



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

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

INSTITUTO DE BIOTECNOLOGÍA

**DEPARTAMENTO DE GENÉTICA DEL DESARROLLO Y FISIOLÓGIA
MOLECULAR**

**IMPLEMENTACIÓN Y VALIDACIÓN DE UNA METODOLOGÍA
PARA LA DETECCIÓN DE PATÓGENOS ASOCIADOS A
ENFERMEDADES RESPIRATORIAS POR MEDIO DE UN
MICROARREGLO**

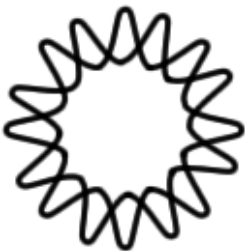
TESIS

QUE PARA OBTENER EL GRADO ACADÉMICO DE

MAESTRO EN CIENCIAS

PRESENTA:

LIC. EN C.G. MARÍA DE LOS DOLORES SOTO DEL RÍO



**DIRECTOR DE TESIS: DR. CARLOS FEDERICO
ARIAS ORTIZ**

**CUERNAVACA, MORELOS, MÉXICO.
AGOSTO 2011**



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

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 F. Arias y Susana López. Se contó con los apoyos de DGPA (IN210807-3), del Instituto de Ciencia y Tecnología del Distrito Federal (PICOSI09-209), del Consejo Nacional de Ciencia y Tecnología y la Secretaría de Salud (S0008-126823).

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

Comité Tutorial:

Dr. Carlos Federico Arias Ortiz

Dra. Rosa María Gutiérrez Ríos

Dr. José Luis Reyes Taboada

Jurado de examen:

Dr. Jean Louis Charli Casalonga

Dr. Ramón González García Conde

Dr. Ernesto Pérez Rueda

Dr. Federico Sánchez Rodríguez

Dr. Lorenzo Segovia

DEDICATORIAS

El presente trabajo de tesis está dedicado a las siguientes personas en agradecimiento por todo lo que me han brindado y significan en mi vida.

A mi madre por haberme enseñado a luchar día a día por los objetivos planteados, por ayudarme a ser una mejor persona, mujer y estudiante. Pero sobre todo por la más grande lección de vida que he tenido: su partida. Gracias Princesa!

A mi padre por ser mi gran apoyo, mi soporte constante e incondicional, mi cómplice, amigo, consejero fiel y el mejor maestro de vida. Gracias Güerito por siempre creer en mi!

A mis padrinos por ser ese soporte emocional y amoroso, por todas esas sonrisas, atenciones y detalles tan hermosos que me han regalado.

A mis hermanas Adriana y Guadalupe y a mi hermano Alfredo por siempre estar al pendiente, por sus atenciones, cariño, y su ejemplo de entrega.

A mis tías Rosa y Consuelo por ser mis ángeles guardianes en todo momento y en todos aspectos.

A mis sobrinos Desireé, Astrid, Alyn y Omar por todas esas pláticas, discusiones, risas, travesuras y lágrimas. Pero sobre todo porque siempre han visto en mi a una amiga, un cómplice y un confidente más que a una tía.

A mis tíos Yolanda, Martha, Basilio y María Elena por ayudarme a mantener siempre los pies en la tierra.

A Martha y a Pepe por ser los mejores amigos y compañeros desde la licenciatura, por todos esos grandes momentos compartidos, por cada sonrisa, chiste, confesión. Los quiero crotalitos!

Al Dr. Carlos Arias por darme la oportunidad de conocer el fascinante mundo de la investigación bajo su tutela, por sus enseñanzas, motivaciones, consejos, anécdotas, por siempre confiar en mí, por brindarme su apoyo incondicional, pero sobre todo gracias por tu amistad Jefe!

A TODOS Y CADA UNO, MUCHAS GRACIAS DE TODO CORAZÓN, NO HUBIERA PODIDO SIN USTEDES!

AGRADECIMIENTOS

A la Dra. Susana López por permitirme trabajar en su laboratorio, por todos sus comentarios y críticas constructivas al proyecto.

Al Dr. Pavel Isa por haberme enseñado a trabajar en el laboratorio, por tenerme mucha paciencia y por siempre estar al pendiente del proyecto.

A Federico Sánchez Quinto y Selene Fernández Valverde por haber sentado las bases para la realización de este proyecto.

A los integrantes del laboratorio Arias/López: Doña Lore, Silvia, Tomás, Luis, Joaquín, Yunuen, Andrea M, Ana, Daniela, Alfonso, Héctor, Enrique, Paula, Jesús, Vicenta, Rafa, Andrea P, Rosa y Rodrigo. Gracias a Luis Felipe, Marco Antonio y a Miguel M. por el gran trabajo en equipo, sugerencias, discusiones y por considerar valiosas mis observaciones. Gracias a Don Pedro, Ernesto, Miguel Olvera y Marco Aurelio por sus atenciones, detalles y por su amistad. Y de manera muy especial a Margarito, Marina, Yina, Diego y Lili por haberme permitido compartir con ellos la mesa más feliz del laboratorio.

A los integrantes de mi Comité Tutoral por sus consejos, críticas y esfuerzo durante la realización de este proyecto de tesis.

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

A la Dra. Rosa María Gutiérrez por sus asesorías en la parte de estadística.

A Jerome por su gran ayuda, disponibilidad, eficiencia y amabilidad con el manejo del cluster.

A los integrantes de la Unidad de Docencia: Jalil, Gloria y Antonio por su disposición, amabilidad y eficiencia.

A los que me faltaron por mi demencia juvenil (mil disculpas, pero ya saben que mi memoria es de teflón)

ÍNDICE GENERAL

	Página
ABREVIATURAS	1
ÍNDICE DE FIGURAS	3
ÍNDICE DE TABLAS	4
I. RESUMEN	5
II. INTRODUCCIÓN	7
1. Métodos de diagnóstico	7
2. Microarreglos	8
3. Enfermedades Respiratorias	10
4. Virus causantes de infecciones respiratorias	12
5. Virus de influenza	15
6. Pandemias de influenza	18
7. Necesidad de nuevos métodos diagnósticos	22
III. HIPÓTESIS	23
IV. OBJETIVOS	23
1. Objetivo general	23
2. Objetivos particulares	23
V. MATERIALES Y MÉTODOS	24
1. Selección de patógenos virales a incluir	24
2. Obtención de secuencias	24
3. Selección de secuencias	24
a. Construcción de árboles filogenéticos	24
b. Selección de secuencias por porcentajes de identidad	25
4. Creación de setentámeros	25
5. Selección de oligonucleótidos	25
6. Selección final de oligonucleótidos	26
7. Preparación de oligonucleótidos	27
8. Impresión de microarreglos	27
9. Post-procesamiento de los microarreglos	27
10. Virus utilizados para validar el microarreglo	28
11. Muestras clínicas de humanos	29
12. Virus de influenza aislados de cerdos	29
13. Extracción de RNA y amplificación de material genético presente en las muestras	30

14. Limpieza y marcaje del DNA	32
15. Hibridación	33
16. Lavados y lecturas de los microarreglos	33
17. Análisis de datos	34
18. Análisis de agrupamiento (clustering)	36
VI. RESULTADOS Y DISCUSIÓN	37
1. Selección de los virus a incluir en el microarreglo	37
2. Obtención de las secuencias de genomas o genomas virales	37
3. Selección de secuencias	37
4. Generación de setentámeros a partir de las secuencias virales seleccionadas	41
5. Selección de oligonucleótidos específicos	41
6. Selección final de oligonucleótidos para el microarreglo	42
7. Validación del microarreglo con cepas de referencia	43
8. Implementación de un método estadístico para determinar la reproducibilidad de los ensayos	48
9. Identificación del subtipo de virus de influenza A	51
10. Visualización gráfica de los resultados	54
11. Análisis de muestras clínicas	57
12. Optimización de la sensibilidad del ensayo de microarreglo	59
13. Re-análisis de muestras clínicas con el protocolo con sensibilidad mejorada	63
14. Análisis de muestras de campo de origen porcino	66
VII. CONCLUSIONES Y PERSPECTIVAS	71
VIII. REFERENCIAS	75
IX. ANEXOS	80
I. Familias, géneros y especies virales incluidos en el microarreglo	80
II. Resultados del análisis de producto de rangos (RP) de cepas de referencia	85

ABREVIATURAS

aa-dUTP	5-(3-aminoalil)-2'-deoxiuridinda 5'-trifosfato
BLAST	Herramienta de búsqueda de alineamientos locales (Basic Local Alignment Search Tool)
CDC	Centro para el Control y Prevención de Enfermedades (Centers for Control Disease and Prevention)
CMN-IMSS	Centro Médico Nacional Siglo XXI del Instituto Mexicano del Seguro Social
Cy3	Cianina 3
Cy5	Cianina 5
dG	Energía libre de Gibbs
DNA	Ácido desoxirribonucleico (desoxyribonucleic acid)
dNTPs	Desoxiribonucleótidos trifosfato
dTTP	Desoxitimina trifosfato
ELISA	Ensayo por inmunoadsorción ligado a enzimas (Enzyme-Linked ImmunoSorbent Assay)
FDR	Tasa de falsos positivos (False Discovery Rate)
FMVZ	Facultad de Medicina Veterinaria y Zootecnia
gal	Lista de arreglo de GenePix (GenePix Array List)
gpr	Resultados de GenePix (GenePix Results)
H	Hemaglutinina
HCoV	Coronavirus humano
HEPES	Ácido 2-[4-(2-hidroxietil)-1-pireracil-etanosulfónico
hMPV	Metapneumovirus humano
HPIV2	Virus de parainfluenza humano tipo 2
HPIV3	Virus de parainfluenza humano tipo 3
HRV	Rinovirus humano
HuAdV6	Adenovirus humano tipo 6
ICTV	Comité Internacional de Taxonomía Viral (International Committee on Taxonomy of Virus)
INEGI	Instituto Nacional de Estadística, Geografía e Informática
IRA	Infecciones respiratorias agudas
M1	Proteína de matriz 1
M2	Proteína de matriz 2

N	Neuraminidasa
NCBI	Centro nacional de información biotecnológica
NEP	Proteína de exportación nuclear
NN	Vecino más cercano (Nearest Neighbor)
NP	Nucleoproteína
NS1	Proteína no estructural 1
NS2	Proteína no estructural 2
nt	Núcleotido
OMS	Organización Mundial de la Salud
PA	Subunidad ácida de la RNA polimerasa de influenza
PB1	Subunidad básica 1 de la RNA polimerasa de influenza
PB2	Subunidad básica 2 de la RNA polimerasa de influenza
PCR	Reacción en cadena de la polimerasa (polymerase chain reaction)
RNA	Ácido ribonucleico (ribonucleic acid)
RP	Productos de rango (Rank Products)
RSV	Virus sincicial respiratorio (respiratory sincitial virus)
RSVA	Virus sincicial respiratorio tipo A (respiratory sincitial virus type A)
RSVB	Virus sincicial respiratorio tipo B (respiratory sincitial virus type B)
RT-PCR	Transcripción reversa seguida de la reacción en cadena de la polimerasa (Reverse Transcription-Polymerase Chain Reaction)
SARS	Síndrome respiratorio agudo severo (Severe Acute Respiratory Syndrome)
SDS	Duodecil sulfato de sodio
SIDA	Síndrome de inmuno deficiencia adquirida
SSC	Solución salina de citrato de sodio
TIFF	Formato de archivo de imagen etiquetada (Tagged Image File Format)
UCSF	Universidad de California en San Francisco
UV	Ultra violeta

ÍNDICE DE FIGURAS

	Página
Figura 1. Orígenes de las cepas pandémicas de influenza A en humanos.	19
Figura 2. Orígenes de los genes de la cepa pandémica de influenza A/H1N1/2009.	21
Figura 3. Árboles filogenéticos para H1.	39
Figura 4. Rondas de amplificación.	44
Figura 5. Esquema del proceso de fabricación y análisis de los microarreglos de diagnóstico.	46
Figura 6. Imágenes obtenidas de experimentos de hibridación de microarreglos de cepas de referencia.	47
Figura 7. Análisis de agrupamiento de las muestras de referencia de las 16 variantes de H y 9 de N de influenza tipo A.	56
Figura 8. Análisis de muestras clínicas con método comercial para detección de virus causantes de enfermedades respiratorias.	58
Figura 9. Análisis de agrupamiento de las amplificaciones de muestras diluidas con diferentes proporciones de nucleótido modificado.	62
Figura 10. Análisis de agrupamiento de cepas porcinas de campo, estacional y pandémica humana.	70

ÍNDICE DE TABLAS

	Página
Tabla 1. Patógenos asociados a enfermedades respiratorias.	11
Tabla 2. Características de las familias virales clásicamente asociadas a enfermedades respiratorias.	12
Tabla 3. Secuencias y oligonucleótidos seleccionados de hemaglutininas y neuraminidasas de influenza tipo A.	40
Tabla 4. Incorporación específica del fluoróforo en el DNA amplificado.	45
Tabla 5. Resultados del análisis de producto de rangos de cepas de referencia.	49
Tabla 6. Análisis de los datos de hibridación del panel inicial de cepas de referencia.	50
Tabla 7. Análisis de los datos de hibridación de las muestras de referencia de las 16 variantes de H y 9 de N de influenza tipo A.	53
Tabla 8. Cuantificación de fluoróforo incorporado en el DNA sintetizado.	61
Tabla 9. Análisis de muestras clínicas por producto de rangos.	64
Tabla 10. Muestras colectadas y su análisis para estudio epidemiológico de cepas de influenza A circulantes en cerdos en México.	68

I. RESUMEN

Las infecciones respiratorias son una de las principales causas de muertes a nivel mundial. La detección temprana y un tratamiento eficiente de estas infecciones pudiera contribuir a una disminución en estos fallecimientos. Es por ello que pruebas efectivas de diagnóstico, que permitan la identificación del mayor número de patógenos posible, de una manera rápida sencilla y costeable es de mayor importancia. Actualmente con la disponibilidad de un gran número de secuencias de genomas virales depositadas en las bases de datos es posible diseñar métodos para la detección molecular y subtipificación de los agente etiológicos causantes de estos padecimientos, que pudieran funcionar también como herramientas de vigilancia epidemiológica. Es por ello que los microarreglos han sido adoptados como una importante plataforma para el diagnóstico de infecciones respiratorias. Con esta finalidad, hemos diseñado y construido un microarreglo para detectar virus que han sido asociados a infecciones del tracto respiratorio de vertebrados. Estos forman una lista de 189 especies virales, comprendidos en 51 géneros, pertenecientes a 15 familias. En el caso del virus de influenza A, el microarreglo fue diseñado para detectar todas las variantes de hemaglutininas (H) y neuraminidasas (N).

El microarreglo consiste de oligonucleótidos de 70 bases que reconocen de manera específica cada uno de los virus. El diseño de estas sondas está sustentado en el reconocimiento específico de todas las secuencias disponibles en las bases de datos, usando diversas herramientas bioinformáticas como análisis filogenéticos, BLAST y parámetros termodinámicos. El análisis de los datos se basa en la normalización de éstos, que posteriormente son sometidos al algoritmo de producto de rangos.

El microarreglo fue validado con cepas de referencia de influenza A que comprende todos los diferentes subtipos de H y N, así como con diferentes cepas de referencia de virus como SARS, adenovirus humano tipo 6, sarampión, parainfluenza humana tipo 2 y 3. Se hicieron modificaciones al protocolo para aumentar la sensibilidad del microarreglo, con lo cual fue posible caracterizar un par de muestras clínicas; el resultado fue corroborado por RT-PCR. Se estudiaron también un conjunto de muestras de cepas de influenza A porcinas circulantes en México, pudiendo detectar y subtipificar el virus.

Este trabajo presenta el diseño, implementación y validación de una valiosa herramienta de identificación, caracterización y vigilancia epidemiológica de una amplia gama de virus asociados a enfermedades respiratorias.

I. INTRODUCCIÓN

Las enfermedades infecciosas son el resultado de la interacción entre los microorganismos, el hospedero, y el medio ambiente [1]. Existen diversos factores que determinan la frecuencia y el control de las enfermedades, como son la patogenicidad y virulencia del agente infeccioso; la capacidad inmune del hospedero y su susceptibilidad genética, el medio ambiente; además de diversos factores culturales y económicos.

Estos padecimientos ocasionan más de una cuarta parte de las defunciones a nivel mundial. En particular el SIDA, la tuberculosis, la diarrea, el paludismo, el sarampión y las enfermedades respiratorias, ocasionan 90% de estas muertes [2]. Sin embargo, no se conoce con exactitud la magnitud real del problema, ya que especialmente en países en vías de desarrollo, muchas de las muertes causadas por enfermedades infecciosas no se clasifican como tales y su seguimiento es deficiente [3].

