no está pensado para detección del virus de influenza A de una muestra desconocida, sino la determinación del origen de los segmentos que componen el genoma del virus, por lo que es necesario saber que la muestra es positiva para el virus de influenza A.

#### ***3.2.1 Muestras de origen porcino***

Todas las muestras analizadas cayeron en uno de dos grupos: A/H1N1 y A/H3N2. Dentro de virus del mismo subtipo todos tuvieron patrones iguales para los 6 segmentos restantes, por lo que se les puede asignar el mismo origen. Como se mencionó anteriormente, en uno de los resultados de hibridación (**A/swine/Mexico/Ver31/2010(H1N1)** en la Tabla anexa 2), para el segmento 6 resultaron dos clados identificados. Sin embargo, a diferencia del caso antes mencionado, los clados

identificados fueron el clado 4 (con 17 oligonucleótidos para identificarlo) y el clado 3, por lo que este resultado apoya la hipótesis de oligonucleótidos inespecíficos, mas no se puede concluir hasta realizar un análisis más detallado de esos oligonucleótidos. Los resultados de las hibridaciones fueron comparados con experimentos de secuenciación de las mismas muestras, y los resultados fueron congruentes. En dos de las muestras no se pudo identificar ninguno de los segmento del genoma del virus, tanto en los amplificados con nonúmeros al azar como el amplificado de genoma completo. Estos resultados concuerdan con los datos de secuenciación en donde hubo un número bajos de lecturas contra el genoma de influenza en esas muestras, aunque sí se pudo mapear parte de éste. Con estos resultados concluimos que la cantidad de material genético que se encontraba en la muestra era muy bajo, por debajo de los límites de detección del microarreglo.

### *3.2.2 Muestras de origen humano*

Seis de las ocho muestras analizadas resultaron ser influenza A/H3N2 estacional, una, la muestra A/Mexico/ibt2/2009/(H1N1) fue influenza A/H1N1 pandémica 2009 y una muestra fue influenza A/H1N1 estacional. Entre sí, todas las muestras A/H3N2 tuvieron patrones iguales para los 8 segmentos. Estas muestras también fueron secuenciadas, sin embargo aún no se ha realizado la comparación de resultados.

### *3.2.3 Muestras de origen aviar*

Para el caso de las muertas aviares, cada una dio un resultado distinto en cuanto a los subtipos de HA y NA. Sin embargo para los 6 segmentos restantes los clados fueron muy parecidos. Esto nos muestra una pequeña parte de la gran variedad de subtipos que probablemente circulan en aves en nuestro país, por lo que la hibridación de un mayor número de muestras de origen aviar puede dar un reflejo de la variabilidad que existe, no sólo para HA y NA, sino para los ocho segmentos en su conjunto. Estas muestras fueron secuenciadas, sin embargo aún no se ha realizado la comparación de resultados. Un análisis muy superficial de los datos de secuencia, mostró una coincidencia con los subtipos de HA y NA reportados en los experimentos de hibridación. De las muestras en las que en el microarreglo no se obtuvo resultado, los resultados de de secuenciación

masiva nos darán una idea de la calidad de la muestra y la cantidad de material genético específico de influenza A en cada una de ellas.

### *3.2.4 Comparación con el método de secuenciación masiva*

Los resultados de secuenciación, además de corroborar los resultados del microarreglo, nos dieron indicios del porqué no todas las muestras dieron resultado en el microarreglo. Cuando una de las muestras secuenciadas dio como resultado un bajo número de lecturas positivas a influenza A, los experimentos de hibridación dieron un resultado negativo; a su vez cuando los resultados de secuenciación masiva un número alto de lecturas específicas al virus de influenza A, en los experimentos de hibridación fue posible detectar los 8 segmentos de forma específica. Esto se puede explicar por la cantidad de material genético inicial que se encontró en la muestra, entre mayor sea la cantidad de material genético del virus, mejor será el resultado que se obtendrá en ambos experimentos. Una forma de visualizarlo es mediante electroforesis geles de agarosa después de la amplificación de genoma completo en donde el enriquecimiento de la muestra es dirigido al genoma viral. Una segunda forma de corroborar que la cantidad de material genético inicial afecta el resultado obtenido en los experimentos son las diluciones seriadas, en las cuales se observó la disminución de la intensidad de fluorescencia en el experimento y la detección de un menor número de segmentos en cada dilución.

## **4. Asignación de origen a los clados**

Una parte importante de este trabajo fue la asignación de un origen de hospedero a cada uno de los clados definidos. En un caso ideal cada uno de los clados tendría un origen único, y al realizar un experimento de hibridación se podría decir de dónde viene cada uno de los segmentos. Sin embargo esto no fue así. De los 88 clados definidos 19 no pudieron ser asignados con el primer método usado.

Aun así, saber el origen de cada segmento aislado nos daba sólo cierta información de la historia de cada virus analizado. Conoceríamos el origen, aunque cómo llegaron esos 8 segmentos a estar juntos sería una incógnita. Con la segunda estrategia de los genomas

únicos obtenemos mayor información. No sólo están incluidos los 88 clados, sino que para cada genoma hay registro de año, lugar y especie de origen donde se observó por primera vez esa combinación, además de los años y locaciones donde ha circulado. Además de que algunos de los genomas están asociados a referencias bibliográficas de donde se puede recabar más información.

Adicionalmente, con este análisis se puede observar la variabilidad genética encontrada en un grupo de hospederos, siendo las aves las que tienen mayor variabilidad, aunque haciendo los cálculos de saturación, aun pudieran ocurrir un gran número de nuevas combinaciones (1 260 000 en total).

## VIII. Conclusión y Perspectivas

En el presente trabajo se diseñó y validó una metodología para la detección y clasificación de los segmentos que conforman el genoma del virus de influenza A basado en microarreglos de DNA, el cual dio resultados muy favorables que podrían culminar en su implementación para estudios de epidemiología en humanos y cerdos, así como estudios de ecología en aves silvestres. Dicha metodología permite clasificar cada uno de los 8 segmentos en distintos grupos o clados definidos por identidad en secuencia de nucleótidos y análisis filogenéticos. Se han realizado una gran cantidad de trabajos basados en microarreglos de DNA para la identificación de influenza A, ya sea algunos subtipos de interés, o los 16 subtipos de HA y 9 de NA, pero ninguno toma en cuenta los 8 segmentos.

En el diseño se tomaron en cuenta dos propiedades de las secuencias, su identidad en nucleótidos, debido a la naturaleza de la metodología la cual se basa en hibridación de secuencias; y su agrupación en los clados dentro de una filogenia, con la que podemos trazar una línea de procedencia para cada uno de los segmentos del genoma de influenza A y así poder detectar rearrreglos genómicos.

Como parte del análisis una vez terminada la selección de los oligonucleótidos, se realizaron hibridaciones teóricas con diversas bases de datos, con las que se observó que los oligonucleótidos sólo reconocen secuencias del virus de influenza A. A su vez, con las hibridaciones computacionales tenemos un punto de comparación y predicción de las hibridaciones que se realizaron con cepas de las cuales se tiene reportada su secuencia de nucleótidos. Estos resultados sirven como una segunda validación del método presentado, ya que en ambos casos se obtiene el mismo resultado.

El diseño separa a los 8 segmentos del genoma de influenza A en 88 clados, con los que tenemos la posibilidad de detectar 1 260 000 genomas distintos de influenza A. Sin embargo, explorando la diversidad reportada en las bases de datos, esta no sobrepasa los 250, por lo que el diseño puede identificar un importante número de variantes aun no descritas.

Para la validación se hibridaron muestras provenientes de los tres hospederos que mayormente infecta el virus de influenza A: humano, cerdo y aves. Para los tres casos se logró identificar, en la mayoría de las muestras, a los 8 segmentos y a cada uno se le asignó únicamente un clado, con lo que se demuestra la eficacia del método descrito. En otros casos no se lograron detectar algunos segmentos, por lo que se realizaron ensayos para determinar los límites de detección del microarreglo y se llegó a la conclusión de que la calidad de la muestra es factor influyente en el resultado.

El diseño actual tiene aspectos importantes donde mejorar. Uno es el añadir a la cepa recientemente encontrada en marzo de 2012 de influenza A proveniente de murciélago la cual tiene un nuevo subtipo tanto de HA como de NA. Realizando una hibridación *in silico* de esta con el diseño actual no fue reconocida por ninguno de los oligonucleótidos. También es importante cambiar la estrategia con los segmentos 7 y 8, de los cuales solo se pudo dividir en dos clados cada uno. Con una estrategia distinta se podría obtener mayor información de estos. Si bien el diseño actual cumple con los objetivos propuestos, una segunda versión sería una gran oportunidad para mejorar en aquellos aspectos donde los resultados no fueron los deseados en términos de sensibilidad y especificidad.