En la actualidad, el diagnóstico y control de muchos padecimientos infecciosos es posible gracias al conocimiento que se tiene de los agentes causales y la forma en que el cuerpo humano se defiende de ellas. El conocimiento derivado de la secuenciación del genoma completo de alrededor de 1300 bacterias y 2700 diferentes virus al momento [4], así como la secuencia completa del genoma humano [5], ha generado nuevas herramientas que ayudan al entendimiento y control las enfermedades infectocontagiosas.

1. Métodos de diagnóstico

Las técnicas de diagnóstico para enfermedades infecciosas se basan principalmente en la identificación de la presencia de los microorganismos. Para ello, se buscan componentes de éstos, tales como lípidos, proteínas y/o material genético; o la respuesta inmune que evocan [1]. Los métodos más comunes son el examen microscópico directo, el aislamiento de virus en cultivos celulares, el crecimiento en medios selectivos de bacterias, la detección de antígenos o subproductos microbianos por métodos inmunológicos o moleculares y la detección de anticuerpos específicos contra componentes del patógeno.

La detección de microorganismos por medio de técnicas de biología molecular, tales como la reacción en cadena de la polimerasa (PCR) y la detección de segmentos específicos del material genético mediante hibridación de ácidos nucleicos han mejorado el diagnóstico de los patógenos asociados a las enfermedades infecciosas, en donde las técnicas estándar presentan limitaciones. Estas nuevas tecnologías tienen como principales ventajas la rapidez, la sensibilidad y la especificidad de detección, además de requerir poca cantidad de muestra, y de no ser necesario crecer al microorganismo en cultivo. Además de esto, ofrecen la posibilidad de analizar la presencia de varios microorganismos en un mismo ensayo.

Variaciones a la técnica de PCR, como es el PCR múltiple, han sido de gran utilidad para el diagnóstico viral, ya que permiten la detección simultánea de varios patógenos en un mismo ensayo [6]. Estas pruebas han mostrado tener una sensibilidad igual o incluso mayor a los ensayos de cultivo [7]. El PCR en tiempo real también se ha adoptado para estos fines y ha mostrado una sensibilidad mayor que los ensayos clásicos de cultivo para distintos virus causantes de enfermedades respiratorias [8, 9].

Otra aproximación dentro de las técnicas de biología molecular corresponde a las metodologías basadas en sondas de ácidos nucleicos. Éstas consisten en fragmentos de DNA o RNA obtenidos del genoma del patógeno de interés. Estos fragmentos se pueden ligar a un fluoróforo, permitiendo la detección en la muestra estudiada de la secuencia complementaria a la sonda cuando ocurre el proceso de hibridación, bajo condiciones óptimas de temperatura y fuerza iónica.

2. Microarreglos

Los microarreglos son soportes sólidos en los cuales se encuentran inmovilizados, en áreas pequeñas de cientos a miles de secuencias de DNA de manera ordenada [10]. Los soportes pueden ser laminillas de vidrio o membranas de nylon. En los microarreglos el DNA puede ser depositado o sintetizado directamente sobre la superficie sólida, siendo la primera forma la más utilizada, la cual también se conoce como “impresión”.

La impresión del microarreglo la realiza un robot que aplica al soporte sólido un cierto volumen de las secuencias seleccionadas. Las superficies en las cuales se deposita el

ácido nucléico están recubiertas con, por ejemplo, poli-lisina, aminas, aldehídos, epoxy, etc., que permiten la unión química del DNA a la superficie. Una vez que se imprimieron las laminillas el DNA se fija a éstas por temperatura o luz UV. Con esto el microarreglo queda listo para la hibridación.

Los pasos a seguir para la preparación de la muestra a hibridar son la extracción y purificación de los ácidos nucléicos, la amplificación del material genético, y el marcaje del DNA amplificado para permitir su detección. Los marcadores más comunes son los fluorescentes, pero también puede utilizarse el marcaje radiactivo o la quimioluminiscencia. La siguiente fase es la hibridación, siendo la etapa clave de la prueba, ya que es ahí donde se produce la reacción de afinidad en la que hibridan las hebras de DNA de las muestras marcadas, con sus complementarias inmovilizadas en la superficie del microarreglo.

El último paso es la lectura del microarreglo. Para ello se utilizan lectores con cámaras digitales o confocales. En ambos casos se usa un láser que excita a las moléculas fluorescentes para obtener la imagen de cada uno de los canales. Se localizan los sitios en los cuales hibridó el DNA y se emitió fluorescencia, se obtiene la intensidad relativa de cada canal y se cuantifican los perfiles. En este paso se analizan las imágenes para cada longitud de onda apreciando así los puntos en los que la reacción de hibridación ha sido positiva y aquellos en los que no hubo tal. Los datos deben ser normalizados para poder ser comparables y analizados, encontrando así los genes o secuencias presentes en la muestra de interés.

Es importante mencionar que los microarreglos para diagnóstico suelen utilizar secuencias que funcionan como control de ubicación, ya que se espera que sólo se tenga señal en algunos puntos del microarreglo. En nuestro caso (ver adelante) todos los puntos tienen impresos, además de las secuencias blanco, un oligonucleótido llamado “spike-70”, que es un setentámero sin homólogos biológicos reportados hasta el momento [11]. Durante el proceso de hibridación, además de la muestra a analizar, se incluye una secuencia que es complementaria al “spike-70” a la cual se le llama “probe-70”. El objetivo de estas secuencias es el de servir como referencia para evaluar la homogeneidad de la impresión y la hibridación del microarreglo, además de permitir la alineación de los puntos impresos [11].

Esta metodología es una herramienta muy poderosa para analizar simultáneamente la presencia de cientos de patógenos en un tiempo reducido. Es por ello que se ha adoptado a los microarreglos como una técnica clave para el diagnóstico de enfermedades infecciosas. Existen diversos trabajos que han descrito microarreglos que detectan un gran número de patógenos virales relacionados con diferentes enfermedades infecciosas [12-21]. Este trabajo se centra en enfermedades infecciosas respiratorias causadas por virus.

3. Enfermedades respiratorias

Las infecciones respiratorias agudas (IRA) representan un problema importante de salud pública por las muertes que originan cada año, principalmente en niños menores de 5 años y adultos de más de 65; además del impacto que tienen sobre los servicios de salud, las incapacidades laborales, escolares, etc. La manifestación y gravedad de las IRA depende de varios factores como la edad de los pacientes (afectando primordialmente al grupo de individuos en el rango de edad ya mencionado), la localización de la infección en el aparato respiratorio, la resistencia del individuo, la estación del año, el medio ambiente, el agente causal y su virulencia, entre otros. En nuestro país se desconoce con precisión la etiología en estos casos ya que la gran mayoría de éstos no se diagnostican.

Estos padecimientos incluyen afecciones pulmonares, de pleura, bronquios, tráquea y de vías respiratorias superiores. Van desde enfermedades leves hasta severas, como es el caso del resfriado común, la neumonía o embolia pulmonar.

En 2005 la OMS calculó que 10.6 millones de niños menores de 5 años mueren anualmente por enfermedades infecciosas, el 19% de éstos por neumonía [2]. Según datos del Centro para el Control y Prevención de Enfermedades (CDC por sus siglas en inglés), durante el año 2005 las enfermedades respiratorias causaron aproximadamente el 7.1% de los decesos en los Estados Unidos; el 5.3% de estas muertes por infecciones crónicas del tracto respiratorio inferior y el 2.6% por influenza y neumonía bacteriana. Lo anterior implica que en 2005 alrededor de 193,000 personas murieron en dicho país por estos padecimientos [22]. En México, según datos de la Secretaría de Salud y el INEGI, en el año 2004 hubo 3,598 defunciones por neumonía en niños menores de 5 años, 3,227 en el año

2005 y 2,245 en el año 2006. Mientras que para adultos, durante 2005, el 2.8% de las muertes se debieron a influenza y pulmonía, causando alrededor de 84,000 defunciones.

La tasa mundial de mortalidad anual por neumonías en niños de 1 a 4 años en 2006 fue de 7.2 por 100,000 niños y en niños menores de 1 año de 109 por 100,000 niños [2]. La tasa de incidencia anual de neumonía en Estados Unidos es de 40 por 1000 en niños menores de 5 años y 7 por 1000 en adolescentes [23]. En México la tasa de incidencia de neumonías y bronconeumonías reportada en 2008 fue de 157 por cada 1,000 niños menores de 1 año y de 411 por cada 100,000 niños de 1 a 4 años de edad [24].

Los agentes etiológicos más comunes de las enfermedades respiratorias pueden ser bacterias, hongos y virus, aunque los agentes virales son causantes de cerca del 70% de estos casos [26]. En la Tabla 1 se enlistan los patógenos más conocidos, en Estados Unidos, para dichos padecimientos [modificado de 25].

Tabla 1. Patógenos asociados a enfermedades respiratorias.

Rasgos clínicos	Bacterias	Virus	Hongos
Resfriado común	Poco frecuentes	Rhinovirus	Poco frecuente
		Coronavirus	
		Parainfluenza	
		Adenovirus	
		Virus respiratorio sincial Influenza	
Faringitis, amigdalitis	Estreptococos grupo A B hemolítico	Adenovirus	<i>Candida albicans</i>
	<i>Corynebacterium diphtheriae</i>	Coxsackievirus	
	<i>Neisseria gonorrhoeae</i>	Influenza	
	<i>Mycoplasma pneumoniae</i>	Rhinovirus	
	<i>Mycoplasma hominis</i>	Coronavirus	
	diversas anaerobias	Parainfluenza	
		Epstein-Barr, cytomegalovirus Herpes simple	
Epiglottitis, laringotraqueitis	<i>Haemophilus influenzae</i>	Virus respiratorio sincial	Poco frecuente
	<i>Corynebacterium diphtheriae</i>	Parainfluenza	
Bronquitis y bronqueolitis	<i>Haemophilus influenzae</i>	Parainfluenza	Poco frecuente
	<i>Streptococcus pneumoniae</i>	Virus respiratorio sincial	
	<i>Mycoplasma pneumoniae</i>	Adenovirus	
		Herpes simple	
Neumonía	<i>Streptococcus pneumoniae</i>	Adenovirus	<i>Histoplasma capsulatum</i>
	<i>Staphylococcus aureus</i>	Parainfluenza	<i>Blastomyces dermatidis</i>
	<i>Streptococcus pyogenes</i>	Virus respiratorio sincial	<i>Paracoccidioides brasiliensis</i>
	<i>Haemophilus influenzae</i>	Influenza	<i>Coccidioides immitis</i>
	<i>Klebsiella pneumoniae</i>	Varicella-zoster	<i>Candida albicans</i>
	<i>Escherichia coli</i>	Sarampión	<i>Aspergillus fumigatus</i>
	<i>Pseudomonas aeruginosa</i>	Cytomegalovirus	<i>Aspergillus spp</i>
	<i>Mycoplasma pneumoniae</i>	Herpes simple	<i>Filobasidiella spp</i>
	<i>Mycobacterium tuberculosis</i>	Hantavirus	
	<i>Coxiella burnetii</i>		
	<i>Chlamydia psittaci</i>		
	<i>Chlamydia trachomatis</i>		
	<i>Chlamydia pneumoniae</i>		
	<i>Legionella spp</i>		
	<i>Mycoplasma spp</i>		
	diversas anaerobias		

4. Virus causantes de infecciones respiratorias

Los agentes virales asociados de manera clásica a enfermedades respiratorias pertenecen a cinco familias: *Adenoviridae*, *Paramyxoviridae*, *Orthomyxoviridae*, *Coronaviridae* y *Picornaviridae* (Tabla 2) [modificada de 26]. Algunos virus pertenecientes a otras familias también se asocian a síntomas respiratorios en las fases tempranas de la infección, aunque la enfermedad principal que causan es fuera del tracto respiratorio, por ejemplo virus de la familia *Herpesviridae*.

Tabla 2. Características de las familias virales clásicamente asociadas a enfermedades respiratorias.

Familia	Genoma	Tamaño del virión (nm)	Replicación	Algunos miembros
<i>Orthomyxoviridae</i>	RNA cadena sencilla polaridad negativa segmentado	~100	nuclear y citoplásmica	Influenza A, B, C
<i>Paramyxoviridae</i>	RNA cadena sencilla polaridad negativa	150-200	citoplásmica	Virus respiratorio sincicial, parainfluenza virus 1 a 4
<i>Picornaviridae</i>	RNA cadena sencilla polaridad positiva	~30	citoplásmica	Rhinovirus, Coxsackievirus, Echovirus
<i>Coronaviridae</i>	RNA cadena sencilla polaridad postiva	60-220	citoplásmica	Coronavirus, SARS
<i>Adenoviridae</i>	DNA doble cadena	80-90	nuclear	Adenovirus

A continuación se describen de manera particular los virus más importantes, comúnmente asociados a infecciones respiratorias:

Virus sincicial respiratorio El RSV (por sus siglas en inglés) pertenece a la familia *Paramyxoviridae*, tiene un genoma de RNA de cadena sencilla de polaridad negativa. Infecta a pacientes de todas las edades, aunque a los 2 años de edad, la mayoría de los niños ya han sido infectados por este virus, siendo el patógeno más importante en este grupo de pacientes [27]. En climas templados se presenta de manera anual en los meses de invierno, mientras que en climas tropicales el mayor número de casos se da en la temporada de lluvias.

En la mayoría de los casos pediátricos el RSV produce sólo síntomas menores, indistinguibles de un resfriado común o enfermedades menores. Para algunos niños, sin embargo, el virus puede causar bronquiolitis, produciendo un cuadro respiratorio grave que puede requerir hospitalización y, en algunos casos ocasiona la muerte. Esto ocurre

mayormente en pacientes con problemas inmunológicos o en bebés prematuros. Para el caso de adultos, los estudios muestran que RSV es el principal responsable de neumonía y bronquiolitis [28].

Los métodos de detección están basados en el uso de la inmunofluorescencia, por medio de ELISA, la neutralización, el PCR, o el aislamiento y confirmación por cultivo, este último puede tardar hasta 5 días [29, 30, 31].

Virus de parainfluenza. Estos virus pertenecen a la familia *Paramyxoviridae* y junto con RSV, representan los patógenos de vías aéreas superiores más importantes en niños y lactantes. Su genoma es de RNA de cadena sencilla de polaridad negativa. Se han descrito 4 serotipos virales que presentan relaciones antigénicas entre sí: parainfluenza tipo 1, tipo 2, tipo 3, y tipo 4. Los serotipos 1 y 2 son la causa más importante de laringotraqueobronquitis y producen enfermedad severa en niños menores de 2 años [32]. El tipo 3 es asociado de manera frecuente a casos de bronquiolitis y neumonía, siendo más severa en niños menores de un año de edad [32]. El tipo 4 se asocia con enfermedades respiratorias de vías superiores en adultos y en niños.

El diagnóstico para los serotipos 1, 2 y 3 es por cultivo celular; mientras que el serotipo 4 es detectado por técnicas de inmunofluorescencia y PCR [32].

Adenovirus. Pertenecen a la familia *Adenoviridae*, su genoma está compuesto por una molécula lineal de DNA doble cadena. Son causa frecuente de gripe, faringitis y amigdalitis en lactantes y niños pequeños, y aproximadamente el 10% de las neumonías en niños es causada por adenovirus [33]. Los adenovirus también son causa de bronquitis y bronquiolitis. En niños menores de 5 años son responsables de alrededor del 5% de los casos de infección respiratoria aguda (IRA), presentando síntomas como congestión nasal, tos, amigdalitis, fiebre y mialgia. Es importante mencionar que algunos miembros de esta familia (adenovirus 40 y 41) están relacionados a padecimientos gastrointestinales, siendo responsables del 7 al 17% de todas las gastroenteritis infantiles [34, 35].

Los adenovirus pueden ser cultivados y aislados en una variedad de líneas celulares e identificados por distintos métodos: la inmunofluorescencia, los inmunoensayos enzimáticos, la inhibición de la hemaglutinación y el método de neutralización, así como por técnicas de PCR.

Metapneumovirus. Como miembro de la familia *Paramyxoviridae*, este virus tiene un genoma de RNA cadena sencilla. Existen dos variantes (A y B) que infectan humanos, y se ha encontrado que un alto porcentaje de la población mayor a 5 años es seropositiva para alguna de las dos variantes, las cuales han circulado en la población por lo menos durante los últimos 50 años [36]. La prevalencia anual varía año con año, con epidemias periódicas, y se ha encontrado que varias cepas de las dos diferentes variantes pueden circular de manera simultánea en la misma temporada. En zonas templadas el pico estacional de infecciones por hMPV ocurre en invierno y primavera, lo cual se traslapa con la temporada de RSV. Varios estudios de niños hospitalizados por infecciones respiratorias en diversas regiones han encontrado este virus en el 6% a 40% de casos [37, 38, 39]. El promedio de prevalencia en la mayoría de las poblaciones pediátricas es de 5 al 10%, aunque puede ser mayor en los meses pico.