De igual manera, es importante tener oligonucleótidos que nos identifiquen la sola presencia de influenza A en la muestra como control positivo, ya que solo se tiene oligonucleótidos para los clados en que se divide cada segmento. Este control nos sería de gran ayuda para los casos en donde no detectamos ninguno de los segmentos como control de la presencia del virus.

Otro aspecto importante es la implementación de esta metodología en estudios de influenza. Como se mencionó anteriormente, esta metodología se puede aplicar para epidemiología de influenza A en cerdos de granja y traspatio, con el fin de conocer la diversidad de influenza A que circula en cerdos mexicanos. Así mismo en estudios de ecología de influenza en aves silvestres acuáticas y de corral con el fin de conocer la diversidad natural circulante en México de este virus en su hospedero natural.

## IX. Bibliografía

Arias CF, Escalera-Zamudio M, Soto-Del Rio MD, Cobian-Güemes AG, Isa P, and Lopez S. *Molecular Anatomy of 2009 Influenza Virus A (H1N1)*. Archives of Medical Research 40, 2009.

Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, and Lipman D. *The Influenza Virus Resource at the National Center for Biotechnology Information*. Journal of Virology 82 (2), 2008.

Bastien N, Antonishyn NA, Brandt K, Wong CE, Chokani K, Vegh N, Horsman GB, Tyler S, Graham MR, Plummer FA, Levett PN and Li Y. *Human Infection with a Triple- Reassortant Swine Influenza A(H1N1) Virus Containing the Hemagglutinin and Neuraminidase Genes of Seasonal Influenza Virus*. The Journal of Infectious Diseases 201 (8), 2010

Bouviera NM, Palesea P. *The biology of influenza viruses*. Vaccine 26 Suppl. 4, 2008.

Chen EC, Miller SA, DeRisi JL, Chiu CY. *Using a Pan-Viral Microarray Assay (Virochip) to Screen Clinical Samples for Viral Pathogens*. J. Vis. Exp. (50), e2536, 2011.

Drake JW. *Rates of spontaneous mutation among RNA viruses*. ProcNatl Acad Sci USA 1993; 90:4171e4175.

Edgar RC. *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Research 32 (5), 2004.

Gall A, Hoffmann B, Harder T, Grund C, Höper D, and Beer M. *Design and Validation of a Microarray for Detection, Hemagglutinin Subtyping, and Pathotyping of Avian Influenza Viruses*. Journal Of Clinical Microbiology 47 (2), 2009.

Gardner SN, Jaing CJ, McLoughlin KS, Slezak TR. *A microbial detection array (MDA) for viral and bacterial detection*. BMC Genomics. 2010 Nov 25 ;11:668.

Gauthier M, Bonnaud B, Arsac M, Lavocat F, Maisetti J, Kay A, Simon F, Zoulim F, and Vernet G. *A Microarray For Hepatitis B Virus Genotyping And The Detection Of 994 Mutations Along The Genome*. Journal of Clinical Microbiology 48 (11), 2010.

Gray DM. *Derivation of Nearest-Neighbor Properties from Data on Nucleic Acid Oligomers. II. Thermodynamic Parameters of DNArRNA Hybrids and DNA Duplexes*. Biopolymers 42, 1997.

Gray GC, McCarthy T, Capuano AW, Setterquist SF, Olsen CW, Alavanja MC, and Lynch CF. *Swine Workers and Swine Influenza Virus Infections*. Emerging Infectious Diseases 13 (12), 2007.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W and Gascuel O. *New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0*. Systematic Biology 59 (3), 2010.

Hay AJ. *The action of adamantanamines against influenza A viruses: Inhibition of the M2 ion channel protein*. Semin. Virol. 3, 1992.

Horimoto T, Kawaoka Y. Influenza: lessons from past pandemics, warnings from current incidents. Nature Rev. Microbiol. 3, 591–600, 2005.

Helenius A. *Unpacking the incoming influenza virus*. Cell 69 (4), 1992.

Inglis SC, Lamb RA, Carroll AR, Mahy BWJ. *Polypeptides specified by the influenza virus genome. Evidence for eight distinct gene products specified by fowl plague virus*. Virology 74, 1976.

Lamb RA, Choppin PW. *Synthesis of influenza virus proteins in infected cells: Translation of viral polypeptides, including three P polypeptides, from RNA produced by primary transcription*. Virology 74, 1976.

Lamb RA and Krug RM. *Orthomyxoviridae: the viruses and their replication*. In: Knipe DM, Howley PM, editors. Fields virology. Philadelphia: Lippincott Williams & Wilkins, 2007.

Lenz O, Petrzik K, Spak J. *Investigating the sensitivity of a fluorescence-based microarray for the detection of fruit-tree viruses*. Journal of Virological Methods 148 (1-2), 2008.

Li J, Chen S, and Evans DH. *Typing and Subtyping Influenza Virus Using DNA Microarrays and Multiplex Reverse Transcriptase PCR*. Journal Of Clinical Microbiology 39 (2), 2001.

Li W, Godzik A. *Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences*. Bioinformatics 22, 2006.

Li X, Qi X, Miao L, Wang Y, Liu F, Gu H, Lu S, Yang Y and Liu F. *Detection and subtyping of influenza A virus based on a short oligonucleotide microarray*. Diagnostic Microbiology and Infectious Disease 65 (3), 2009.

Miller MB, Tang YW. *Basic Concepts of Microarrays and Potential Applications in Clinical Microbiology*. Clinical Microbiology Reviews 22 (4), 2009.

Ohta T. *Drift: Theoretical Aspects*. Encyclopedia Of Life Sciences, 2005.

Onodera K and Melcher U. *VirOligo: a database of virus-specific oligonucleotides*. Nucleic Acids Res. 2002, 30:203–204.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria 2012 (<http://www.R-project.org>)

Samji T. *Influenza A: Understanding the Viral Life Cycle*. Yale Journal Of Biology And Medicine 82, 2009.

Samuel CE. *Thematic Minireview Series: Toward a Structural Basis for Understanding Influenza Virus-Host Cell Interactions*. The Journal Of Biological Chemistry 285 (37), 2010.

Sengupta S, Onodera K, Lai A, and Melcher U. *Molecular Detection and Identification of Influenza Viruses by Oligonucleotide Microarray Hybridization*. Journal Of Clinical Microbiology 41 (10), 2003.

Sip M, Bystricka D, Kmoch S, Noskova L, Hartmannova H, Dedic P. *Detection of viral infections by an oligonucleotide microarray*. Journal of Virological Methods 165, 2010.

Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghvani J, Bhatt S, Peiris JSM, Guan Y and Rambaut A. *Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic*. Nature 459 (7250), 2009.

Taubenberger JK, Kash JC. *Influenza Virus Evolution, Host Adaptation, and Pandemic Formation*. Cell Host & Microbe 7, 2010

Taubenberger JK, Morens DM. *1918 Influenza: the Mother of All Pandemics*. Emerging Infectious Diseases 12 (1), 2006.

Taubenberger JK, Reid AH, Lourens RM, Wang R, Jin G and Fanning TG. *Characterization of the 1918 influenza virus polymerase genes*. Nature Vol 437, 889-893, 2005.

Tong S, Li Y, Rivaller P, Conrardy C, Castillo DA, Chen LM, Recuenco S, Ellison JA, Davis CT, York IA, Turmelle AS, Moran D, Rogers S, Shi M, Tao Y, Weil MR, Tang K, Rowe LA, Sammons S, Xu X, Frace M, Lindblade KA, Cox NJ, Anderson LJ, Rupprecht CE, Donis RO. *A distinct lineage of influenza A virus from bats*. Proc Natl Acad Sci USA. 2012 Mar 13; 109(11):4269-74.

Townsend MB, Dawson ED, Mehlmann M, Smagala JA, Dankbar DM, Moore CL, Smith CB, Cox NJ, Kuchta RD, and Rowlen KL. *Experimental Evaluation of the FluChip Diagnostic Microarray for Influenza Virus Surveillance*. Journal Of Clinical Microbiology 44 (8), 2006.

Urisman A, Fischer KF, Chiu CY, et al. *E-Predict: a computational strategy for species identification based on observed DNA microarray hybridization patterns*. Genome Biology 2005; 6:R78.

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 USA. 2002 Nov 26; 99(24)

## World Health Organization

2009

- [www.who.int/infectious-disease-report/idr99-spanish/pages/graph5.html](http://www.who.int/infectious-disease-report/idr99-spanish/pages/graph5.html)
- [www.who.int/infectious-disease-report/idr99-spanish/pages/ch2text.html#Anchor7](http://www.who.int/infectious-disease-report/idr99-spanish/pages/ch2text.html#Anchor7)
- [www.who.int/mediacentre/factsheets/fs211/en/](http://www.who.int/mediacentre/factsheets/fs211/en/)
- [www.who.int/vaccine\\_research/diseases/ari/en/index.html](http://www.who.int/vaccine_research/diseases/ari/en/index.html)

2013

- [www.who.int/influenza/surveillance\\_monitoring/updates/en/](http://www.who.int/influenza/surveillance_monitoring/updates/en/)
- [gamapserver.who.int/gareports/Default.aspx?ReportNo=6](http://gamapserver.who.int/gareports/Default.aspx?ReportNo=6)
- [gamapserver.who.int/mapLibrary/](http://gamapserver.who.int/mapLibrary/)

Zhou B, Donnelly ME, Scholes DT, George KS, Hatta M, Kawaoka Y, and Wentworth DE. *Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses*. *Journal of Virology*, Oct. 2009, p. 10309–10313

Todos los análisis computacionales fueron realizados en el cluster del Instituto de Biotecnología – UNAM.

## X. Anexos

**Tabla anexa 1. Resultados de las cepas de referencia de influenza A**

Muestra St. Jude 03 – A_Fort_Monmouth_1_1947_H1N1				
<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	6	2.60E-03
FLUCHP_S2_PB1_C0	FLUCHP_S2_PB1_C0	16	3	5.31E-02
FLUCHP_S3_PA_C4	FLUCHP_S3_PA_C4	8	3	8.49E-03
FLUCHP_S4_H1_C0	FLUCHP_S4_H1_C0	6	4	1.53E-04
FLUCHP_S5_NP_C4	FLUCHP_S5_NP_C4	17	6	2.30E-04
	FLUCHP_S5_NP_C0	35	8	5.23E-04
FLUCHP_S6_N1_C1	FLUCHP_S6_N1_C1	6	3	3.41E-03
FLUCHP_S7_M_C0				
FLUCHP_S8_NS_C1				

Muestra St. Jude 04 – A_New_Jersey_1976_H1N1				
<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	5	1.75E-02
FLUCHP_S2_PB1_C0	FLUCHP_S2_PB1_C0	16	7	2.24E-05
FLUCHP_S3_PA_C4	FLUCHP_S3_PA_C0	7	7	3.31E-09
FLUCHP_S4_H1_C1	FLUCHP_S4_H1_C1	7	6	3.78E-07
FLUCHP_S5_NP_C4	FLUCHP_S5_NP_C4	17	7	3.59E-05
FLUCHP_S6_N1_C2	FLUCHP_S6_N1_C2	9	2	9.52E-02
FLUCHP_S7_M_C0				
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	5	5.01E-04

Muestra St. Jude 05 – A_Singapore_1_1957_H2N2				
<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	12	2.46E-05
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	15	2.55E-12
FLUCHP_S3_PA_C4	FLUCHP_S3_PA_C4	8	8	7.06E-08
	FLUCHP_S4_H2_C0	5	4	1.25E-03
FLUCHP_S5_NP_C4	FLUCHP_S5_NP_C4	17	10	8.30E-06
FLUCHP_S6_N2_C0	FLUCHP_S6_N2_C0	12	8	2.07E-05
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	3	8.18E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	5	1.12E-02

**Muestra St. Jude 08 – A\_equine\_Miami\_1\_1963\_H3N8**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	7	1.72E-03
FLUCHP_S2_PB1_C1	FLUCHP_S2_PB1_C1	4	4	2.97E-05
	FLUCHP_S2_PB1_C2	3	3	4.12E-04
FLUCHP_S3_PA_C2	FLUCHP_S3_PA_C2	5	3	3.55E-03
FLUCHP_S4_H3_C1	FLUCHP_S4_H3_C1	4	3	1.53E-03
FLUCHP_S5_NP_C1	FLUCHP_S5_NP_C1	6	3	6.58E-03
FLUCHP_S6_N8_C0	FLUCHP_S6_N8_C0	4	4	2.97E-05
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	4	2.59E-03
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	7	5.58E-06

**Muestra St. Jude 10 – A\_duck\_Czechoslovakia\_1956\_H4N6**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	8	2.53E-04
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	7	1.08E-04
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	7	8.68E-03
FLUCHP_S4_H4_C2	FLUCHP_S4_H4_C2	4	3	1.61E-03
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	17	1.70E-11
FLUCHP_S6_N6_C0	FLUCHP_S6_N6_C0	4	4	3.18E-05
FLUCHP_S7_M_C0				
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	6	9.72E-05

**Muestra St. Jude 12 – A\_shearwater\_Australia\_1\_1972\_H6N5**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	16	3.23E-07
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	16	1.17E-11
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	29	4.39E-18
FLUCHP_S4_H6_C0	FLUCHP_S4_H6_C0	6	5	9.76E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	26	1.51E-13
FLUCHP_S6_N5_C1	FLUCHP_S6_N5_C1	3	2	8.25E-02
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	9	1.94E-07
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	8	2.53E-04

**Muestra St. Jude 15 – A\_ruddy\_turnstone\_NJ\_65\_1985\_H7N3**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	13	4.56E-06
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	16	5.76E-14
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	18	1.60E-08
FLUCHP_S4_H7_C0	FLUCHP_S4_H7_C0	4	4	3.09E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	23	1.82E-13
FLUCHP_S6_N3_C0	FLUCHP_S6_N3_C0	4	3	8.21E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	4	1.94E-02
FLUCHP_S8_NS_C0	FLUCHP_S8_NS_C0	5	4	1.35E-03

**Muestra St. Jude 16 – A\_turkey\_Ontario\_6118\_1968\_H8N4**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	22	4.15E-13
FLUCHP_S2_PB1_C0	FLUCHP_S2_PB1_C0	16	11	1.74E-05
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	21	2.69E-08
FLUCHP_S4_H8_C0	FLUCHP_S4_H8_C0	4	4	1.38E-03
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	32	3.49E-21
FLUCHP_S6_N4_C0	FLUCHP_S6_N4_C0	3	3	7.21E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	7	2.21E-04
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	9	3.85E-05

**Muestra St. Jude 17 – A\_turkey\_Wisconsin\_1\_1966\_H9N2**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	15	6.97E-07
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	17	9.92E-15
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	17	1.73E-06
FLUCHP_S4_H9_C0	FLUCHP_S4_H9_C0	4	4	6.14E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	25	5.18E-14
FLUCHP_S6_N2_C0	FLUCHP_S6_N2_C0	10	5	1.04E-02
FLUCHP_S6_N2_C1	FLUCHP_S6_N2_C1	15	6	1.63E-02
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	5	6.13E-03
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	6	4.98E-03

**Muestra St. Jude 18 – A\_chicken\_Germany\_n\_1949\_H10N7**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	22	1.31E-14
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	13	1.47E-07
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	24	5.58E-12
FLUCHP_S4_H10_C1	FLUCHP_S4_H10_C1	3	3	5.72E-03
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	32	2.32E-22
FLUCHP_S6_N7_C1	FLUCHP_S6_N7_C1	4	3	1.88E-02
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	8	6.91E-06
FLUCHP_S8_NS_C0	FLUCHP_S8_NS_C0	5	5	1.78E-04

**Muestra St. Jude 19 – A\_duck\_Memphis\_546\_1974\_H11N9**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	18	1.01E-09
FLUCHP_S2_PB1_C0	FLUCHP_S2_PB1_C0	16	12	2.11E-07
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	21	8.20E-10
FLUCHP_S4_H11_C0	FLUCHP_S4_H11_C0	3	2	6.64E-02
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	28	1.36E-17
FLUCHP_S6_N9_C0	FLUCHP_S6_N9_C0	3	3	4.24E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	4	3.59E-02
FLUCHP_S8_NS_C0	FLUCHP_S8_NS_C0	5	4	2.85E-03

**Muestra St. Jude 20 – IAV\_A\_duck\_Alberta\_60\_1976\_H12N5**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	FLUCHP_S1_PB2_C2	26	21	1.97E-12
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	15	3.71E-10
	FLUCHP_S3_PA_C3	34	27	1.69E-15
FLUCHP_S4_H12_C0	FLUCHP_S4_H12_C0	3	3	5.85E-03
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	30	4.92E-19
FLUCHP_S6_N5_C0	FLUCHP_S6_N5_C0	3	3	5.85E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	7	1.39E-04
FLUCHP_S8_NS_C0	FLUCHP_S8_NS_C0	5	5	1.84E-04