El virus causa una variedad de síntomas, principalmente rinorrea, tos y fiebre, aunque los pacientes también pueden presentar vómito y diarrea. Entre el 5 y el 17% de las infecciones con metapneumovirus se han encontrado coinfecciones con otros virus, principalmente con RSV. El método diagnóstico clásico para estos virus es el uso de PCR.

Coronavirus. Pertenecen a la familia *Coronaviridae* y poseen un genoma de RNA de cadena sencilla de polaridad positiva. Cuatro tipos de coronavirus (HCoV-229E, HCoV-HKU1, HCoV-NL63 y HCoV-OC43) se han asociado con diferentes problemas respiratorios [40]. La cepa 229E del grupo I y la cepa OC43 del grupo II han sido los más estudiados, ya que los coronavirus HCoV-NL63 y HCoV-HKU1 fueron descritos más recientemente. En otoño del 2002 apareció en la provincia de Guandong, China un nuevo virus respiratorio. La enfermedad se asociaba a hipoxia y falla respiratoria rápida y fue designada como síndrome respiratorio agudo severo (SARS). Se encontró que un coronavirus nuevo era la causa del SARS [41, 42]. Entre noviembre de 2002 y julio del 2003, el SARS infectó a más de 8000 personas y causó 800 muertes en 32 países. No hay evidencia serológica de infecciones previas en personas sanas, lo que sugiere que el SARS es un virus emergente en los humanos. La manifestación primaria es fiebre y neumonía con alteración pulmonar difusa e hipoxia importante. Otras manifestaciones son diarrea, linfopenia y alteraciones hepáticas. La mortalidad en adultos es de 10% a 17% [43]

mientras que la enfermedad es mucho mas leve en niños. El método diagnóstico clásico para coronavirus es por PCR.

5. Virus de Influenza

Es importante mencionar el caso particular de los virus de influenza, dado que la enfermedad que causa es considerada una de las principales afecciones respiratorias causadas por virus. Cada año la influenza afecta alrededor de 12% de la población mundial, es decir, alrededor de 500 millones de personas [2]. Sin embargo, las complicaciones generadas por la enfermedad ocasionan cerca de 500 mil muertes anuales, principalmente en niños pequeños, adultos mayores y en la población bajo condiciones médicas especiales [2].

Los virus de influenza se clasifican en tres tipos: A, B, y C, que corresponden a tres géneros de la familia *Orthomyxoviridae*. Los virus de influenza tipo A infectan una amplia variedad de aves y mamíferos, incluyendo humanos, y se dividen en subtipos dependiendo de las glicoproteínas presentes en la superficie del virus, la hemaglutinina (H) y la neuraminidasa (N). Los virus de influenza B infectan humanos principalmente y, dado que sólo se conoce un subtipo, tienen un bajo potencial pandémico, aunque pueden provocar cuadros severos. Los virus de influenza C causan padecimientos moderados en humanos y algunos animales, pero han sido poco estudiados. Debido a su potencial epidémico y pandémico, la atención de las organizaciones internacionales y nacionales para protección de la salud está puesta en el estudio y monitoreo de los virus de influenza tipo A.

El genoma del virus influenza tipo A consta de alrededor de 13,500 nucleótidos repartidos en 8 segmentos de cadena sencilla de RNA de polaridad negativa, que codifican para 11 proteínas: dos proteínas de superficie con actividad de hemaglutinina (H) y de neuraminidasa (N); una nucleoproteína (NP); dos proteínas de matriz (M1 y M2); una proteína no estructural (NS1), la proteína de exportación nuclear (NS2 o NEP) y tres unidades de la RNA polimerasa viral (PA, PB1 y PB2).

El virus infecta la mucosa nasal, la faringe, los bronquios y los alvéolos pulmonares, teniendo como consecuencia directa que estos tejidos se inflamen y necrosen. Los síntomas de la enfermedad van desde dolor de cabeza y garganta, congestión nasal y tos, hasta fiebre aguda. Las infecciones por influenza se pueden complicar con infecciones

bacterianas secundarias resultando en neumonía.

El virus de la influenza tipo A tiene la particularidad de presentar una gran cantidad de variantes antigénicas en sus 2 proteínas externas. De la hemaglutinina existen 16 variantes (H1 a H16); y de la neuraminidasa 9 diferentes subtipos (N1 a N9). Lo anterior resulta en 144 posibles combinaciones de subtipos de proteínas H y N que potencialmente pueden generar virus con capacidades distintas de infección, ya que estas proteínas son las responsables de establecer el contacto inicial con la célula hospedera. La clasificación del virus influenza A se basa en los subtipos de estas dos proteínas [44].

Existen dos mecanismos principales de variación antigénica en los virus de influenza A: la deriva génica y el cambio antigénico. El primero se refiere a cambios puntuales en la secuencia de nucleótidos que ocurren como resultado de los frecuentes errores de la RNA polimerasa del virus al replicar el genoma. La tasa de mutación es de aproximadamente un cambio por genoma copiado (7.3×10^{-5} mutaciones por nucleótido replicado) [44, 45]. Para el caso del virus de influenza tipo A, las tasas de cambio son distintas para cada uno de los segmentos del genoma, lo cual pudiera indicar la diferente presión de selección por parte del hospedero para cada proteína viral. Los genes que codifican las Hs y Ns mutan de manera más rápida que el resto del genoma.

El segundo mecanismo es el cambio antigénico, y se refiere a las llamadas “variaciones mayores” ya que implican un cambio total de un segmento génico. Este mecanismo se ve favorecido por la naturaleza segmentada del genoma de influenza A. Para que se dé una nueva combinación de genes es necesario que un hospedero se infecte con al menos dos subtipos virales distintos al mismo tiempo. Así el hospedero funciona como “tubo de ensayo” en donde pudieran producirse nuevos virus con mezclas de genes de los subtipos originales (Figura 1).

Es importante señalar que los virus de influenza tipo A presentan cierta especificidad por el hospedero que infectan. En general los virus que infectan a las aves no infectan a humanos y viceversa. Sin embargo, los subtipos de los virus que infectan aves son mucho más amplios y variables. Si bien no se han encontrado las 144 posibles combinaciones de H y N en la naturaleza, sí se han aislado todas las variantes de estas dos proteínas, conocidas hasta el momento, en aves acuáticas silvestres. Es por ello que estos hospederos son el reservorio natural de variabilidad genética para el virus de la influenza

tipo A [46] y se considera que son la fuente de los virus que son transmitidos a los demás animales, incluyendo las aves de corral [47]. Muchos de los virus de influenza que infectan aves lo hacen de manera asintomática o causan síntomas moderados; sin embargo, ciertos subtipos del virus, como H5 y H7 de alta patogenicidad, pueden causar síntomas severos y la muerte en algunas especies de aves silvestres, así como en aves domésticas, como pollos y pavos.

Por otra parte, los cerdos se infectan estacionalmente sólo con algunos subtipos del virus (H1N1, H3N2 y H1N2), ya sea de manera asintomática o con manifestaciones similares a los que se presentan en humanos, como tos, fiebre, secreción nasal, etc. [47, 56] Sin embargo se sabe que estos hospederos se pueden infectar también con cepas de influenza aviarias y de humanos [44, 46, 47]. De esta manera, si en un cerdo llegara a coincidir la infección por virus de estos dos orígenes, en principio se pueden generar virus con nuevas combinaciones de genes, incluyendo los que codifican para H y N. Estas nuevas variantes antigénicas representan un riesgo latente como causa potencial de pandemias no sólo en animales sino también en humanos [44, 46].

Este dinamismo en las variaciones del genoma viral tiene como consecuencia la posible susceptibilidad de nuevos hospederos a la infección, siendo ésta la base virológica e inmunológica que subyace en las epidemias causadas por un subtipo viral que no había afectado anteriormente a la comunidad, para el cual no hay inmunidad poblacional previa [48, 49].

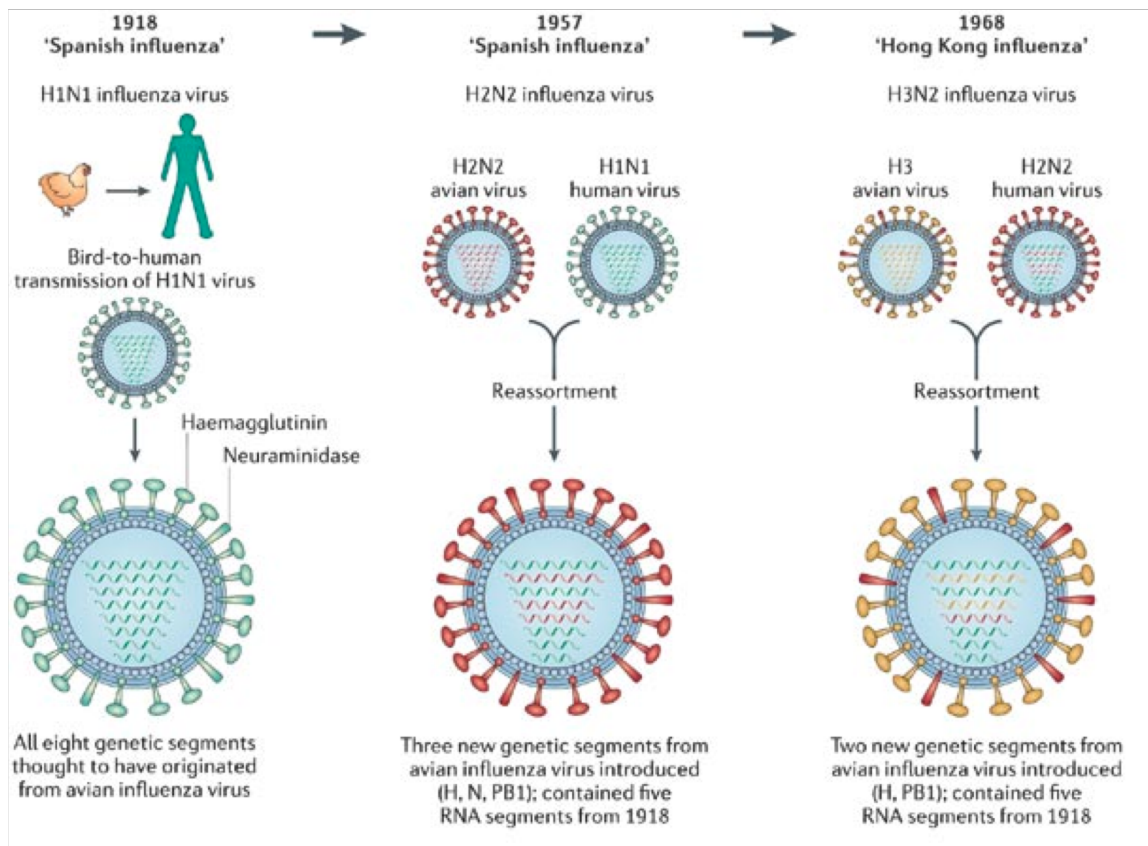
La transmisión de virus de influenza aviar entre aves y humanos puede ocasionar cuadros graves de influenza en humanos como los reportados en 1997 para una cepa de influenza A (H5N1) en Hong Kong [50]. Más recientemente, entre finales del 2003 y principios del 2004, apareció otra variante de influenza A H5N1 de alta virulencia en pollos y patos, que continúa circulando hasta la actualidad. Este virus ha sido capaz de cruzar la barrera de especie, habiendo infectado a la fecha a 553 personas en 15 países, provocando la muerte de 323, la mayoría de ellas en Asia [50]. Debido a ello se han sacrificado más de 200 millones de aves domésticas como principal medida de control de la zoonosis [46]. A pesar de esto, el mayor riesgo de dicha cepa aviar reside en la posibilidad de que recombine con alguno de los virus de influenza que circulan en humanos, como son los virus de subtipo H1N1 o H3N2, ya sea en el humano o en un hospedero intermedio como el cerdo.

De esta manera se podrían generar rearrreglantes virales con potencial de diseminación global por su transmisibilidad directa de persona a persona (capacidad que aún no ha sido adquirida por el virus H5N1), y probablemente con elevada patogenicidad por la ausencia de inmunidad previa en las poblaciones. Afortunadamente la transmisión de humano a humano en los casos descritos hasta el momento de influenza aviar ha sido documentada sólo en decenas de casos [51].

6. Pandemias de influenza

Durante el siglo pasado se introdujeron tres cambios antigénicos en los virus de influenza A circulantes en humanos, causando las siguientes pandemias: en 1918, con la aparición de un virus H1N1; en 1957, cuando el virus H1N1 fue reemplazado por un virus de subtipo H2N2; y en 1968, cuando el virus H3N2 reemplazó al subtipo H2N2 (Figura 1) [49, 52]. En 1977 reapareció el subtipo H1N1 en humanos, aunque en este último caso la variante H1N1 no resultó en una pandemia ni reemplazó al subtipo circulante H3N2, por lo que ambos subtipos (H1N1 y H3N2) circularon hasta el 2009 en todo el mundo, causando brotes estacionales cada año durante la temporada de frío [47]. Como consecuencia de la pandemia de 2009 causada por un virus H1N1 de origen porcino (ver adelante), en el 2010 circularon esencialmente dos cepas de virus de influenza en humanos, el virus estacional H3N2 y la cepa H1N1 pandémica. En los siguientes años podremos saber si el virus estacional H1N1 ha sido desplazado de manera permanente por la cepa pandémica o no.

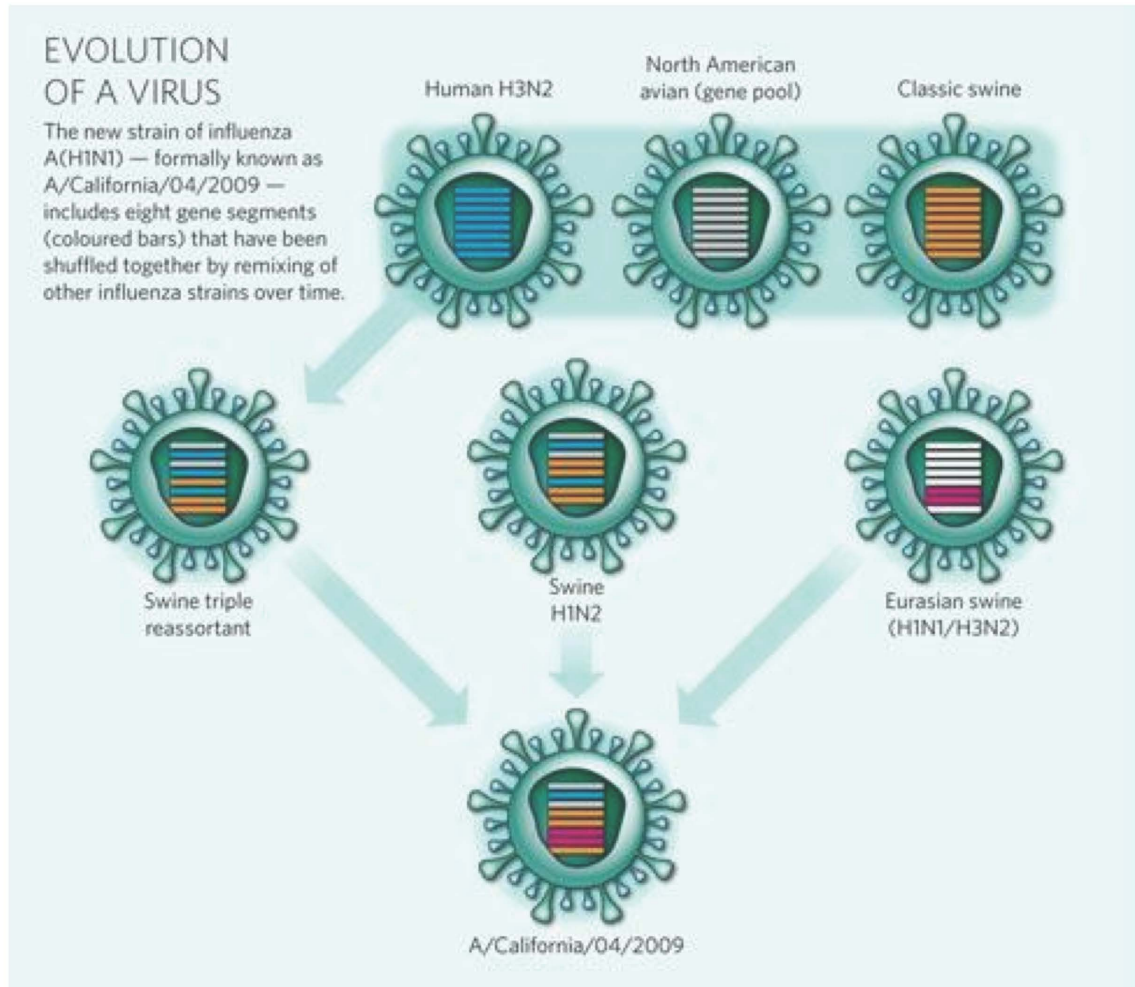
Figura 1. Orígenes de las cepas pandémicas de influenza A en humanos. La pandemia de 1918 fue causada por el virus de influenza A H1N1, que se cree tuvo sus orígenes en aves. Posteriormente la pandemia de 1957 fue causada por un rearreglo de la cepa de 1918 con una cepa aviar H2N2, por un mecanismo conocido como cambio antigénico, dando origen a una nueva cepa de subtipo H2N2 que desplazó a la de 1918. La cepa que originó la pandemia de 1968 fue el rearreglante de un virus aviar H3 y la cepa de 1957, originando un virus subtipo H3N2 que circula hasta nuestros días (Figura modificada de 52).



La pandemia del 2009 fue causada por un virus de origen porcino (Figura 2), ya que sus genes están presentes en linajes virales que han circulado en cerdos por lo menos durante la última década [49, 53, 54]. El precursor inmediato de este virus es un triple rearreglante que adquirió los genes que codifican para la neuraminidasa y la proteína de matriz del virus H1N1 parecido al aviar que circula en cerdos euroasiáticos [49, 53, 54, 57]. No se sabe el sitio geográfico ni el tiempo en que ocurrió el intercambio génico que originó esta variante, pero se considera que pudo haber circulado en cerdos por varios años antes de infectar a humanos [49]. Los reportes con los que se cuentan hasta el momento sugieren que la transmisión de cerdos a humanos se dio a finales de 2008 [49, 55].

Esta última pandemia tuvo una mortalidad baja, aunque una alta transmisibilidad, ya que para agosto de 2010 más de 214 países y territorios de ultramar habían reportado casos confirmados de la cepa pandémica A/H1N1/2009, incluyendo más de 18,500 muertes [58]. El 10 de agosto de 2010 la Organización Mundial de la Salud anunció que el evento pandémico causado por esta cepa, estaba ya en el periodo post-pandémico [59].

Figura 2. Orígenes de los genes de la cepa pandémica de influenza A/H1N1/2009. Los genes de origen clásico porcino son HA, NP, M y NS. El gen PB2 y PA tienen un origen en cepas aviarias de norte América, mientras que los genes NA y M provienen de cepas euroasiáticas porcinas [57].



7. Necesidades de nuevos métodos diagnósticos

A pesar de la gran cantidad de información disponible sobre los agentes virales clásicos causantes de infecciones respiratorias, entre el 15 y el 30% de los casos no pueden ser diagnosticados [1, 2]. Esto pudiera deberse, en parte, a un sesgo en la búsqueda de los patógenos, derivado de las limitaciones en las pruebas rutinarias de diagnóstico.

Por otro lado, existen numerosos reportes en la literatura en los cuales se describe el aislamiento, del tracto respiratorio tanto de humanos como de animales, de virus que no están típicamente asociados a enfermedades respiratorias. En muchos de estos casos no se conoce con exactitud la relación que pudieran tener con padecimientos respiratorios. Sin embargo la exploración de la asociación de su presencia con estos padecimientos es importante para ampliar el conocimiento que se tiene sobre los agentes causales de las enfermedades respiratorias, así como de un diagnóstico temprano certero que derive en un tratamiento efectivo.

Considerando las limitaciones que tienen los métodos diagnósticos actuales para la detección de virus, la evidencia de transmisión de virus de animales a humanos (zoonosis) y la importancia para la salud pública de las infecciones respiratorias agudas; se propuso en este proyecto diseñar un método de diagnóstico que pudiera explorar la presencia de un número amplio de patógenos asociados a enfermedades respiratorias. Este objetivo se cumplió, generando así un valioso elemento de monitoreo de las infecciones virales más frecuentes y posibles nuevas variantes de estos agentes causales de pandemias en vertebrados.

III. HIPÓTESIS

Es posible diseñar e implementar un microarreglo que detecte a los patógenos virales asociados a enfermedades respiratorias, y que permita subtipificar directamente los virus de influenza A.

IV. OBJETIVOS

OBJETIVO GENERAL

Diseñar, implementar y validar una metodología para la detección, caracterización y subtipificación de virus asociados a enfermedades respiratorias por medio de un microarreglo.

OBJETIVOS PARTICULARES

- 1) Diseñar oligonucleótidos específicos para detectar, a nivel de especie, los virus asociados a enfermedades respiratorias en vertebrados.
- 2) Diseñar oligos que permitan determinar el subtipo de las 16 hemaglutininas y las 9 neuraminidasas del virus de influenza A.
- 3) Imprimir el microarreglo con los oligonucleótidos seleccionados.
- 4) Validar el microarreglo con cepas de referencia y muestras clínicas.

IV. MATERIALES Y MÉTODOS

1. Selección de patógenos virales a incluir

Se incluyeron todos aquellos virus que se han aislado del tracto respiratorio de animales vertebrados, sin importar el principal padecimiento al cual se asocian comúnmente (Anexo I).

2. Obtención de secuencias

Se descargaron las secuencias nucleotídicas disponibles hasta octubre de 2007 en la base de datos de NCBI, para los 16 tipos de hemaglutinina (H) y las 9 variantes de neuraminidasa (N) de influenza A en la sección específica para este virus (www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). No hubo restricciones para el organismo hospedero, ni para la región o el país de los que se obtuvieron las secuencias; la única condición para incluirlas en el análisis fue que se tratara de secuencias génicas completas. Se estableció este requisito puesto que hay una gran cantidad de secuencias parciales depositadas en las bases de datos, las cuales generaban ruido en el análisis posterior, haciéndolo más complicado.

En cuanto al resto de los virus seleccionados, diferentes de influenza, se descargaron todas aquellas secuencias disponibles, ya fuera de ciertos genes o del genoma, la única condición fue que se tratara de secuencias completas.

3. Selección de secuencias.

Para el caso de las hemaglutininas y neuraminidasas de influenza tipo A, dada la gran cantidad de secuencias disponibles en las bases de datos y la redundancia existente entre ellas, se siguió la siguiente estrategia para filtrar las secuencias:

a. Construcción de árboles filogenéticos

Con la base de datos de influenza se alinearon las secuencias con el programa ClustalX [60], usando los parámetros por defecto. Estos alineamientos se

analizaron con el programa MEGA 3.0 [61] generando árboles filogenéticos, con 1000 pseudo réplicas de bootstrap, bajo Neighbor Joining [62], asumiendo el modelo Tamura Nei [63].

b. Selección de secuencias por porcentajes de identidad

A partir de estos árboles se seleccionaron aquellas secuencias que presentaran un nivel de identidad menor o igual al 90% a nivel de nucleótidos; esto con la finalidad de eliminar aquellas secuencias que proporcionaran información redundante. Para ello se realizó un BLAST de 2 secuencias (www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi) entre cada par de secuencias mostradas en el árbol filogenético como más cercanas. Además de buscar aquellas cepas que tuvieran al menos un 10% de divergencia, buscamos quedarnos con las más recientes, las que tuvieran secuencias más largas y que pertenecieran a regiones geográficas cercanas a México. Este análisis se hizo de forma manual para todas las secuencias existentes de H y N reportadas hasta octubre de 2007.

Para el resto de las especies virales no fue necesario el análisis filogenético, ya que el número de secuencias reportadas para esa fecha era del orden de decenas, por lo que se trabajó con todas las secuencias génicas completas que estaban disponibles.

4. Creación de setentámeros

Una vez que se seleccionó el juego de datos no redundante de H y N, y las secuencias del resto de virus de interés, se generaron oligonucleótidos de setenta bases a partir de las secuencias elegidas. Para esto se utilizó una ventana de 70 nucleótidos, la cual se fue recorriendo de 5 en 5 posiciones a todo lo largo de las secuencias completas.

5. Selección de oligonucleótidos

Todos los setentámeros generados fueron sometidos a un BLASTN [64] con los parámetros por defecto, contra la base de datos no redundante de NCBI, mientras que para el caso de la subtipificación de influenza se hizo el BLAST contra una base que únicamente

contenía las secuencias de H o N virales dependiendo de qué oligos se estuvieran analizando. Esta base contiene los datos iniciales con los que se construyeron los árboles filogenéticos. El resultado del BLAST de cada oligo fue analizado para determinar su especificidad de reconocimiento. Esta característica fue definida en base a que los mejores hits pertenecieran a la misma especie viral, o al mismo subtipo (para el caso de H y N de influenza A) del cual fueron diseñados. Además se determinó que el alineamiento tuviera una cobertura de al menos 85% (60 de los 70 nucleótidos) y el valor de expectancia (e-value) fuera menor o igual a 0.0001. Los oligos que pasaron este filtro los llamamos “específicos”.

Un segundo filtro para la selección de los setentámeros es la predicción de la hibridación del oligo con las secuencias recuperadas por BLAST. Lo anterior está basado en la energía libre de Gibbs (ΔG) que se calculó utilizando el método de Nearest Neighbor (NN)[65]. Por trabajos anteriores en el laboratorio se ha visto que aquellos alineamientos que tuvieran un valor de dG menor o igual a -75 kcal/mol pueden considerarse como hibridaciones potencialmente positivas y específicas [66].

6. Selección final de oligonucleótidos

Se realizó un BLASTN contra toda la base de datos no redundante de NCBI (nr) con los parámetros por defecto, para corroborar que los oligos seleccionados no alinearan con secuencias del hospedero, o con alguna otra especie viral.

Por último se diseñó un algoritmo para seleccionar un grupo final de oligos que en conjunto reconociera, por lo menos dos veces, a todo el universo de secuencias reportadas de todas las especies de interés. De este modo se asegura, teóricamente, que cada una de las secuencias reportadas en la base de datos de NCBI pueda ser reconocida por, al menos, dos oligos del microarreglo. También se buscó que los oligonucleótidos seleccionados cubrieran la mayor parte del gen contra el cual fueron diseñados, tratando de que no traslaparan entre sí.

Todos los análisis bioinformáticos se realizaron en el cluster del Instituto de Biotecnología de la UNAM campus Morelos.

7. Preparación de oligonucleótidos

Los oligos se sintetizaron químicamente por la compañía Invitrogen-Illumina, San Diego (dispuestos en placas de 384 pozos, a una concentración de 40mM). Los oligonucleótidos presentes en placas se resuspendieron en 3x SSC (0.45 M cloruro de sodio, 0.045 M citrato de sodio, pH 7) a una concentración de 50 $\mu\text{mol/l}$. Cada uno de los puntos contiene también el oligo control denominado “spike-70” sintetizado por la misma compañía (5'-ACCTCGCTAACCTCTGTATTGCTTGCCGGACGCGAGACAAACCTG AACATTGAGAGTCACCCTCGTTGTT-3') impreso a una proporción de 1:50 con respecto al oligo presente en cada uno de los pozos de la placa [11].

8. Impresión de microarreglos

La impresión de los microarreglos fue hecha por personal de la Universidad de California en San Francisco (UCSF) con el robot Linear Servo Motor Microarrayer, utilizando el programa ArrayMaker versión 2.64, en las facilidades de dicha institución. Se utilizaron puntas de silicón y portaobjetos cubiertos de poli-L-lisina. El microarreglador coloca gotas de menos de 250 pL de oligo en el microarreglo, las cuales se evaporan rápidamente, dejando los oligos unidos al vidrio por interacciones electrostáticas.

9. Post-procesamiento de los microarreglos

Los oligos impresos en el portaobjetos de vidrio se someten a un proceso de preparación previo a la hibridación. El primer paso es marcar con una pluma de punta de diamante la zona del vidrio en la que está impreso el microarreglo por la parte trasera. Esto debido a que con los pasos subsecuentes los puntos se harán invisibles y es necesario poder identificar la zona en la que está impreso el microarreglo. Posteriormente, para unir los oligos de manera covalente al vidrio, se entrecruzan mediante un pulso de luz ultravioleta de 60mJ. A continuación los portaobjetos se sumergen en una solución salina de citrato de sodio (SSC) 3X y duodecil sulfato de sodio (SDS) 0.2%, durante 5 minutos a 65°C. Se enjuagan dos veces con agua para eliminar el exceso de SDS y posteriormente se lavan con etanol al 95%, se secan por centrifugación por 5 minutos a 50 x g.

Los portaobjetos se sumergen a continuación en una solución con 5.5 gramos de anhídrido succínico (Sigma) disueltos en 15 ml de 1M de borato de sodio a pH 8.0 y 355 ml de 1-metil-2-pirrolidinona (Sigma) por 15 minutos. Lo anterior con la finalidad de bloquear las aminas expuestas de la poli-L-lisina del vidrio, a través de la unión del anhídrido succínico a las cargas positivas libres. Posteriormente los microarreglos se lavan una vez con agua, una última vez con etanol al 95% y se secan por centrifugación durante 5 minutos a 50 x g. Estos pasos deben hacerse lo más rápido posible para evitar que los vidrios se sequen antes de la centrifugación y por tanto queden manchados. Es importante mencionar que este paso de bloqueo debe realizarse en una campana de extracción puesto que la pirrolidinona es un compuesto tóxico. Una vez procesados los vidrios se guardan en un desecador para evitar que se humedezcan. Las laminillas pueden permanecer almacenados hasta por 3 meses antes de ser hibridados.

10. Virus utilizados para validar el microarreglo

La Dra. Blanca Lilia Barrón del Departamento de Microbiología de la Escuela Nacional de Ciencias Biológicas del Instituto Politécnico Nacional, proporcionó lisados de células infectadas con adenovirus humano 6 (HuAdV6 ATCCVR-6) y sarampión (cepa vacunal Edmonton-Zagreb). El Dr. Julio Santiago Cruz, del Departamento de Inmunogenética y Alergia del Instituto Nacional de Enfermedades Respiratorias (INER) facilitó RNA purificado de coronavirus SARS. En el laboratorio se cuenta con cepas de referencia de virus influenza B, influenza estacional humana H1N1 (A/New Caledonia/20/99) y H3N2 (A/Panama/2007/99), una cepa pandémica de H1N1 2009 (A/Mexico/IBt-1/2009(H1N1)), parainfluenza humana tipo 2 y 3 (HPIV2, HPIV3). Por otra parte, el Doctor Robert G. Webster, del Departamento de Enfermedades Infecciosas del Hospital Pediátrico de Saint Jude en Memphis Tennessee, donó una colección de muestras de RNA genómico de 33 cepas de influenza crecidas en huevo, que representan los 16 subtipos de H y los 9 de N.

11. Muestras clínicas de humanos

Se cuenta con 40 muestras nasofaríngeas (colectadas con hisopos BBL™ CultureSwab™) de pacientes pediátricos con enfermedad respiratoria que acudieron a consulta externa al Hospital de Pediatría del Centro Médico Nacional Siglo XXI del Instituto Mexicano del Seguro Social (CMN-IMSS), así como a un consultorio privado en el estado de Veracruz, que presentaban síntomas respiratorios como tos, escurrimiento nasal, ardor de garganta, etc.

La presencia de virus respiratorios en estas muestras se detectó utilizando el kit “RV15 OneStep ACE Detection” de la compañía Seegene el cual detecta de manera específica 15 virus asociados a enfermedades respiratorias (adenovirus A/B/C/D/E, coronavirus 229E/NL63 y OC43, parainfluenza tipo 1,2,3 y 4, rinovirus A/B/C, influenza tipo A y B, virus sincicial respiratorio tipo A y B, bocavirus 1/2/3/4, enterovirus). Los ácidos nucleicos presentes en las muestras se amplificaron por RT-PCR de acuerdo al protocolo indicado por la compañía y los fragmentos de DNA amplificados se analizaron en geles de agarosa al 2%.

12. Virus de influenza aislados de cerdos

Por otra parte el Dr. Iván Sánchez de la Facultad de Medicina Veterinaria y Zootecnia (FMVZ) del Departamento de Producción Porcina de la UNAM, proporcionó virus de influenza A, aislados y crecidos en embrión de pollo, a partir de biopsias pulmonares de cerdos con secreción nasal, tos y/o estornudo y lesión pulmonar. Las muestras fueron colectadas en rastros de 17 estados del país, de noviembre de 2009 a febrero de 2011. 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. Nos proporcionaron seis de estos virus, de los cuales sólo se pudo determinar el subtipo N1 para dos de ellas, no pudiendo identificar el subtipo de H ni de N en el resto de los aislados.