**Muestra St. Jude 21 – IAV\_A\_gull\_Maryland\_704\_1977\_H13N6**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	16	1.16E-09
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	14	6.26E-11
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	18	6.77E-09
FLUCHP_S4_H13_C0	FLUCHP_S4_H13_C0	3	3	2.03E-03
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	19	1.36E-09
FLUCHP_S6_N6_C0	FLUCHP_S6_N6_C0	4	4	2.52E-04
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	7	1.28E-05
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	3	1.34E-01

**Muestra St. Jude 22 – IAV\_A\_mallard\_duck\_Astrakhan\_263\_1982\_H14N5**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C1	25	11	4.72E-05
	FLUCHP_S2_PB1_C3	17	16	1.19E-14
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	20	2.71E-11
FLUCHP_S4_H14_C0	FLUCHP_S4_H14_C0	4	4	2.13E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	26	1.97E-18
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	3	7.00E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	4	3.85E-02

**Muestra St. Jude 23 – IAV\_A\_w-t\_shearwater\_WAustralia\_2576\_1979\_H15N9**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	13	3.65E-06
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	14	1.96E-10
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	28	1.93E-20
FLUCHP_S4_H15_C0	FLUCHP_S4_H15_C0	4	4	3.47E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	14	6.15E-05
FLUCHP_S6_N9_C1	FLUCHP_S6_N9_C1	3	3	2.58E-03
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	3	9.05E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	8	3.00E-05

**Muestra St. Jude 24 – IAV\_A\_shorebird\_Delaware\_168\_06\_H16N3**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	5	1.62E-02
FLUCHP_S3_PA_C3	FLUCHP_S2_PB1_C3	17	14	3.04E-15
	FLUCHP_S3_PA_C3	34	4	1.16E-01
FLUCHP_S4_H16_C2	FLUCHP_S4_H16_C2	3	3	2.78E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	13	4.81E-08
FLUCHP_S6_N3_C1	FLUCHP_S6_N3_C1	3	2	1.22E-02
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	4	1.60E-03
FLUCHP_S8_NS_C1				

**Muestra St. Jude 27 – IAV\_A\_tern\_South\_Africa\_1961\_H5N3**

<i>in silico</i>	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C1	FLUCHP_S1_PB2_C1	25	13	5.66E-06
FLUCHP_S2_PB1_C3	FLUCHP_S2_PB1_C3	17	17	1.49E-15
FLUCHP_S3_PA_C3	FLUCHP_S3_PA_C3	34	14	6.49E-05
FLUCHP_S4_H5_C0	FLUCHP_S4_H1_C0	6	5	2.90E-04
FLUCHP_S5_NP_C0	FLUCHP_S5_NP_C0	35	27	7.46E-18
FLUCHP_S6_N3_C2	FLUCHP_S6_N3_C2	6	5	2.90E-04
FLUCHP_S7_M_C0	FLUCHP_S7_M_C0	9	4	2.40E-02
FLUCHP_S8_NS_C1	FLUCHP_S8_NS_C1	12	10	1.27E-07

Tabla anexa 2. Resultados de las cepas de origen porcino

AW	Muestra	A/swine/Mexico/Mex50/2010(H3N2)	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
			FLUCHP_S1_PB2_C2	26	5	0.054616
			FLUCHP_S2_PB1_C3	17	10	2.36E-07
			FLUCHP_S3_PA_C3	34	5	0.110437
			FLUCHP_S4_H3_C0	8	3	0.025829
			FLUCHP_S5_NP_C4	17	6	0.002212
			FLUCHP_S6_N2_C0	12	9	4.60E-08
			FLUCHP_S7_M_C0	9	2	0.153835
			FLUCHP_S8_NS_C1	12	6	0.000257

AW	Muestra	A/swine/Mexico/Mex51/2010(H3N2)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	13	4.47E-07
	FLUCHP_S2_PB1_C3	17	6	0.005797
	FLUCHP_S3_PA_C3	34	17	5.62E-09
	FLUCHP_S4_H3_C0	8	5	0.000603
	FLUCHP_S5_NP_C4	17	10	1.60E-06
	FLUCHP_S6_N2_C0	12	12	1.66E-12
	FLUCHP_S8_NS_C1	12	4	0.028418

AL	Muestra	A/swine/Mexico/Mex52/2010(H1N1)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	15	3.54E-05
	FLUCHP_S2_PB1_C3	17	15	4.63E-09
	FLUCHP_S3_PA_C3	34	27	1.71E-13
	FLUCHP_S4_H1_C1	7	7	1.84E-05
	FLUCHP_S5_NP_C4	17	12	1.32E-05
	FLUCHP_S6_N1_C2	9	9	7.83E-07
	FLUCHP_S7_M_C0	9	9	7.83E-07
	FLUCHP_S8_NS_C1	12	12	6.59E-09

-	Muestra	A/swine/Mexico/Mich40/2010(H3N2)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	18	3.69E-12
	FLUCHP_S2_PB1_C3	17	9	3.44E-05
	FLUCHP_S3_PA_C3	34	23	4.41E-15
	FLUCHP_S4_H3_C0	8	6	5.66E-05
	FLUCHP_S5_NP_C4	17	10	3.42E-06
	FLUCHP_S6_N2_C0	12	7	0.000127

AL	Muestra	A/swine/Mexico/Qro32/2010(H1N1)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	4	0.092947
	FLUCHP_S2_PB1_C3	17	9	8.50E-07
	FLUCHP_S3_PA_C3	34	8	0.002481
	FLUCHP_S4_H1_C1	7	7	1.28E-08
	FLUCHP_S5_NP_C4	17	5	0.006257
	FLUCHP_S6_N1_C2	9	6	1.22E-05
	FLUCHP_S7_M_C0	9	3	0.02406
	FLUCHP_S8_NS_C1	12	3	0.049824

AL	Muestra	A/swine/Mexico/Ver29/2010(H1N1)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	3	0.20715
	FLUCHP_S2_PB1_C3	17	13	5.12E-12
	FLUCHP_S3_PA_C3	34	7	0.013
	FLUCHP_S4_H1_C1	7	7	2.06E-08
	FLUCHP_S5_NP_C4	17	6	0.001391
	FLUCHP_S6_N1_C2	9	7	6.35E-07
	FLUCHP_S7_M_C0	9	3	0.02831
	FLUCHP_S8_NS_C1	12	3	0.057626

AM/AL	Muestra	A/swine/Mexico/Ver31/2010(H1N1)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	5	0.047758
	FLUCHP_S2_PB1_C3	17	4	0.041203
	FLUCHP_S3_PA_C3	34	8	0.004861
	FLUCHP_S4_H1_C1	7	7	2.88E-08
	FLUCHP_S5_NP_C4	17	10	1.52E-07
	FLUCHP_S6_N1_C3	10	4	0.006696
	FLUCHP_S6_N1_C2	9	3	0.031723
	FLUCHP_S7_M_C0	9	4	0.004382
	FLUCHP_S8_NS_C1	12	10	7.64E-10

AL	Muestra	A/swine/Mexico/Ver37/2010(H1N1)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	16	1.03E-07
	FLUCHP_S2_PB1_C3	17	13	3.74E-08
	FLUCHP_S3_PA_C3	34	24	4.13E-13
	FLUCHP_S4_H1_C1	7	7	2.50E-06
	FLUCHP_S5_NP_C4	17	11	6.27E-06
	FLUCHP_S6_N1_C2	9	4	0.035029
	FLUCHP_S7_M_C0	9	5	0.006578
	FLUCHP_S8_NS_C1	12	5	0.024754

Tabla anexa 3. Resultados de las cepas de origen humano.

G	Muestra	A/Mexico/ibt2/(H1N1)		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	10	4.42E-05
	FLUCHP_S2_PB1_C3	17	11	2.09E-08
	FLUCHP_S3_PA_C3	34	17	4.20E-10
	FLUCHP_S4_H1_C1	7	6	3.92E-06
	FLUCHP_S5_NP_C4	17	8	5.07E-05
	FLUCHP_S6_N1_C3	10	2	1.84E-01
	FLUCHP_S7_M_C0	9	5	5.96E-04
	FLUCHP_S8_NS_C1	12	6	3.26E-04

-	Muestra	2664		
	Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
	FLUCHP_S1_PB2_C2	26	13	7.95E-06
	FLUCHP_S2_PB1_C3	17	15	7.34E-12
	FLUCHP_S3_PA_C4	8	7	6.42E-06
	FLUCHP_S4_H3_C0	8	5	1.85E-03
	FLUCHP_S5_NP_C4	17	12	1.10E-07
	FLUCHP_S6_N2_C0	12	8	3.45E-05
	FLUCHP_S7_M_C0	9	3	9.40E-02

Muestra		2662		
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar	
FLUCHP_S1_PB2_C2	26	8	2.02E-02	
FLUCHP_S2_PB1_C3	17	15	1.97E-11	
FLUCHP_S3_PA_C4	8	8	2.07E-07	
FLUCHP_S4_H3_C0	8	6	2.09E-04	
FLUCHP_S5_NP_C4	17	12	2.30E-07	
FLUCHP_S6_N2_C0	12	12	7.77E-11	
FLUCHP_S7_M_C0	9	7	3.85E-05	
FLUCHP_S8_NS_C1	12	9	4.07E-06	

Muestra		25216		
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar	
FLUCHP_S1_PB2_C2	26	11	2.64E-04	
FLUCHP_S2_PB1_C3	17	14	2.56E-10	
FLUCHP_S3_PA_C4	8	8	1.23E-07	
FLUCHP_S4_H3_C0	8	2	2.24E-01	
FLUCHP_S5_NP_C4	17	11	1.48E-06	
FLUCHP_S6_N2_C0	12	9	2.34E-06	
FLUCHP_S7_M_C0	9	6	3.76E-04	
FLUCHP_S8_NS_C1	12	8	3.45E-05	

Muestra Veracruz 1				
Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar	
FLUCHP_S1_PB2_C2	26	4	1.80E-01	
FLUCHP_S2_PB1_C3	17	5	2.75E-02	
FLUCHP_S3_PA_C4	8	7	1.51E-06	
FLUCHP_S4_H3_C0	8	5	7.11E-04	
FLUCHP_S5_NP_C4	17	9	2.37E-05	
FLUCHP_S6_N2_C0	12	9	3.74E-07	
FLUCHP_S7_M_C0	9	7	6.05E-06	
FLUCHP_S8_NS_C1	12	6	9.09E-04	

K

**Muestra Santos 33**

Resultado de hibridación	Oligos Totales	Oligos Vistos	Probabilidad de Azar
FLUCHP_S1_PB2_C2	26	7	1.76E-02
FLUCHP_S2_PB1_C3	17	8	2.34E-04
FLUCHP_S3_PA_C4	8	5	7.90E-04
FLUCHP_S4_H3_C0	8	4	7.74E-03
FLUCHP_S5_NP_C4	17	10	2.77E-06
FLUCHP_S6_N2_C0	12	9	4.58E-07
FLUCHP_S7_M_C0	9	7	7.08E-06
FLUCHP_S8_NS_C1	12	10	1.65E-08

Tabla anexa 4a. Definición de clado – origen en valores absolutos. Número de secuencias por hospedero que cada uno de los clados identificó.

Gen / Clado	aviar	porcino	humano	equino	canino	otro	TOTAL
FLUCHP_S1_PB2_C0	0	0	0	12	0	0	<b>12</b>
FLUCHP_S1_PB2_C1	15378	1398	1335	2	93	116	<b>18322</b>
FLUCHP_S1_PB2_C2	18490	2687	26954	494	112	76	<b>48813</b>
FLUCHP_S2_PB1_C0	100	1831	7292	0	0	0	<b>9223</b>
FLUCHP_S2_PB1_C1	0	8	0	380	115	4	<b>507</b>
FLUCHP_S2_PB1_C2	0	0	0	9	0	0	<b>9</b>
FLUCHP_S2_PB1_C3	44857	3684	37276	13	106	235	<b>86171</b>
FLUCHP_S3_PA_C0	48	1542	34	0	0	0	<b>1624</b>
FLUCHP_S3_PA_C1	0	0	0	9	0	0	<b>9</b>
FLUCHP_S3_PA_C2	0	10	0	404	109	4	<b>527</b>
FLUCHP_S3_PA_C3	50667	3856	15010	22	121	396	<b>70072</b>
FLUCHP_S3_PA_C4	0	119	15479	0	0	12	<b>15610</b>
FLUCHP_S4_H10_C0	324	0	0	0	0	0	<b>324</b>
FLUCHP_S4_H10_C1	21	0	0	0	0	2	<b>23</b>
FLUCHP_S4_H11_C0	314	0	0	0	0	0	<b>314</b>
FLUCHP_S4_H11_C1	85	2	0	0	0	0	<b>87</b>
FLUCHP_S4_H12_C0	146	0	0	0	0	0	<b>146</b>
FLUCHP_S4_H12_C1	22	0	0	0	0	0	<b>22</b>
FLUCHP_S4_H13_C0	50	0	0	0	0	0	<b>50</b>
FLUCHP_S4_H13_C1	17	0	0	0	0	3	<b>20</b>
FLUCHP_S4_H14_C0	12	0	0	0	0	0	<b>12</b>
FLUCHP_S4_H15_C0	32	0	0	0	0	0	<b>32</b>
FLUCHP_S4_H16_C0	8	0	0	0	0	0	<b>8</b>
FLUCHP_S4_H16_C1	26	0	0	0	0	0	<b>26</b>
FLUCHP_S4_H16_C2	11	0	0	0	0	0	<b>11</b>

FLUCHP_S4_H16_C3	3	0	0	0	0	0	<b>3</b>
FLUCHP_S4_H1_C0	5	740	7550	0	0	12	<b>8307</b>
FLUCHP_S4_H1_C1	62	3725	16291	0	0	75	<b>20153</b>
FLUCHP_S4_H1_C2	650	6	0	0	0	0	<b>656</b>
FLUCHP_S4_H1_C3	5	422	6	0	0	0	<b>433</b>
FLUCHP_S4_H2_C0	277	0	345	0	0	0	<b>622</b>
FLUCHP_S4_H2_C1	623	10	0	0	0	0	<b>633</b>
FLUCHP_S4_H3_C0	562	1138	15297	4	19	11	<b>17031</b>
FLUCHP_S4_H3_C1	0	8	0	489	16	12	<b>525</b>
FLUCHP_S4_H3_C2	737	0	0	0	0	0	<b>737</b>
FLUCHP_S4_H3_C3	24	0	0	0	0	0	<b>24</b>
FLUCHP_S4_H3_C4	600	5	0	0	0	12	<b>617</b>
FLUCHP_S4_H4_C0	884	8	0	0	0	4	<b>896</b>
FLUCHP_S4_H4_C1	10	0	0	0	0	0	<b>10</b>
FLUCHP_S4_H4_C2	168	0	0	0	0	3	<b>171</b>
FLUCHP_S4_H4_C3	30	0	0	0	0	0	<b>30</b>
FLUCHP_S4_H5_C0	9839	114	1226	5	0	150	<b>11334</b>
FLUCHP_S4_H5_C1	1126	0	0	0	0	0	<b>1126</b>
FLUCHP_S4_H6_C0	3242	3	0	0	0	0	<b>3245</b>
FLUCHP_S4_H6_C1	152	0	0	0	0	0	<b>152</b>
FLUCHP_S4_H6_C2	40	0	0	0	0	0	<b>40</b>
FLUCHP_S4_H7_C0	1164	0	8	0	0	12	<b>1184</b>
FLUCHP_S4_H7_C1	0	0	0	60	0	0	<b>60</b>
FLUCHP_S4_H7_C2	1077	0	20	0	0	0	<b>1097</b>
FLUCHP_S4_H7_C3	51	0	0	0	0	0	<b>51</b>
FLUCHP_S4_H8_C0	110	0	0	0	0	0	<b>110</b>
FLUCHP_S4_H9_C0	40	0	0	0	0	0	<b>40</b>
FLUCHP_S4_H9_C1	321	3	0	0	0	0	<b>324</b>
FLUCHP_S4_H9_C2	2533	65	9	4	4	4	<b>2619</b>
FLUCHP_S5_NP_C0	29732	1250	1375	20	42	215	<b>32634</b>
FLUCHP_S5_NP_C1	0	12	0	550	252	6	<b>820</b>
FLUCHP_S5_NP_C2	0	0	0	20	0	0	<b>20</b>
FLUCHP_S5_NP_C4	164	3736	39205	0	6	61	<b>43172</b>
FLUCHP_S5_NP_C5	6	0	0	0	0	0	<b>6</b>
FLUCHP_S6_N1_C0	32	0	0	0	0	0	<b>32</b>
FLUCHP_S6_N1_C1	6	26	5407	0	0	6	<b>5445</b>
FLUCHP_S6_N1_C2	69	2920	77	0	0	6	<b>3072</b>
FLUCHP_S6_N1_C3	14339	2346	16499	0	6	221	<b>33411</b>
FLUCHP_S6_N2_C0	121	1381	15154	0	0	22	<b>16678</b>
FLUCHP_S6_N2_C1	5770	100	14	1	48	3	<b>5936</b>
FLUCHP_S6_N3_C0	808	16	4	0	0	4	<b>832</b>
FLUCHP_S6_N3_C1	16	0	0	0	0	0	<b>16</b>