En nuestro laboratorio se hizo un RT-PCR dirigido para amplificar una región del gen M de influenza A y así comprobar la presencia del virus en las muestras. Para esto se utilizaron los oligos (F 5'-CTTCTAACCGAGGTCGAAACG-3' y R 5'-CATGCAACTGGCAAGTGCACC-3') utilizando el kit SuperScript One-Step RT-PCR Platinum Taq III de Invitrogen. Las muestras de influenza porcina se analizaron también con el microarreglo diseñado en el presente trabajo. Aquéllas que dieron positivas por microarreglo para H1N1 se analizaron por RT-PCR con oligos específicos para detectar la N1 pandémica (F 5'-AGTGGATGGGCTATATACACA-3' y R 5'-ATTTGTACGAGGCGGCCTT-3') usando el kit antes mencionado. Se corrieron geles de agarosa al 1% para observar los productos amplificados.

13. Extracción de RNA y amplificación del material genético presente en las muestras.

La extracción de RNA viral a partir de lisados celulares se hizo usando el kit PureLink™ RNA Mini Kit de Invitrogen, de acuerdo al protocolo de los fabricantes, partiendo de 200 µl de muestra inicial y eluyendo en 30 µl de agua libre de RNAsas. Una vez extraído el material genético, éste se amplificó a través de 3 rondas (Figura 4) que se describen a continuación:

Ronda A) Transcripción reversa del RNA total presente en la muestra y síntesis de DNA doble cadena. Para este paso se utilizó el cebador A (5'-CACTGGAGGATANNNNNNNNN-3') en donde N es cualquier nucleótido (A,T,C o G), lo que permite que el cebador se pegue en zonas azarosas a las moléculas de RNA presentes en la muestra. Se utilizó el kit SuperScript III Reverse Transcriptase de Invitrogen con las siguientes cantidades de reactivos por muestra: 2 µl de 5x RT buffer, 0.5 µl de 100 mM DTT, 1 µl de 12.5 mM dNTPs, 1 µl de H₂O libre de RNAsas y 0.5 µl de la transcriptasa reversa SuperScript III RT (200 unidades/µl).

Del RNA extraído se toman 5 µl (sin importar la concentración del mismo), se mezcla con 1 µl del cebador A (40 pmol/µl), se hierve por 5 minutos y se añaden 5 µl de la preparación previamente descrita. Se incuba una hora a 42°C, seguido de 10 minutos a 94°C. El siguiente paso consiste en la síntesis de DNA de doble cadena, para lo cual 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:

<p>94 °C---2 min 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</p>	}	<p>2ciclos</p>	<p>* En el primer ciclo añadir a cada muestra 5 µl de la <i>Mezcla de reactivos I</i> y en el segundo ciclo añadir 0.6 µl de la <i>Mezcla de reactivos II</i></p>
---	---	----------------	---

Mezcla de reactivos I

3.8 µl de H₂O, 1 µl de 5x Sequenase buffer, 0.15 µl de enzima Sequenasa (13 unidades/µl)

Mezcla de reactivos II

0.45 µl de Sequenase diluting buffer, 0.15 µl de enzima Sequenasa (13 unidades/µl)

Ronda B) Esta es la primera ronda de amplificación. El templado son 5µl del producto de la Ronda A. Para ésta se utiliza el cebador B (5'-GTTCCCACTGGAGGATA-3'), el cual es complementario a la parte conocida del cebador A y tiene 6 bases más en el extremo 5'. Para la reacción en cadena de la polimerasa se utiliza el kit KlenTaq LA de Sigma (mezcla de una mutante de la Taq polimerasa con una deleción 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 estándar de reactivos para amplificar una muestra es la siguiente: 5 µl de 10X buffer para KlenTaq LA, 1 µl de 12.5 mM dNTPs, 1 µl de 100 pmol/µl cebador B, 37 µl de H₂O y 1 µl de KlenTaqLA (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- ∞

Ronda C) Se toman 5µl del producto de la ronda anterior como templado, se utiliza la misma enzima de la ronda B. Las únicas variaciones son la concentración del cebador B, que para este caso es de 40 pmol/µl y los dNTPs utilizados. En este paso se emplea 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 2:3 aa-dUTP/dTTP o 7:3 aa-dUTP/dTTP. El protocolo de amplificación es igual al descrito, previamente sólo que con 20 ciclos.

Finalmente se corre un gel de agarosa 1% con los productos de PCR para confirmar la presencia de los amplicones, que se observarán como barridos de DNA con un tamaño mayor a 300 pares de bases.

14. Limpieza y marcaje del DNA

Los fragmentos de DNA amplificado se recuperaron por medio de columnas de afinidad (Zymo columns DNA Clean & Concentrator™-5, Zymo Research, que purifican segmentos de DNA de 50 pares de bases hasta 23Kbp) de acuerdo al protocolo del fabricante. El DNA se eluye en 11µl de agua.

El siguiente paso es 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). Posteriormente se preparan dos distintas mezclas de reacción. Primero, los 11µl del producto purificado de PCR de la ronda C de amplificación se mezclan con 1.5 µl de amortiguador de carbonato de sodio 1M (pH9) y 5µl del colorante resuspendido Cy5 (Amersham). Por otro lado, se prepara una mezcla de 1.5 µl del amortiguador de carbonatos de sodio 1M (pH9), 5 µl de Cy3 (Amersham), 8 µl de agua y 1 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-dT]GTTCAGGTTTGTC[AmC6-dT]CGCGTCCGGCAAGCAA[AmC6-dT]ACAGAGGT[AmC6-dT]AGCGAGGT-3'), es complementario al spike70 descrito previamente [11] 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 limpieza para eliminar el colorante no acoplado y enriquecer la muestra, utilizando las columnas de Zymo. El DNA producto

de las amplificaciones se eluye en 16 μ l de agua y la sonda “probe-70” marcada se eluye en 30 μ l. Posteriormente, se hace una cuantificación de la cantidad de fluoróforo incorporado por μ g de DNA usando un lector de absorbancia NanoDrop (ND-1000 Thermo Scientific).

15. Hibridación

Se mezclan 14.5 μ l de agua, 2.5 μ l de la preparación de Probe70 marcado con Cy5 y 15 μ l de cada muestra marcada con Cy3. A continuación se le añaden a cada muestra 6 μ l de SSC 20x pH 7, 1 μ l de 1M HEPES (ácido 2-[4-(2-hidroxietil)-1-pireracini-etanosulfónico) pH 8 y 1 μ l de SDS al 10%. Se hierven las muestras por dos minutos, se centrifugan rápidamente y se dejan enfriar antes de colocarlas sobre el microarreglo.

Se colocan los vidrios ya procesados en una cámara de hibridación (Die Tech 5005) y se pone un cubreobjetos Mseries Lifterslip (Thermo Scientific) de manera que cubra completamente el área donde se encuentra impreso el microarreglo. Entre cada cubreobjeto y laminilla se introduce por capilaridad 30 μ l de la mezcla de productos de PCR y “probe 70” marcados con los fluoróforos. Se vierten 20 μ l 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 horas.

16. 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 los vidrios se transfieren a una rejilla y se sumergen vigorosamente en esta solución durante un minuto. El segundo y tercer lavado son en una solución SSC 1X y finalmente un único lavado con SSC 0.2X. Los vidrios se secan por centrifugación durante 5 min. a 50 x g. Las laminillas fueron escaneadas en un lector GenePix 4000B utilizando en general condiciones de lectura con una ganancia de 450 para Cy3 con un poder de láser de 100, y una ganancia de 750 para Cy5 con un poder de láser de 100.

17. Análisis de datos

Las imágenes en formato TIFF producto del escaneo se analizaron mediante del software GenePix6.0.2. El programa hace un mapeo de la posición de cada oligo, 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 (nombre de los oligos, media y mediana para cada punto de la intensidad de los dos canales, ruido de cada punto de los dos canales, desviaciones estándar de los dos canales, porcentajes de saturación de cada canal, etc.)

A partir del archivo .gpr se calcula el valor normalizado de la intensidad de cada punto para el canal marcado con Cy3. Para ello se suma el valor de la media de los 5185 oligos impresos en el microarreglo, así como los valores de la media del ruido. Posteriormente el valor de la intensidad de cada oligo se divide entre la sumatoria de intensidades. De igual manera, el ruido de cada punto se divide entre la suma del ruido de todos los puntos. Finalmente se hace una resta de los dos cocientes antes mencionados. De esta manera tenemos que el valor para cada punto del microarreglo es:

$$\text{Valor normalizado} = (\text{media de señal Cy3} / \sum \text{media de señal Cy3 de todos los oligos}) - (\text{media de ruido Cy3} / \sum \text{media de ruido Cy3 de todos los oligos})$$

Todos los resultados mostrados en el trabajo están basados en este cálculo normalizado para cada uno de los puntos impresos en el microarreglo.

El principal objetivo del análisis de los datos es la identificación de aquellos oligos que tengan una señal significativamente superior con respecto al resto, para poder determinar la presencia de material genético de un virus en particular en la muestra estudiada. Esta tarea se complica debido al ruido inherente al método experimental, además del gran número de oligos que se examinan de manera simultánea en cada experimento. Hay varios métodos reportados en la literatura para analizar los resultados de microarreglos de diagnóstico [11-21] los cuales se basan principalmente en corregir las intensidades por el ruido, transformarlas a \log_2 , convertirlas a valores-z y sus correspondientes valores p , estableciendo un corte arbitrario a partir del cual todos los valores superiores a él se consideran como positivos. Otra aproximación es la comparación de los patrones de hibridación observados con respecto a lo que se observa al hibridar *in silico* cepas de referencia; incluso usando hasta pruebas no paramétricas como McNemar [67].

En este trabajo utilizamos la metodología de análisis propuesta por Breitling R. *et al* llamado “Rank Products” (RP) o productos de rangos [68], en el cual contamos con la excelente asesoría de la Dra. Rosa María Gutiérrez del Instituto de Biotecnología de la UNAM. La técnica está diseñada para microarreglos de expresión, de uno o dos canales, pero es fácilmente trasladable a microarreglos de diagnóstico. Se basa en identificar aquellos genes expresados de manera diferencial (diferentes intensidades de los oligos diagnósticos en nuestro caso) con un modelo estadístico sencillo, pero con un trasfondo biológico importante. Para calcular los valores de rango es necesario tener réplicas (al menos 3 de acuerdo a Breitling R. *et al*) de los experimentos. Este método también provee de un valor estadístico estricto, capaz de determinar el nivel de significancia para la intensidad obtenida de cada oligo, permitiendo un control flexible en la tasa de detección de falsos positivos. Esta técnica ha mostrado tener un comportamiento más confiable y consistente que las variantes no paramétricas de la prueba-t [68], principalmente porque no asume una distribución normal en los datos. Además de que los resultados obtenidos con este método son confiables aún en presencia de datos ruidosos. El modelo asume que serán pocos los oligos que darán intensidades altas, que las réplicas representan experimentos independientes, que los cambios son independientes unos de otros y que la variación en las mediciones es la misma para todos los oligos.

Brevemente, la estadística detrás del modelo de RP es la siguiente: para un experimento con n oligos y k réplicas, si ordenamos los oligos de mayor a menor intensidad, podemos decir que la probabilidad de que cierto oligo aparezca en el primer lugar/rango en cada una de las réplicas es $1/n^k$ si la lista fuera totalmente azarosa. Puesto de otra manera, es poco probable observar a un mismo oligo en las primeras posiciones de las listas de todas las réplicas si se tratara de un setentámero cuya secuencia complementaria no fue amplificada por no estar presente en la muestra analizada. De manera general, para cada oligo g en k réplicas, cada una examinando n_i oligos, uno puede calcular la probabilidad combinada correspondiente como un producto de rangos $RP_g^{up} = \prod_{i=1}^k (r_{i,g}^{up}/n_i)$, donde $r_{i,g}^{up}$ es la posición del oligo g en la lista de oligos en la i -ésima réplica ordenada de manera decreciente por la intensidad del oligo ($r^{up}=1$ para el oligo con mayor intensidad, $r^{up}=2$ para el segundo oligo de la lista, y así consecutivamente). Estos valores RP de cada oligo pueden ordenarse de acuerdo a la probabilidad de observarlos en esas posiciones de las listas por azar. Sin embargo el método va más allá, ya que busca la

probabilidad combinada de que los patrones de hibridación sean similares para todas las réplicas del experimento. Hace también estimaciones basadas en permutaciones, puesto que un mismo valor de RP puede resultar de la multiplicación de más de un par de posiciones; este paso origina un valor similar al e-valor de BLAST [64]. Es a partir de este parámetro que se puede estimar la tasa de falsos positivos (false discovery rate -FDR) lo que permite asignar un valor de significancia a cada uno de los oligos.

Los datos mostrados en el trabajo son todos aquellos que, a partir de su valor normalizado, se sometieron a análisis de producto de rangos (RP) y cuya tasa de falsos positivos (FDR) fue igual a cero.

18. Análisis de agrupamiento (clustering)

Para presentar los resultados de los microarreglos de manera gráfica se utilizaron los valores normalizados de las intensidades de los oligos, redondeadas a su cienmilésimo más cercano. Estos datos se sometieron a algoritmos de agrupamiento con el programa Cluster 3.0 [69] normalizando los oligos y agrupándolos jerárquicamente de manera central. El archivo de salida de este programa fue visualizado en el programa Java TreeView.

VI. RESULTADOS Y DISCUSIÓN

1. Selección de los virus a incluir en el microarreglo

Se hizo una revisión de la literatura para seleccionar los virus asociados a enfermedades respiratorias. Se incluyeron todos aquellos que hubieran sido aislados del tracto respiratorio de animales vertebrados. De esta manera se obtuvo una lista de 189 especies virales, correspondientes a 51 géneros comprendidos en 15 familias (Anexo I).

Según el Comité Internacional de Taxonomía Viral (ICTV por sus siglas en inglés) una especie viral normalmente puede estar representada por un grupo de cepas de diferentes orígenes, o una población de cepas de un origen en particular, que tienen en común un conjunto determinado de propiedades estables que los relaciona entre sí y que los separa de otro grupo de cepas [70].

2. Obtención de las secuencias de genes o genomas virales

Las información con la que se comenzó este análisis fue descargada de la base de datos de influenza (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>). Para el caso particular de influenza tipo A, en octubre de 2007 estaban disponibles 3009 secuencias de hemaglutininas (H), y 4121 secuencias de neuraminidasas (N) (Tabla 3). En cuanto al resto de las especies virales seleccionadas se descargaron las secuencias génicas completas disponibles en la base de datos no redundante de NCBI a la misma fecha.