FLUCHP_S6_N3_C2	581	0	0	0	0	0	<b>581</b>
FLUCHP_S6_N4_C0	65	0	0	0	0	0	<b>65</b>
FLUCHP_S6_N4_C1	21	0	0	0	0	0	<b>21</b>
FLUCHP_S6_N4_C2	51	0	0	0	0	6	<b>57</b>
FLUCHP_S6_N5_C0	242	0	0	0	0	0	<b>242</b>
FLUCHP_S6_N5_C1	18	0	0	0	0	0	<b>18</b>
FLUCHP_S6_N6_C0	401	3	0	0	0	3	<b>407</b>
FLUCHP_S6_N6_C1	716	3	0	0	0	0	<b>719</b>
FLUCHP_S6_N7_C0	365	0	0	0	0	6	<b>371</b>
FLUCHP_S6_N7_C1	180	0	8	0	0	0	<b>188</b>
FLUCHP_S6_N7_C2	0	3	0	30	0	3	<b>36</b>
FLUCHP_S6_N8_C0	0	8	0	384	224	4	<b>620</b>
FLUCHP_S6_N8_C1	210	0	0	4	0	0	<b>214</b>
FLUCHP_S6_N8_C2	17	0	0	0	0	0	<b>17</b>
FLUCHP_S6_N8_C3	1078	0	0	0	0	0	<b>1078</b>
FLUCHP_S6_N9_C0	438	0	0	0	0	3	<b>441</b>
FLUCHP_S6_N9_C1	21	0	0	0	0	0	<b>21</b>
FLUCHP_S7_M_C0	31143	6329	25722	767	234	265	<b>64460</b>
FLUCHP_S7_M_C1	0	4	0	44	0	0	<b>48</b>
FLUCHP_S8_NS_C0	5135	10	5	5	0	5	<b>5160</b>
FLUCHP_S8_NS_C1	40790	6698	31817	1070	1006	510	<b>81891</b>

**Tabla anexa 4b. Definición de clado – origen en porcentajes. Porcentaje de secuencias por hospedero que cada uno de los clados identificó.**

Gen / Clado	%aviar	%porcino	%humano	%equino	%canino	%otro
FLUCHP_S1_PB2_C0	0	0	0	1	0	0
FLUCHP_S1_PB2_C1	0.8393	0.0763	0.0729	0.0001	0.0051	0.0063
FLUCHP_S1_PB2_C2	0.3788	0.055	0.5522	0.0101	0.0023	0.0016
FLUCHP_S2_PB1_C0	0.0108	0.1985	0.7906	0	0	0
FLUCHP_S2_PB1_C1	0	0.0158	0	0.7495	0.2268	0.0079
FLUCHP_S2_PB1_C2	0	0	0	1	0	0
FLUCHP_S2_PB1_C3	0.5206	0.0428	0.4326	0.0002	0.0012	0.0027
FLUCHP_S3_PA_C0	0.0296	0.9495	0.0209	0	0	0
FLUCHP_S3_PA_C1	0	0	0	1	0	0
FLUCHP_S3_PA_C2	0	0.019	0	0.7666	0.2068	0.0076
FLUCHP_S3_PA_C3	0.7231	0.055	0.2142	0.0003	0.0017	0.0057
FLUCHP_S3_PA_C4	0	0.0076	0.9916	0	0	0.0008
FLUCHP_S4_H1_C0	0.0006	0.0891	0.9089	0	0	0.0014

FLUCHP_S4_H1_C1	0.0031	0.1848	0.8084	0	0	0.0037
FLUCHP_S4_H1_C2	0.9909	0.0091	0	0	0	0
FLUCHP_S4_H1_C3	0.0115	0.9746	0.0139	0	0	0
FLUCHP_S4_H2_C0	0.4453	0	0.5547	0	0	0
FLUCHP_S4_H2_C1	0.9842	0.0158	0	0	0	0
FLUCHP_S4_H3_C0	0.033	0.0668	0.8982	0.0002	0.0011	0.0006
FLUCHP_S4_H3_C1	0	0.0152	0	0.9314	0.0305	0.0229
FLUCHP_S4_H3_C2	1	0	0	0	0	0
FLUCHP_S4_H3_C3	1	0	0	0	0	0
FLUCHP_S4_H3_C4	0.9724	0.0081	0	0	0	0.0194
FLUCHP_S4_H4_C0	0.9866	0.0089	0	0	0	0.0045
FLUCHP_S4_H4_C1	1	0	0	0	0	0
FLUCHP_S4_H4_C2	0.9825	0	0	0	0	0.0175
FLUCHP_S4_H4_C3	1	0	0	0	0	0
FLUCHP_S4_H5_C0	0.8681	0.0101	0.1082	0.0004	0	0.0132
FLUCHP_S4_H5_C1	1	0	0	0	0	0
FLUCHP_S4_H6_C0	0.9991	0.0009	0	0	0	0
FLUCHP_S4_H6_C1	1	0	0	0	0	0
FLUCHP_S4_H6_C2	1	0	0	0	0	0
FLUCHP_S4_H7_C0	0.9831	0	0.0068	0	0	0.0101
FLUCHP_S4_H7_C1	0	0	0	1	0	0
FLUCHP_S4_H7_C2	0.9818	0	0.0182	0	0	0
FLUCHP_S4_H7_C3	1	0	0	0	0	0
FLUCHP_S4_H8_C0	1	0	0	0	0	0
FLUCHP_S4_H9_C0	1	0	0	0	0	0
FLUCHP_S4_H9_C1	0.9907	0.0093	0	0	0	0
FLUCHP_S4_H9_C2	0.9672	0.0248	0.0034	0.0015	0.0015	0.0015
FLUCHP_S4_H10_C0	1	0	0	0	0	0
FLUCHP_S4_H10_C1	0.913	0	0	0	0	0.087
FLUCHP_S4_H11_C0	1	0	0	0	0	0
FLUCHP_S4_H11_C1	0.977	0.023	0	0	0	0
FLUCHP_S4_H12_C0	1	0	0	0	0	0
FLUCHP_S4_H12_C1	1	0	0	0	0	0
FLUCHP_S4_H13_C0	1	0	0	0	0	0
FLUCHP_S4_H13_C1	0.85	0	0	0	0	0.15
FLUCHP_S4_H14_C0	1	0	0	0	0	0
FLUCHP_S4_H15_C0	1	0	0	0	0	0
FLUCHP_S4_H16_C0	1	0	0	0	0	0
FLUCHP_S4_H16_C1	1	0	0	0	0	0
FLUCHP_S4_H16_C2	1	0	0	0	0	0
FLUCHP_S4_H16_C3	1	0	0	0	0	0
FLUCHP_S5_NP_C0	0.9111	0.0383	0.0421	0.0006	0.0013	0.0066

FLUCHP_S5_NP_C1	0	0.0146	0	0.6707	0.3073	0.0073
FLUCHP_S5_NP_C2	0	0	0	1	0	0
FLUCHP_S5_NP_C4	0.0038	0.0865	0.9081	0	0.0001	0.0014
FLUCHP_S5_NP_C5	1	0	0	0	0	0
FLUCHP_S6_N1_C0	1	0	0	0	0	0
FLUCHP_S6_N1_C1	0.0011	0.0048	0.993	0	0	0.0011
FLUCHP_S6_N1_C2	0.0225	0.9505	0.0251	0	0	0.002
FLUCHP_S6_N1_C3	0.4292	0.0702	0.4938	0	0.0002	0.0066
FLUCHP_S6_N2_C0	0.0073	0.0828	0.9086	0	0	0.0013
FLUCHP_S6_N2_C1	0.972	0.0168	0.0024	0.0002	0.0081	0.0005
FLUCHP_S6_N3_C0	0.9712	0.0192	0.0048	0	0	0.0048
FLUCHP_S6_N3_C1	1	0	0	0	0	0
FLUCHP_S6_N3_C2	1	0	0	0	0	0
FLUCHP_S6_N4_C0	1	0	0	0	0	0
FLUCHP_S6_N4_C1	1	0	0	0	0	0
FLUCHP_S6_N4_C2	0.8947	0	0	0	0	0.1053
FLUCHP_S6_N5_C0	1	0	0	0	0	0
FLUCHP_S6_N5_C1	1	0	0	0	0	0
FLUCHP_S6_N6_C0	0.9853	0.0074	0	0	0	0.0074
FLUCHP_S6_N6_C1	0.9958	0.0042	0	0	0	0
FLUCHP_S6_N7_C0	0.9838	0	0	0	0	0.0162
FLUCHP_S6_N7_C1	0.9574	0	0.0426	0	0	0
FLUCHP_S6_N7_C2	0	0.0833	0	0.8333	0	0.0833
FLUCHP_S6_N8_C0	0	0.0129	0	0.6194	0.3613	0.0065
FLUCHP_S6_N8_C1	0.9813	0	0	0.0187	0	0
FLUCHP_S6_N8_C2	1	0	0	0	0	0
FLUCHP_S6_N8_C3	1	0	0	0	0	0
FLUCHP_S6_N9_C0	0.9932	0	0	0	0	0.0068
FLUCHP_S6_N9_C1	1	0	0	0	0	0
FLUCHP_S7_M_C0	0.4831	0.0982	0.399	0.0119	0.0036	0.0041
FLUCHP_S7_M_C1	0	0.0833	0	0.9167	0	0
FLUCHP_S8_NS_C0	0.9952	0.0019	0.001	0.001	0	0.001
FLUCHP_S8_NS_C1	0.4981	0.0818	0.3885	0.0131	0.0123	0.0062