3. Selección de las secuencias

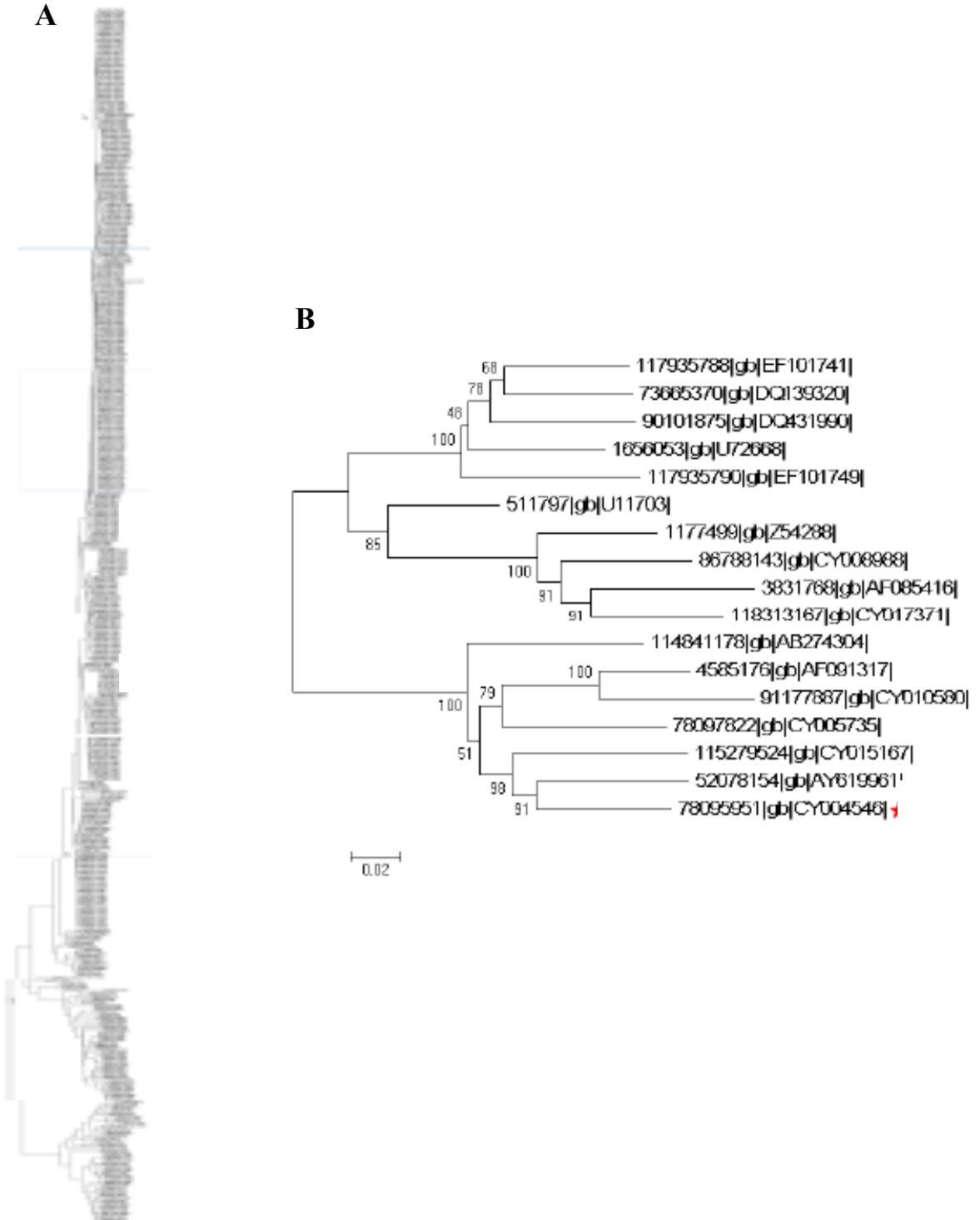
Dada la gran cantidad de secuencias disponibles para el virus de influenza tipo A y la redundancia que éstas presentan, manejar todas las secuencias de un subtipo como un grupo no parecía ser la mejor estrategia. Por otra parte se tienen reportes de que la diversidad entre las secuencias de las cepas de influenza que han circulado desde 1918 a la fecha es significativa; por ejemplo para el caso de H1 la divergencia a nivel de nucleótidos puede ser de hasta cerca de un 30% [44]. Por ello era necesario recuperar un subgrupo de datos no redundantes que fuera representativo de la divergencia dentro de cada subtipo.

Para lograr esto se alinearon las secuencias iniciales con el programa ClustalX [60] con los parámetros por defecto y se obtuvieron árboles filogenéticos por Neighbor Joining [62]. El modelo que se asumió para los árboles fue el de Tamura Nei [63] dado que las proporciones de purinas y pirimidinas de las secuencias utilizadas no eran iguales. Las pseudo réplicas de bootstrap se hicieron para estimar un error de muestreo y poder tener idea de qué tan bien fundamentadas estaban las topologías.

Una vez que contamos con los árboles comenzamos a seleccionar de forma manual aquellas secuencias que tuvieran una identidad menor o igual al 90% a nivel de nucleótidos. Cabe mencionar que el corte en el porcentaje fue arbitrario. Se realizó un BLAST de 2 secuencias entre cada par de secuencias mostradas en el árbol filogenético como más cercanas. En caso de que dos de ellas tuvieran una identidad mayor al 90% nos quedábamos con la que fuera más larga, más reciente, y/o que fuera aislada de una zona geográfica cercana a México.

En la Figura 3A aparece el árbol inicial para H1, reducido en tamaño, con sus 522 secuencias. Es importante señalar que la estructura tan vertical que se observa en este árbol es típica de las secuencias de influenza [71, 72]. En el panel B de la misma figura vemos el árbol de las 17 secuencias que resultaron de la selección basada en identidad menor o igual al 90% a nivel de nucleótidos.

Figura 3. Árboles filogenéticos para H1. Las topologías se construyeron con alineamientos obtenidos en ClustalX. Se generaron con el programa MEGA 4.0 por Neighbor Joining asumiendo el modelo Tamura Nei con 1000 pseudo réplicas de bootstrap (expresadas en porcentajes). En **A** se muestra el árbol filogenético del juego inicial de las 522 secuencias de H1 reportadas hasta octubre de 2007 en la base de influenza de NCBI. En **B** observamos la filogenia de las 17 secuencias no redundantes para este subtipo de hemaglutinina.



Esto mismo se hizo para todos los diferentes subtipos de hemaglutinina y neuraminidasa. A este juego de datos lo llamamos secuencias finales (Tabla 3). Después de este procedimiento la base de datos se redujo a 107 secuencias para H y 94 para N.

Tabla 3. Secuencias y oligonucleótidos seleccionados de hemaglutininas y neuraminidasas de influenza tipo A.

Hemaglutinina	No. secuencias iniciales*	No. secuencias finales	Oligos generados	Oligos “específicos”	Oligos finales
H1	522	17	5722	3040	8
H2	75	7	2381	914	6
H3	737	9	3000	1869	9
H4	75	9	2947	1615	6
H5	873	8	2719	1096	10
H6	177	11	3667	1954	6
H7	206	13	4323	2030	10
H8	10	1	327	199	6
H9	166	11	3629	2364	6
H10	26	6	1978	1148	7
H11	23	3	1006	609	9
H12	15	2	646	413	6
H13	15	4	1351	366	7
H14	75	2	665	331	6
H15	7	1	339	130	6
H16	7	3	1014	396	6
TOTAL	3009	107	35,714	18,474	114

Neuraminidasa	No. secuencias iniciales *	No. secuencias finales	Oligos generados	Oligos “específicos”	Oligos finales
N1	1311	20	5450	3729	7
N2	2369	30	8175	6106	6
N3	110	9	2431	2030	6
N4	20	5	1366	622	6
N5	38	3	821	496	6
N6	86	6	1655	1413	6
N7	37	6	1652	1199	6
N8	114	11	3001	1847	6
N9	36	4	1114	720	6
TOTAL	4121	94	25,665	18,162	55

*Las secuencias iniciales de H y N fueron descargadas de <http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html> en octubre 2007

En cuanto al resto de los virus a incluir en el microarreglo, la disponibilidad de secuencias génicas completas no es tan amplia como para el virus de influenza A. De estos virus se obtuvieron las secuencias completas de todos los genes disponibles, o genomas si era el caso. En general el número de secuencias por gen es del orden de decenas. Debido a lo anterior no se requirió de un proceso de eliminación de redundancia, por lo que se continuó el análisis con el grupo de secuencias obtenidas directamente de la búsqueda.

4. Generación de setentámeros a partir de las secuencias virales seleccionadas

A partir del grupo de secuencias no redundantes de H y N, y las disponibles para las otras especies virales, se generaron oligonucleótidos de 70 bases de longitud. Para ello se fragmentó cada secuencia en ventanas de 70 bases que se fueron recorriendo de 5 en 5 posiciones a todo lo largo de la secuencia. Optamos por recorrer la ventana en 5 posiciones ya que se ha reportado que al menos con este número de desapareamientos continuos, en los extremos 5' y 3' del oligo la señal de hibridación se ve disminuida significativamente [13]. Trabajos previos en el laboratorio, en donde se recorre la ventana de base en base y cada 3 nucleótidos, si bien exploran un mayor número de oligos que cubren toda la secuencia, arrojan resultados similares a los obtenidos en este análisis [Martínez-Mercado M. 2011, UNAM IBt, comunicación personal].

Una vez que contamos con este universo de oligos era necesario calificarlos para seleccionar únicamente aquellos que fueran específicos para cada virus o para cada subtipo de H y de N de influenza A.

5. Selección de oligonucleótidos específicos

Todos los setentámeros obtenidos se sometieron a un BLASTN [64] contra la base de datos inicial de hemaglutininas, neuraminidasas o no redundante, según fuera el caso. Se analizó cada resultado de BLAST para definir a aquellos oligos específicos. Los criterios para esta clasificación fueron que los mejores hits pertenecieran al mismo subtipo de proteína a partir del cual habían sido diseñados, que tuvieran una identidad de al menos 60 nucleótidos de los 70 posibles y que el valor de expectancia (e-value) fuera menor o igual a

0.0001. Con estos filtros se redujo el número de oligos a casi la mitad, con los que continuamos el análisis.

El siguiente parámetro de filtrado que utilizamos se basó en el cálculo de la energía libre de Gibbs (ΔG), estableciendo el corte en -75 kcal/mol. Este punto de corte se determinó por un trabajo previo [66] en el que se eligieron 1000 setentámeros al azar, a los cuales se les calculó el valor para ΔG con su secuencia complementaria exacta; posteriormente se generaron de uno hasta 20 desapareamientos aleatorios en posiciones azarasas y en cada caso se calculó el valor de ΔG . Los resultados previos [73, 74] permitieron concluir que valores para ΔG inferiores a -75 kcal/mol presentan una buena predicción teórica para la hibridación específica de un oligonucleótido con su secuencia complementaria, siendo éste nuestro principal indicador de una potencial hibridación experimental positiva.

6. Selección final de oligonucleótidos para el microarreglo

Como último paso realizamos un BLASTN contra toda la base de datos de NCBI para descartar que los oligos tuvieran secuencias similares en los genomas de los hospederos y/o de otras especies virales. Al final se seleccionó un grupo de oligos que en conjunto reconocieran al menos dos veces el universo de secuencias disponibles en la red para cada gen de las distintas especies virales. Se buscó también que no hubiera traslape entre ellos y que cubrieran diferentes regiones del gen contra el cual fueron diseñados.

Terminamos con un conjunto de 1883 oligos para la detección de las 189 especies virales asociadas a enfermedades respiratorias. Dentro de éstos, 114 oligos son para la subtipificación de H y 55 para la de N (Tabla 3). Estos 1883 oligos se imprimieron en laminillas de vidrio en el laboratorio del Dr. Joseph DeRisi de la Universidad de California en San Francisco (UCSF) y se trataron para hibridación de acuerdo a lo descrito en materiales y métodos.

Cabe mencionar que se imprimieron además 3301 oligos del grupo de San Francisco, que están diseñados con una metodología distinta a la descrita en este trabajo [11, 12], basada primordialmente en la comparación de patrones de hibridación de las muestras analizadas con hibridaciones teóricas de cepas de referencia. De manera general,

la principal diferencia con respecto a nuestro diseño es que un oligo de UCSF es capaz de reconocer a varias especies virales al mismo tiempo, siendo menos específicos para una especie viral determinada. El conjunto de setentámeros del laboratorio de UCSF que se imprimieron en los microarreglos, está diseñado para la identificación de virus y bacterias asociadas a enfermedades respiratorias y gastrointestinales. Es importante hacer notar que, a pesar de que los dos juegos de oligonucleótidos (los de UCSF y los de nuestro laboratorio) comparten especies virales a detectar, las secuencias de los oligonucleótidos fueron distintas entre un diseño y otro, por lo que no hay redundancia a nivel de secuencia, sino complementación de regiones del genoma a detectar, así como de los subtipos virales incluidos en el microarreglo. Sin embargo, los resultados obtenidos con los oligos de UCSF no se discutirán en este trabajo, sólo se mostrarán en algunos experimentos, dado que su análisis es distinto y complicado con respecto al aquí propuesto. Un estudio más detallado de los resultados de los oligos de UCSF sería interesante, pero escapa de los objetivos de este proyecto.

7. Validación del microarreglo con cepas de referencia

Los virus usados para probar el diseño de nuestro microarreglo fueron: adenovirus humano 6 (HuAdV6), sarampión, parainfluenza humano tipos 2 y 3 (HPIV2, HPIV3), SARS, e influenza A y B. En el caso de influenza A se usó la cepa pandémica A/Mexico/IBt-1/2009(H1N1), dos cepas estacionales [H1N1 (A/NewCaledonia/20/99) y H3N2 (A/Panama/2007/99)] y un juego de 33 virus de influenza que representan todos los subtipos de H y de N.

El material genético de todos los virus, con excepción del virus de SARS y de las 33 cepas de influenza, de los cuales nos donaron RNA viral purificado, se extrajo a partir de lisados de células infectadas conforme a lo descrito en materiales y métodos. El material genético, se sometió a las rondas de amplificación descritas en la Figura 4, y los productos de PCR obtenidos se marcaron con el fluoróforo Cy3, se cuantificó la cantidad de DNA sintetizado así como de fluoróforo incorporado. La Tabla 4 muestra un resultado típico, para algunas de estas cepas de referencia, de la actividad específica (cantidad de fluoróforo/ μ g de DNA) que se obtiene en la reacción de acoplamiento.

Figura 4. Rondas de amplificación. Esquema de las rondas de amplificación del material genético para ser usado en la hibridación de los microarreglos.

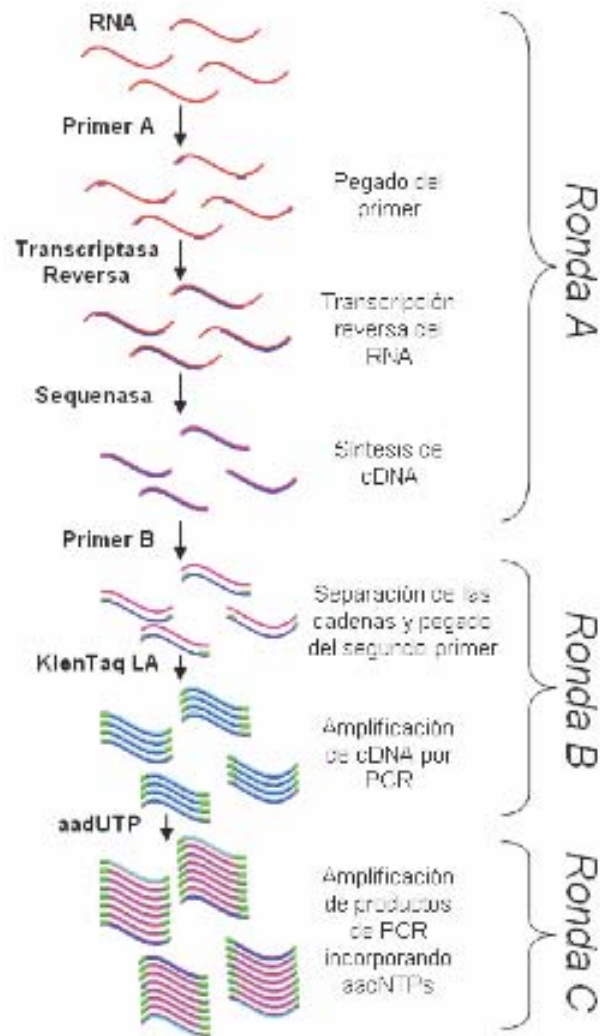


Tabla 4. Incorporación específica del fluoróforo en el DNA amplificado.

Muestra	ng DNA/ μ l >RondaC	ng DNA/ μ l >acoplar	DNArecuperado >acoplar(%)	Cy3 pmol/ μ l	pmol Cy3/ ng DNA	pmol bases (nt) incorporadas/ μ l*	pmol Cy3/ pmol nt incorporado
HuAdV6	123	37	30%	1.23	0.0332	112.12	1/91
HPIV2	111	38	34%	1.31	0.0345	115.15	1/88
HPIV3	136	43	32%	1.37	0.0319	130.30	1/95
Sarampión	136	35	26%	1.17	0.0334	106.06	1/91

*pmol bases (nt) incorporadas/ μ l = pmol de nt incorporados en dcDNA y recuperados por μ l, después de acoplar Cy3 [33 ng de nt (MW=333) = 100 pmol]

Como se puede ver en la Tabla 4, después de las rondas de amplificación y marcaje obtuvimos aproximadamente 1 molécula de fluoróforo Cy3 incorporado por cada 90 bases. Esta relación es cercana a la recomendada de 1:60 [75 -78]. Por otro lado, la cantidad total de Cy3 incorporado, del orden de 20 pmoles, está en el límite inferior de lo recomendado para usar en un experimento de hibridación, lo cual está en el rango de 20 y 200 pmoles de fluoróforo totales por microarreglo [75 -78] .

A partir de los productos purificados se continuó con el protocolo de hibridación, ilustrado de manera resumida en la Figura 5 y descrito en materiales y métodos. Posteriormente las laminillas se leyeron en el lector GenePix 4000B usando las condiciones descritas en materiales y métodos. En la Figura 6 se presenta un ejemplo de las imágenes que se obtienen al llevar a cabo los experimentos de hibridación y su posterior lectura. El panel A muestra la imagen de la hibridación de DNA amplificado a partir de RNA del parainfluenza humana tipo 3 (HPIV3), a un lado se muestra la lista de los oligos que dieron una mayor intensidad de la señal cruda. El panel B muestra la hibridación de una muestra de SARS. En la figura se observan imágenes claras, con puntos homogéneos (en rojo) para el control “spike70-probe70”, así como puntos discretos (en verde) para la señal de la muestra viral.

Figura 5. Esquema del proceso de fabricación y análisis de los microarreglos de diagnóstico.

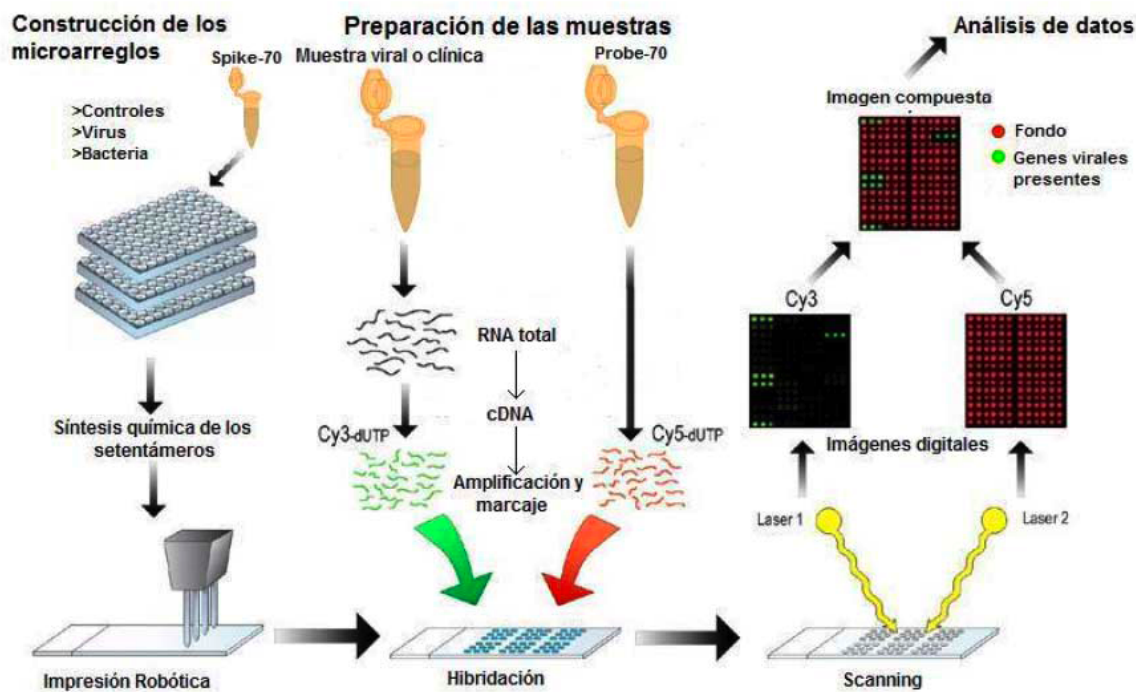
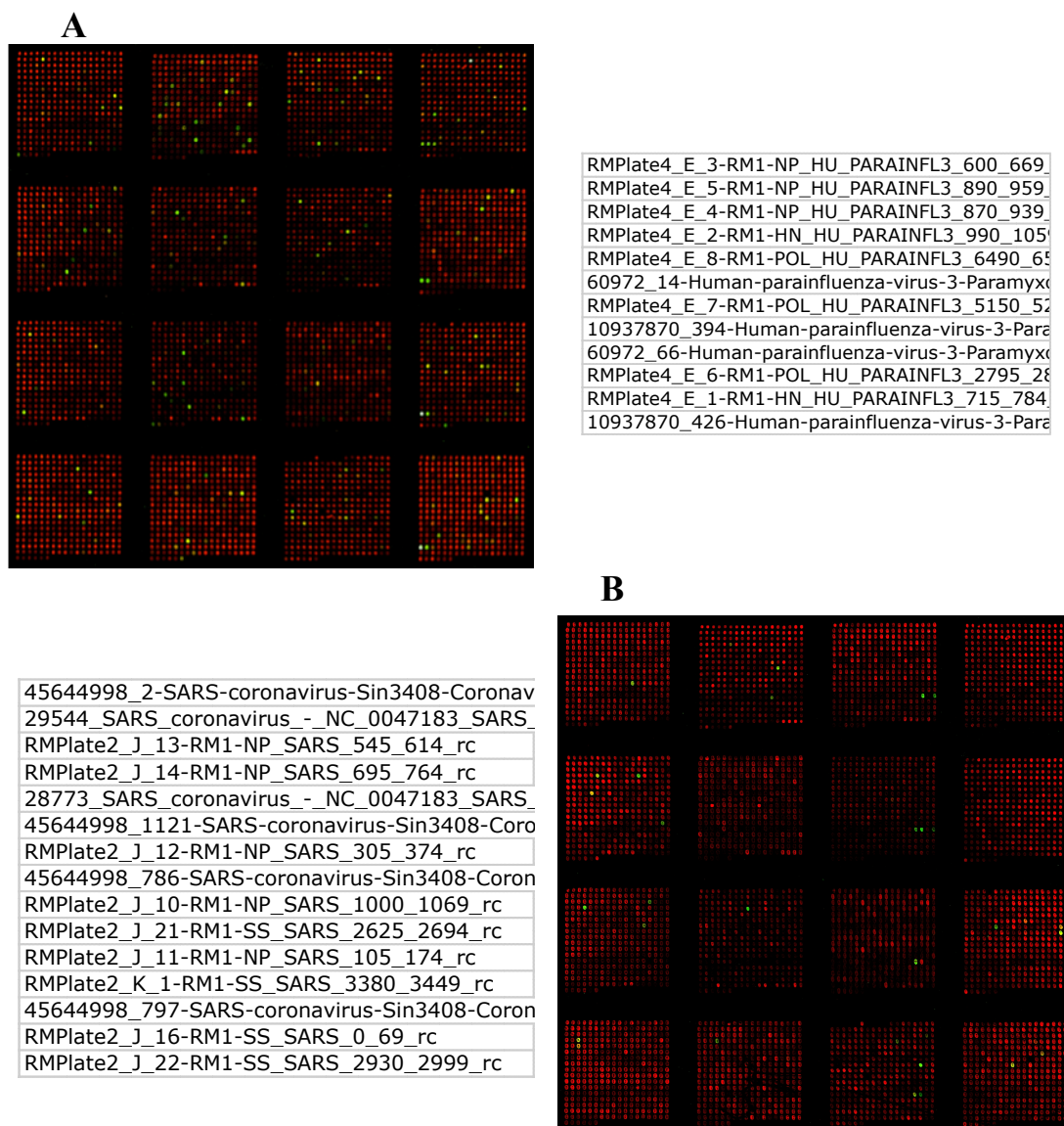


Figura 6. Imágenes obtenidas de experimentos de hibridación de microarreglos de cepas de referencia. En rojo se muestran los puntos correspondientes al control “spike70-probe70”, en verde los que dieron una señal positiva para la muestra particular. Las lecturas se hicieron con el equipo GenePix 4000B. En ambos casos la ganancia del láser para Cy3 es de 650 y para Cy5 es de 600. Junto a las imágenes están listados los oligos que dieron intensidades más altas, correspondientes a los puntos verdes observados en las imágenes. En **A** se hibridó parainfluenza humana tipo 3 (HPIV3). **B** Hibridación de una muestra de SARS.



8. Implementación de un método estadístico para determinar la reproducibilidad de los ensayos

Con el objeto de determinar la reproducibilidad de los experimentos, así como de la eficiencia de los oligos, se implementó el método de producto de rangos (RP) descrito previamente [68]. Para esto hicimos tres experimentos independientes partiendo del mismo material genético extraído. Para implementar el método se seleccionaron las cepas de adenovirus humano tipo 6 (HuAdV6), parainfluenza tipo 2 (HPIV2), tipo 3 (HPIV3) y sarampión, dado que eran las muestras de las que contábamos con una mayor cantidad de material inicial. Los resultados de cada cepa fueron normalizados y sometidos al algoritmo RP.

La Tabla 5 muestra los oligos con una tasa de falsos positivos (FDR-false Discovery rate) igual a cero, resultado del análisis de rangos (ver Materiales y Métodos). Como se puede observar para el caso de adenovirus (Tabla 5A), todos los oligos diseñados en este trabajo, que aparecen en la lista corresponden a adenovirus tipo C que es la especie a la que pertenece la cepa hibridada. En el caso de parainfluenza tipo 2 todos los oligos que diseñamos están presentes en la lista de aquellos con FDR igual a cero (Tabla 5B). Cabe señalar que hacia el final de la lista aparecen oligos para otras especies virales como rubulavirus porcino y virus sincicial respiratorio humano. Sin embargo el oligo “RMplate4_I_1-RM1-POL_PO_RUBULA_1120_1189” refiere una región de la polimerasa de esta especie viral, que pertenece a la misma familia y género que HPIV2. Dado que los genes que presentan un mayor nivel de conservación dentro de las especies de un mismo género viral son los de las polimerasas, no es sorprendente que se obtenga una señal consistente de este único oligo al hibridar material amplificado de una especie tan cercana.

Estos mismos análisis se hicieron para las cepas de PIV3 y sarampión, obteniendo resultados similares (Anexo II), es decir que todos los oligos con FDR igual a cero correspondieron a aquellos diseñados para la cepa cuyo material genético fue amplificado e hibridado. Es importante mencionar que dentro de este grupo se encuentran todos los oligonucleótidos diseñados en este trabajo para cada una de las cepas, así como otros diseñados por la Universidad de California en San Francisco.

Tabla 5. Resultados del análisis de producto de rangos de cepas de referencia Se muestran los resultados de productos de rangos de tres experimentos de hibridación independientes con las cepas indicadas. Sólo se muestran aquellos oligos con un FDR=0. Los oligos cuyo nombre comienzan con “RMPlate” corresponden a los diseñados en este trabajo, el resto es diseño de UCSF. RP es el valor de producto de rangos, FDR es la tasa de falsos positivos.

A

Adenovirus humano 6			
OLIGO	RP	E-VALUE	FDR
748_Human_adenovirus_C02_-_J019171Adenovirustyp	151.284809	5.74E-11	0
567_Human_adenovirus_C02_-_J019171Adenovirustyp	167.344699	6.46E-11	0
535_Human_adenovirus_C02_-_J019171Adenovirustyp	125.55166	3.23E-10	0
558_Human_adenovirus_C02_-_J019171Adenovirustyp	144.106153	4.59E-10	0
RMPlate1_F_14-RM1-HEXON_HU_ADENO_C_370_439	77.4563102	1.03E-09	0
RMPlate1_F_18-RM1-POL_HU_ADENO_C_935_1004_rc	68.5814125	3.16E-09	0
9632275_875-Porcine-adenovirus-A-Adenoviridae-V4 (68.5815803	3.45E-09	0
RMPlate1_F_15-RM1-HEXON_HU_ADENO_C_915_984	67.8390864	3.66E-09	0
9626158_846-Human-adenovirus-C-Adenoviridae-V4 (64.6646807	3.87E-09	0
9626158_868-Human-adenovirus-C-Adenoviridae-V4 (66.6431949	3.88E-09	0
RMPlate1_F_17-RM1-POL_HU_ADENO_C_615_684_rc	62.9792744	6.03E-09	0
9626158_863-Human-adenovirus-C-Adenoviridae-V4 (44.7763779	1.43E-08	0
9626553_707-Human-adenovirus-F-Adenoviridae-V4 (44.5746291	1.69E-08	0
9626187_866-Human-adenovirus-5-Adenoviridae-V4 (44.6344136	1.79E-08	0
9632407_785-Human-adenovirus-D-Adenoviridae-V4 (43.1061138	2.09E-08	0
RMPlate1_F_16-RM1-POL_HU_ADENO_C_100_169_rc	32.3533791	2.58E-08	0
9626158_754-Human-adenovirus-C-Adenoviridae-V4 (37.1245609	2.77E-08	0
9626187_754-Human-adenovirus-5-Adenoviridae-V4 (33.7777768	3.13E-08	0
9626553_810-Human-adenovirus-F-Adenoviridae-V4 (27.2576935	4.71E-08	0

B

virus de parainfluenza humano tipo 2			
OLIGO	RP	E-VALUE	FDR
RMPlate4_E_21-RM1-NP_HPIV_2_1240_1309_rc	146.639796	7.18E-11	0
RMPlate4_E_20-RM1-NP_HPIV_2_1120_1189_rc	132.104774	1.44E-10	0
RMPlate4_E_22-RM1-NP_HPIV_2_1360_1429_rc	147.304244	1.72E-10	0
RMPlate4_E_24-RM1-POL_HPIV-2_0_69_rc	140.421745	1.72E-10	0
RMPlate4_E_19-RM1-NP_HPIV_2_1000_1069_rc	131.70356	1.55E-09	0
26655521_nt11341-Human-parainfluenza-virus-	117.28618	1.74E-09	0
RMPlate4_F_5-RM1-POL_HPIV-2_3180_3249_rc	130.493994	2.87E-09	0
RMPlate4_F_3-RM1-POL_HPIV-2_1845_1914_rc	109.843989	6.83E-09	0
RMPlate4_E_18-RM1-NP_HPIV_2_0_69_rc	90.1458596	7.24E-09	0
19525721_498-Human-parainfluenza-virus-2-Pa	119.09004	8.29E-09	0
RMPlate4_E_23-RM1-NP_HPIV_2_145_214_rc	91.2262654	1.05E-08	0
RMPlate4_F_4-RM1-POL_HPIV-2_2150_2219_rc	96.7032934	1.55E-08	0
19525721_445-Human-parainfluenza-virus-2-Pa	87.7766746	1.75E-08	0
26655523_265-Human-parainfluenza-virus-2-Pa	67.4257585	1.76E-08	0
26655523_567-Human-parainfluenza-virus-2-Pa	81.5928802	1.80E-08	0
19525721_292-Human-parainfluenza-virus-2-Pa	95.8636155	1.94E-08	0
26655521_nt11356-Human-parainfluenza-virus-	51.4741039	2.05E-08	0
19525721_36-Human-parainfluenza-virus-2-Para	82.2307312	2.23E-08	0
RMPlate4_F_2-RM1-POL_HPIV-2_1585_1654_rc	57.0643215	2.56E-08	0
26655523_113-Human-parainfluenza-virus-2-Pa	70.5309629	3.72E-08	0
26655523_154-Human-parainfluenza-virus-2-Pa	52.7914468	4.59E-08	0
RMPlate4_F_1-RM1-POL_HPIV-2_1000_1069_rc	50.4630171	4.72E-08	0
RMPlate4_I_1-RM1-POL_PO_RUBULA_1120_1189	47.8000172	8.37E-08	0
19525721_433-Human-parainfluenza-virus-2-Pa	44.1216383	9.29E-08	0
12060914_11-Human-respiratory-syncytial-virus	33.676359	1.21E-07	0

La Tabla 6 muestra el número de los oligos diseñados en este trabajo que resultaron positivos por el método de rangos para cada una de las especies de referencia hibridadas por triplicado, así como para influenza B y SARS.

Tabla 6. Análisis de los datos de hibridación del panel inicial de cepas de referencia. *Número de oligos que dieron una señal positiva, dividido entre el total de oligos diseñados para detectar cada especie viral en particular. Para el caso de influenza B y SARS únicamente se contaron todos aquellos oligos cuyo valor normalizado fuera superior a cero. Los números para el resto de los virus corresponden a los resultados de productos de rangos.

Virus	No. Oligos *
HPIV2	12/12
HPIV3	8/8
HuAdV6	6/12
Influenza B	18/22
Sarampión	12/12
SARS	17/21

Los resultados mostrados en la Tabla 5 y 6 indican que la metodología implementada para el diseño de los setentámeros es suficientemente estricta y adecuada como para que los oligos que resultan a partir de ella hibriden únicamente con el material genético de la especie viral a partir de la cual fueron diseñados. En consecuencia, los oligos seleccionados para identificar estas especies no presentan hibridación cruzada o inespecífica, lo cual aumenta la confianza que se puede tener en los resultados obtenidos con este microarreglo.

El estudio de tres experimentos independientes de amplificación e hibridación de material genético de cepas de referencia crecidas en cultivo arrojó resultados similares en cada uno de ellos, en el sentido de que todos los oligos con un FDR igual a cero correspondieron a la especie analizada. Esto sugiere una alta reproducibilidad de los resultados, a pesar de la gran cantidad de ruido que suelen tener los microarreglos. Finalmente, el método de producto de rangos (RP) mostró ser de gran utilidad para el análisis de los datos obtenidos, aportando una metodología robusta y estadísticamente confiable para futuros experimentos.

Sin embargo no es posible hacer réplicas con todas las muestras que se pretenden examinar por microarreglos, principalmente por disponibilidad de material inicial y por costos. Por lo tanto el análisis de los datos obtenidos debe ser lo suficientemente sólido como para tener una interpretación confiable a partir de una menor cantidad de réplicas, idealmente en un solo experimento.