**Tabla anexa 4c Definición de clado – origen en valores binarios. Aquellos marcados con un “1” son los clados que tienen un origen asignado por tener 90% o más de secuencias reconocidas de un mismo origen**

Gen / Clado	Aviar	Porcino	Humano	Equino	Canino	Otro	Origen definido
FLUCHP_S1_PB2_C0	0	0	0	1	0	0	1
FLUCHP_S1_PB2_C1	0	0	0	0	0	0	0
FLUCHP_S1_PB2_C2	0	0	0	0	0	0	0
FLUCHP_S2_PB1_C0	0	0	0	0	0	0	0
FLUCHP_S2_PB1_C1	0	0	0	0	0	0	0
FLUCHP_S2_PB1_C2	0	0	0	1	0	0	1
FLUCHP_S2_PB1_C3	0	0	0	0	0	0	0
FLUCHP_S3_PA_C0	0	1	0	0	0	0	1
FLUCHP_S3_PA_C1	0	0	0	1	0	0	1
FLUCHP_S3_PA_C2	0	0	0	0	0	0	0
FLUCHP_S3_PA_C3	0	0	0	0	0	0	0
FLUCHP_S3_PA_C4	0	0	1	0	0	0	1
FLUCHP_S4_H1_C0	0	0	1	0	0	0	1
FLUCHP_S4_H1_C1	0	0	0	0	0	0	0
FLUCHP_S4_H1_C2	1	0	0	0	0	0	1
FLUCHP_S4_H1_C3	0	1	0	0	0	0	1
FLUCHP_S4_H2_C0	0	0	0	0	0	0	0
FLUCHP_S4_H2_C1	1	0	0	0	0	0	1
FLUCHP_S4_H3_C0	0	0	0	0	0	0	0
FLUCHP_S4_H3_C1	0	0	0	1	0	0	1
FLUCHP_S4_H3_C2	1	0	0	0	0	0	1
FLUCHP_S4_H3_C3	1	0	0	0	0	0	1
FLUCHP_S4_H3_C4	1	0	0	0	0	0	1
FLUCHP_S4_H4_C0	1	0	0	0	0	0	1
FLUCHP_S4_H4_C1	1	0	0	0	0	0	1
FLUCHP_S4_H4_C2	1	0	0	0	0	0	1
FLUCHP_S4_H4_C3	1	0	0	0	0	0	1
FLUCHP_S4_H5_C0	0	0	0	0	0	0	0
FLUCHP_S4_H5_C1	1	0	0	0	0	0	1
FLUCHP_S4_H6_C0	1	0	0	0	0	0	1
FLUCHP_S4_H6_C1	1	0	0	0	0	0	1
FLUCHP_S4_H6_C2	1	0	0	0	0	0	1
FLUCHP_S4_H7_C0	1	0	0	0	0	0	1
FLUCHP_S4_H7_C1	0	0	0	1	0	0	1
FLUCHP_S4_H7_C2	1	0	0	0	0	0	1
FLUCHP_S4_H7_C3	1	0	0	0	0	0	1
FLUCHP_S4_H8_C0	1	0	0	0	0	0	1
FLUCHP_S4_H9_C0	1	0	0	0	0	0	1

FLUCHP_S4_H9_C1	1	0	0	0	0	0	1
FLUCHP_S4_H9_C2	1	0	0	0	0	0	1
FLUCHP_S4_H10_C0	1	0	0	0	0	0	1
FLUCHP_S4_H10_C1	1	0	0	0	0	0	1
FLUCHP_S4_H11_C0	1	0	0	0	0	0	1
FLUCHP_S4_H11_C1	1	0	0	0	0	0	1
FLUCHP_S4_H12_C0	1	0	0	0	0	0	1
FLUCHP_S4_H12_C1	1	0	0	0	0	0	1
FLUCHP_S4_H13_C0	1	0	0	0	0	0	1
FLUCHP_S4_H13_C1	0	0	0	0	0	0	0
FLUCHP_S4_H14_C0	1	0	0	0	0	0	1
FLUCHP_S4_H15_C0	1	0	0	0	0	0	1
FLUCHP_S4_H16_C0	1	0	0	0	0	0	1
FLUCHP_S4_H16_C1	1	0	0	0	0	0	1
FLUCHP_S4_H16_C2	1	0	0	0	0	0	1
FLUCHP_S4_H16_C3	1	0	0	0	0	0	1
FLUCHP_S5_NP_C0	1	0	0	0	0	0	1
FLUCHP_S5_NP_C1	0	0	0	0	0	0	0
FLUCHP_S5_NP_C2	0	0	0	1	0	0	1
FLUCHP_S5_NP_C4	0	0	1	0	0	0	1
FLUCHP_S5_NP_C5	1	0	0	0	0	0	1
FLUCHP_S6_N1_C0	1	0	0	0	0	0	1
FLUCHP_S6_N1_C1	0	0	1	0	0	0	1
FLUCHP_S6_N1_C2	0	1	0	0	0	0	1
FLUCHP_S6_N1_C3	0	0	0	0	0	0	0
FLUCHP_S6_N2_C0	0	0	1	0	0	0	1
FLUCHP_S6_N2_C1	1	0	0	0	0	0	1
FLUCHP_S6_N3_C0	1	0	0	0	0	0	1
FLUCHP_S6_N3_C1	1	0	0	0	0	0	1
FLUCHP_S6_N3_C2	1	0	0	0	0	0	1
FLUCHP_S6_N4_C0	1	0	0	0	0	0	1
FLUCHP_S6_N4_C1	1	0	0	0	0	0	1
FLUCHP_S6_N4_C2	0	0	0	0	0	0	0
FLUCHP_S6_N5_C0	1	0	0	0	0	0	1
FLUCHP_S6_N5_C1	1	0	0	0	0	0	1
FLUCHP_S6_N6_C0	1	0	0	0	0	0	1
FLUCHP_S6_N6_C1	1	0	0	0	0	0	1
FLUCHP_S6_N7_C0	1	0	0	0	0	0	1
FLUCHP_S6_N7_C1	1	0	0	0	0	0	1
FLUCHP_S6_N7_C2	0	0	0	0	0	0	0
FLUCHP_S6_N8_C0	0	0	0	0	0	0	0
FLUCHP_S6_N8_C1	1	0	0	0	0	0	1

FLUCHP_S6_N8_C2	1	0	0	0	0	0	1
FLUCHP_S6_N8_C3	1	0	0	0	0	0	1
FLUCHP_S6_N9_C0	1	0	0	0	0	0	1
FLUCHP_S6_N9_C1	1	0	0	0	0	0	1
FLUCHP_S7_M_C0	0	0	0	0	0	0	0
FLUCHP_S7_M_C1	0	0	0	1	0	0	1
FLUCHP_S8_NS_C0	1	0	0	0	0	0	1
FLUCHP_S8_NS_C1	0	0	0	0	0	0	0

Total de Clados 88

Total de Clados con origen definido 69

Número de clados no asignados 19

**Tabla anexa 5. Genomas únicos de virus de influenza A cuyo hospedero es humano. Estos fueron escogidos en base a ensayos de hibridación *in silico* contra genomas de influenza A provenientes de humano con los oligonucleótidos que fueron impresos en el microarreglo.**

Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5	Segmento 6	Segmento 7	Segmento 8
PB2_C2	PB1_C0	PA_C4	H1_C0	NP_C4	N1_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C0	NP_C4	N1_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C4	H1_C0	NP_C4	N1_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C4	H1_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C0	PA_C4	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C4	H2_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C0	PA_C4	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C4	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_C0	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_C0	N3_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_C0	N7_C1	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C2	NP_C0	N2_C1	M_C0	NS_C1

**Tabla anexa 6. Genomas únicos de virus de influenza A cuyo hospedero es cerdo. Estos fueron escogidos en base a ensayos de hibridación *in silico* contra genomas de influenza A provenientes de cerdo con los oligonucleótidos que fueron impresos en el microarreglo.**

Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5	Segmento 6	Segmento 7	Segmento 8
PB2_C1	PB1_C3	PA_C3	H1_C0	NP_C0	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C0	PA_C4	H1_C0	NP_C4	N1_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C0	H1_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C0	NP_C4	N1_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C4	H1_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H1_C1	NP_C0	N1_C3	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H1_C1	NP_C0	N2_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H1_C1	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C0	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N1_C3	M_C0	NS_C0
PB2_C1	PB1_C3	PA_C3	H1_C3	NP_C0	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C3	NP_C4	N1_C3	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_C0	N6_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C4	N3_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H3_C0	NP_C0	N2_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C0	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C4	H3_C0	NP_C4	N2_C0	M_C0	NS_C1
PB2_C2	PB1_C1	PA_C2	H3_C1	NP_C1	N8_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_C0	N3_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_C0	N1_C3	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_C0	N2_C1	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H5_C0	NP_C4	N2_C1	M_C0	NS_C0
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_C0	N2_C1	M_C0	NS_C0
PB2_C1	PB1_C3	PA_C3	H9_C1	NP_C0	N2_C1	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C2	NP_C0	N2_C1	M_C0	NS_C1

**Tabla anexa 7. Genomas únicos de virus de influenza A cuyo hospedero son aves. Estos fueron escogidos en base a ensayos de hibridación *in silico* contra genomas de influenza A provenientes de aves con los oligonucleótidos que fueron impresos en el microarreglo.**

Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5	Segmento 6	Segmento 7	Segmento 8
PB2_C2	PB1_C0	PA_C0	H1_C1	NP_C4	N1_C2	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C1	NP_C4	N2_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H1_C2	NP_C0	N1_C3	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N1_C2	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N1_C3	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N2_C1	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N3_C0	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N5_C0	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N6_C1	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N9_C0	M_C0	NS_C0
PB2_C1	PB1_C3	PA_C3	H1_C2	NP_C0	N1_C3	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H1_C2	NP_C0	N4_C2	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H1_C2	NP_C0	N9_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N2_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N3_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N6_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H1_C2	NP_C0	N9_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H2_C0	NP_C0	N2_C1	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H2_C0	NP_C0	N3_C2	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H2_C0	NP_C0	N9_C0	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H2_C0	NP_C0	N9_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C0	NP_C0	N1_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C0	NP_C0	N6_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C0	NP_C0	N8_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N2_C1	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N3_C0	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N7_C0	M_C0	NS_C0
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N1_C3	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N3_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N4_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N5_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N7_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C0	N9_C0	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H2_C1	NP_C5	N1_C3	M_C0	NS_C1
PB2_C1	PB1_C3	PA_C3	H3_C0	NP_C0	N2_C1	M_C0	NS_C0
PB2_C1	PB1_C3	PA_C3	H3_C0	NP_C0	N8_C1	M_C0	NS_C0
PB2_C1	PB1_C3	PA_C3	H3_C0	NP_C0	N8_C1	M_C0	NS_C1

PB2_C1	PB1_C3	PA_C3	H3_C0	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C0	NP_C4	N2_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N1_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N5_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N6_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N8_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C2	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C0	PA_C3	H3_C3	NP_CO	N8_C2	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C3	NP_CO	N8_C2	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N6_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N8_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N9_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N7_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C4	NP_CO	N8_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H3_C5	NP_CO	N6_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H3_C5	NP_CO	N8_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C5	NP_CO	N6_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H3_C5	NP_CO	N8_C3	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H3_C5	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H3_C5	NP_CO	N8_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C5	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N1_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N5_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N6_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N8_C3	M_CO	NS_CO
PB2_C2	PB1_C0	PA_C3	H4_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N5_C0	M_CO	NS_C1

PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N6_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N7_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H4_C0	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H4_C1	NP_CO	N2_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H4_C1	NP_CO	N8_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H4_C2	NP_CO	N6_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H4_C2	NP_CO	N8_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H4_C2	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H4_C2	NP_CO	N3_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H4_C2	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H4_C3	NP_CO	N8_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N1_C3	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N3_C2	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N6_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N3_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N7_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N8_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H5_C0	NP_CO	N9_C1	M_CO	NS_C1
PB2_C2	PB1_C0	PA_C3	H5_C1	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C0	PA_C3	H5_C1	NP_CO	N9_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N1_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N3_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N5_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N8_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N9_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H5_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C0	PA_C3	H5_C1	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N4_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N7_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H5_C1	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N1_C3	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N8_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N1_C3	M_CO	NS_CO

PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N8_C3	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N5_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N8_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N8_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C0	NP_CO	N9_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C0	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N3_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N8_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N4_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C1	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C2	NP_CO	N1_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C2	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H6_C2	NP_CO	N5_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C2	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C2	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H6_C2	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N1_C3	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N3_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N6_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N7_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N8_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H7_C0	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N1_C3	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N3_C2	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N7_C1	M_CO	NS_CO

PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N8_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N1_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N3_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N7_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C2	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N2_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N3_C2	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N7_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N3_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N4_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H7_C3	NP_CO	N7_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H8_C0	NP_CO	N4_C2	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H8_C0	NP_CO	N4_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H8_C0	NP_CO	N4_C2	M_CO	NS_C1
PB2_C2	PB1_C0	PA_C3	H8_C0	NP_CO	N4_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H8_C0	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H8_C0	NP_CO	N4_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H9_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C1	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H9_C1	NP_CO	N2_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H9_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C1	NP_CO	N2_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C1	NP_CO	N6_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H9_C1	NP_CO	N1_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H9_C1	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H9_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H9_C1	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H9_C1	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C2	NP_CO	N2_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H9_C2	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N3_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N7_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N6_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N7_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H10_C0	NP_CO	N8_C3	M_CO	NS_C1

PB2_C1	PB1_C3	PA_C3	H10_C1	NP_CO	N7_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H10_C1	NP_CO	N4_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H10_C1	NP_CO	N8_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H10_C1	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H10_C1	NP_CO	N9_C1	M_CO	NS_C1
PB2_C2	PB1_C0	PA_C3	H11_C0	NP_CO	N9_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N2_C1	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N3_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N9_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H11_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C0	NP_CO	N9_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N2_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N3_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H11_C0	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N8_C1	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N9_C0	M_CO	NS_CO
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N1_C3	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H11_C1	NP_CO	N9_C1	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H12_C0	NP_CO	N5_C0	M_CO	NS_CO
PB2_C2	PB1_C3	PA_C3	H12_C0	NP_CO	N1_C3	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H12_C0	NP_CO	N4_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H12_C0	NP_CO	N5_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H12_C0	NP_CO	N6_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H12_C1	NP_CO	N3_C2	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H12_C1	NP_CO	N8_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H12_C1	NP_CO	N9_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H13_C0	NP_CO	N6_C0	M_CO	NS_C1
PB2_C2	PB1_C3	PA_C3	H13_C0	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H13_C1	NP_CO	N2_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H13_C1	NP_CO	N6_C0	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H15_C0	NP_CO	N9_C1	M_CO	NS_C1
PB2_C1	PB1_C3	PA_C3	H16_C1	NP_CO	N3_C1	M_CO	NS_C1

**Tabla anexa 8. Genomas únicos de virus de influenza A cuyo hospedero son caballos. Estos fueron escogidos en base a ensayos de hibridación *in silico* contra genomas de influenza A provenientes de aves con los oligonucleótidos que fueron impresos en el microarreglo.**

Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5	Segmento 6	Segmento 7	Segmento 8
PB2_C2	PB1_C1	PA_C2	H3_C1	NP_C1	N8_C0	M_C0	NS_C1
PB2_C0	PB1_C2	PA_C1	H7_C1	NP_C2	N7_C2	M_C1	NS_C1
PB2_C2	PB1_C1	PA_C2	H7_C1	NP_C1	N7_C2	M_C1	NS_C1
PB2_C1	PB1_C3	PA_C3	H9_C2	NP_C0	N2_C1	M_C0	NS_C1

**Tabla anexa 9. Genomas únicos de virus de influenza A cuyo hospedero son perros. Estos fueron escogidos en base a ensayos de hibridación *in silico* contra genomas de influenza A provenientes de aves con los oligonucleótidos que fueron impresos en el microarreglo.**

Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5	Segmento 6	Segmento 7	Segmento 8
PB2_C1	PB1_C3	PA_C3	H3_C0	NP_C0	N2_C1	M_C0	NS_C1
PB2_C2	PB1_C3	PA_C3	H3_C0	NP_C4	N1_C3	M_C0	NS_C1