9. Identificación del subtipo de virus de influenza A

Con los antecedentes descritos se llevó a cabo la hibridación de una colección de 33 variantes del virus de influenza A que representan los diferentes subtipos de las proteínas de superficie del virus, H y N (Tabla 7). Esta colección de genomas de referencia está conformado por virus aislados de humanos, aves y caballos, provenientes de distintas zonas geográficas a lo largo de más de 70 años. En este análisis también se incluyeron las cepas humanas estacionales H1N1 (A/NewCaledonia/20/99) y H3N2 (A/Panama/2007/99), así como una cepa pandémica H1N1 aislada en el laboratorio (A/Mexico/IBt-1/2009(H1N1)).

A partir de los valores normalizados de las intensidades obtenidas en el análisis de cada uno de los virus se corrió el algoritmo de productos de rangos (RP). El resultado de este análisis fue un total de 477 oligos con un FDR igual a cero, dentro de los cuales 72 correspondían a oligos para la detección de hemaglutininas y neuraminidasas de diferentes subtipos, así como de oligos genéricos para influenza tipo A (estos últimos reportados por la Universidad de California en San Francisco). Esto era de esperarse pues se trataba de varias cepas de un mismo virus. Sin embargo, los 405 oligos restantes se diseñaron para reconocer virus distintos a influenza A y, por lo tanto, representan el ruido del ensayo. En otras palabras, podemos decir que estos oligos son “sucios” o “promiscuos” puesto que dan una señal intensa y de manera recurrente cuando se hibrida material genético de una especie viral distinta a su blanco. De especial interés es el hecho que dentro de esta lista, únicamente 103 (25%) corresponden a setentámeros diseñados con la metodología descrita en este trabajo, mientras que los 302 restantes (75%) son del diseño de UCSF.

El anterior grupo de 405 oligos puede considerarse como el ruido consistente presente en los experimentos de hibridación de estas muestras en particular. De manera general, estos oligos se diseñaron para la detección de la subunidad ribosomal pequeña de

bacterias de los géneros *Acinetobacter*, *Actinomyces*, *Bacillus*, *Clostridium*, etc; hongos del género *Aspergillus*; así como para rinovirus, adenovirus, metapneumovirus, calicivirus, enterovirus, sincicial respiratorio, por mencionar algunos. Desconocemos la razón por la cual estos setentámeros den señales de hibridación intensas sin importar el material genético que se esté hibridando, aunque algunas posibles razones serían contaminación de los reactivos utilizados en las amplificaciones (agua, oligos, dNTPs, etc), o incluso que estos setentámeros presenten cierta identidad con los oligos A y/o B utilizados para la amplificación de las muestras o con los concatámeros que estos mismos oligos pudieran estar generando entre si.

Los 405 setentámeros con FDR=0 que no correspondían a la identificación/subtipificación de influenza A se eliminaron del análisis posterior de los datos obtenidos para las hibridaciones de las muestras de referencia de Saint Jude. Se estableció que todos aquellos oligos cuyo valor normalizado (ver Materiales y Métodos) fuera superior a cero se tomaría como señal positiva. Este punto de corte fue definido arbitrariamente. En la Tabla 7 se muestra el conteo de oligos que dieron señal positiva, bajo este criterio, para cada una de las muestras.

Tabla 7. Análisis de los datos de hibridación de las muestras de referencia de las 16 variantes de H y 9 de N de influenza tipo A.

Subtipo	Hospedero	Origen	Fecha de colección	No. oligos Influenza A*	No. oligos HA*	No. oligos NA*
H1N1	humano	Puerto Rico	1934	18/18	6/8	4/7
H1N1	aviar	Alberta	1976	12/18	6/8	6/7
H1N1	humano	Fort Manmouth	1947	16/18	7/8	5/7
H1N1	humano	Nueva Jersey	1976	18/18	8/8	5/7
H1N1	humano	Nueva Jersey	1976	15/18	6/8	5/7
H1N1	humano	México	2009	10/18	7/8	3/7
H1N1	humano	Nueva Caledonia	1999	14/18	6/8	5/7
H1N3	aviar	Sudáfrica	1961	15/18	6/8	6/6
H1N5	aviar	Australia	1973	12/18	6/8	3/6
H2N2	humano	Singapur	1957	14/18	6/6	6/6
H2N2	aviar	Nueva York	1978	14/18	6/6	6/6
H2N2	aviar	Singapur	1957	11/18	6/6	6/6
H3N2	humano	Aichi	1962	13/18	9/9	6/6
H3N2	humano	Panamá	1999	13/18	6/9	4/6
H3N8	aviar	Ucrania	1963	9/18	8/9	4/6
H3N8	equino	Miami	1963	10/18	4/9	5/6
H3N8	aviar	Chabarovsky	1972	8/18	9/9	4/8
H4N6	aviar	Checoslovaquia	1956	9/18	4/6	5/6
H5N2	aviar	Delaware	1991	11/18	8/10	6/6
H6N5	aviar	Australia	1972	11/18	6/6	4/6
H6N2	aviar	Massachusetts	1965	8/18	4/6	6/6
H7N7	equino	Praga	1956	10/18	7/10	5/6
H7N7	equino	Praga	1956	11/18	9/10	6/6
H7N3	aviar	Nueva Jersey	1985	10/18	10/10	6/6
H8N4	aviar	Ontario	1968	8/18	6/6	6/6
H8N4	aviar	Ontario	1968	11/18	6/6	6/6
H9N2	aviar	Wisconsin	1966	9/18	6/6	6/6
H10N7	aviar	Alemania	1949	8/18	7/7	2/6
H11N6	aviar	Inglaterra	1956	9/18	6/9	5/6
H11N9	aviar	Memphis	1974	5/18	9/9	6/6
H11N9	aviar	Memphis	1974	9/18	7/9	3/6
H12N5	aviar	Alberta	1976	8/18	6/6	6/6
H13N6	aviar	Maryland	1977	10/18	7/7	5/6
H14N5	aviar	Astrakhan	1982	6/18	6/6	2/6
H15N9	aviar	Australia	1979	8/18	6/6	6/6
H16N3	aviar	Delaware	2006	5/18	6/6	3/6

En negritas se encuentran las cepas de influenza estacional y pandémica con las que se contaba en el laboratorio. * Número de oligos que dieron señal positiva, dividido entre el total de oligos diseñados para identificar el subtipo particular de HA y NA respectivamente, o para la detección del virus.

Es notorio que, en general, los resultados para las variantes del virus influenza A que se muestran en la tabla son satisfactorios para todos los subtipos, considerando que los oligos están diseñados para que cada secuencia sea reconocida al menos por dos oligos de H y N. Como se puede ver en la tabla 7, las cepas estacionales de influenza A H1N1 y H3N2 fueron detectadas de manera adecuada, al igual que la cepa pandémica A/H1N1/2009, aún cuando el diseño de los oligonucleótidos fue previo al brote.

Sin embargo, el corte establecido para el valor normalizado superior a cero, es totalmente arbitrario, y si bien permite recuperar un alto número de oligonucleótidos para la detección y subtipificación del virus influenza A, también incluye un gran número de oligos que corresponden a otros virus. Lo anterior indica que es necesario un análisis más robusto para estas muestras en las que no se cuenta con réplicas del experimento de hibridación.

10. Visualización gráfica de los resultados

Para representar de manera esquemática los resultados previamente descritos se hicieron análisis de agrupamiento con el programa Cluster 3.0 [69] usando los valores normalizados de las intensidades redondeados a su cienmilésimo más cercano. Las imágenes obtenidas con el visor JavaTreeView se muestran en la Figura 7. Los resultados muestran la especificidad de los oligos diseñados para las 16 variantes de hemaglutininas y 9 de neuraminidasas, así como de los oligos de influenza tipo A. Las tonalidades más brillantes de verde representan aquellos puntos con mayor intensidad.

Los paneles A y C de la Figura 7 muestran los resultados de los oligos diseñados para la subtipificación de H y N. Los datos se ordenaron dependiendo de la numeración de cada subtipo, tanto para los oligos como para las cepas hibridadas. Por esta disposición en el arreglo de las coordenadas se esperaba ver una diagonal verde brillante en los análisis de agrupamiento. De acuerdo a lo esperado, como se puede observar en la Figura 7, en ambos casos se presenta un patrón discreto en diagonal claramente distinguible. Sin embargo, a simple vista pueden también observarse algunos puntos fuera de la diagonal. En particular las hibridaciones de cepas de H1, muestran también una señal positiva importante para los oligos de H10 (Figura 7A). Las secuencias de estos dos subtipos son bastante lejanas entre sí [44], por lo que el análisis de alineamientos pareados de los oligos de H10 con las cepas

de H1 no dan un apareamiento significativo. El análisis de las imágenes del lector de microarreglos nos permitió notar que los oligos de H10 están dispuestos en el microarreglo de forma contigua a los de H1, lo que sugiere que probablemente hubo algún problema en la impresión del mismo. Las puntas del robot pudieron no haberse lavado bien y por tanto se “arrastró” el oligo de H1 y quedó impreso también en el punto adyacente marcado como H10, o bien que las placas en las que vienen sintetizados los oligos se hubiesen contaminado. Lo anterior sólo puede corregirse, para este lote de impresión de microarreglos, al omitir los resultados positivos de los oligos de H10 cuando se hibrida una cepa de H1. Si este fuera el caso, se espera que en una futura impresión no se observe esta “hibridación cruzada” entre los setentámeros de estos subtipos.

En el panel B de la Figura 7 se muestran los resultados de los oligos diseñados para la detección de influenza tipo A. Los oligos marcados como Influenza_A-1 hasta Influenza_A-6 fueron diseñados por UCSF para detectar H1, mientras que el resto (Influenza_A-7 hasta Influenza_A-18) reconocen el gen que codifica para la nucleoproteína (NP). Es claro que algunos oligos, como el A-15, el -16 y el -18 reconocen de manera más eficiente a las diferentes cepas de influenza, mientras que la cepa A/PR/1934(H1N1) es reconocida de manera más homogénea por los diferentes oligos. Esto indica que la variabilidad de la secuencia de las cepas a lo largo de los 70 años en que se colectaron no es menor, aún para el gen que codifica para la nucleoproteína, que es uno de los genes más conservados; pero a pesar de ello el microarreglo es capaz de detectarlas de manera eficiente.

11. Análisis de muestras clínicas

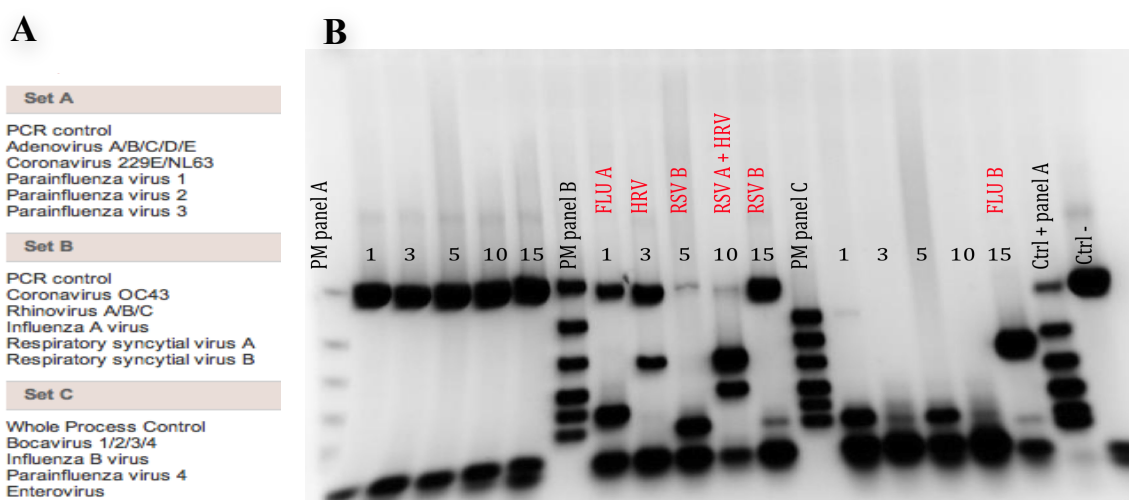
Como parte de la caracterización del microarreglo analizamos también 13 muestras de infantes con infección respiratoria aguda del Centro Médico Nacional Siglo XXI y de un laboratorio privado del estado de Veracruz, con el protocolo descrito en Materiales y Métodos, utilizado también para las cepas de referencia. Los resultados no fueron alentadores puesto que sólo una muestra dio una señal positiva débil para influenza B, la cual posteriormente fue confirmada por RT-PCR. Una segunda muestra fue positiva para sarampión pero aún no se ha analizado por otra metodología. Las otras 11 muestras no fueron claras por hibridación con el microarreglo, aunque hay que mencionar que tres de este conjunto inicial de especímenes fueron positivas para influenza A mediante RT-PCR dirigido contra el gen M del virus.

En general, todas las nuevas herramientas que se diseñan para la detección de patógenos deben ser comparadas con una técnica convencional para así evaluar su desempeño. A la metodología utilizada como referencia se le llama “estándar de oro”. Por muchos años fue el cultivo celular, aunque en la actualidad éste ha sido reemplazado por ensayos de RT-PCR, los cuales son altamente específicos y más rápidos. Por esto era necesario tener un marco de referencia para evaluar tanto los resultados obtenidos del microarreglo como la calidad de las muestras analizadas. Dado que diseñar oligonucleótidos específicos para amplificar segmentos del genoma de cada virus y estandarizarlo es una tarea ardua que se escapaba de los fines de este proyecto, se adquirió el ensayo comercial RV15 OneStep ACE Detection de la compañía Seegene, el cual detecta de manera específica 15 virus comúnmente asociados a enfermedades respiratorias. Esta técnica implica la amplificación del material genético viral a partir de 3 juegos distintos de cebadores (panel A, B, C) que contienen, cada uno, una mezcla de oligos específicos para 5 especies virales (Figura 8). Dependiendo del patrón de migración de las bandas de los amplicones para cada uno de los tres paneles de oligos es que se define qué virus está presente en la muestra.

Con este juego comercial se analizaron 4 muestras de Veracruz (indicadas como 1, 3, 5 y 10) y una del Centro Médico Nacional XXI (indicada como 15) siendo esta última la que mostró una señal positiva para influenza B en los primeros ensayos. Los productos del

RT-PCR obtenidos con el ensayo de Seegene se corrieron en un gel de agarosa, el cual se muestra en la Figura 8.

Figura 8. Análisis de muestras clínicas con método comercial para detección de virus causantes de enfermedades respiratorias. Cinco muestras clínicas proporcionadas por el Centro Médico Nacional Siglo XXI y el estado de Veracruz (indicadas como 1, 3, 5, 10, 15) se amplificaron con el juego comercial RV15 OneStep ACE Detection de la compañía Seegene. En **A** se muestra las especies virales que puede detectar cada uno de los paneles de cebadores. En **B** se presenta el gel de las muestras amplificadas, señalando en rojo la especie viral presente en cada una de ellas, dependiendo del patrón de migración de las bandas para cada panel.



Como puede observarse todas estas muestras fueron positivas para un virus, excepto las muestras 10 y 15 que presentan coinfecciones de virus respiratorio sincicial tipo A (RSVA) con rinovirus humano (HRV), y virus respiratorio sincicial tipo B (RSVB) con influenza tipo B, respectivamente.

Una vez que confirmamos la presencia de virus en las muestras clínicas, podíamos establecer que en efecto el microarreglo no los pudo detectar, muy posiblemente por una baja sensibilidad del mismo. Por esta razón nos enfocamos a aumentarla.

12. Optimización de la sensibilidad del ensayo de microarreglo

Hay varios puntos en el protocolo que pudieran ser blanco de modificaciones para aumentar la sensibilidad del microarreglo. Uno de ellos es la extracción del material genético, ya que se puede partir de un mayor volumen inicial de muestra (400 μ l en lugar de 200 μ l) y se puede eluir en un volumen menor (15 μ l en lugar de 30). Otra es modificar el número de lavados de las columnas en las purificaciones del DNA amplificado y marcado con aa-dUTP, para así encontrar las mejores condiciones para remover efectivamente los nucleótidos libres de la reacción, al mismo tiempo de no perder el DNA unido a la columna por exceso de los mismos. Pero quizá el parámetro más susceptible y probablemente el más efectivo de modificar para aumentar la sensibilidad del método es mejorar la eficiencia de incorporación del fluoróforo al DNA amplificado. Nuestra eficiencia de marcaje es de entre 1-2 pmol de cianina por μ l de muestra, mientras que en el laboratorio del Dr. C. Chiu de UCSF obtienen valores de hasta 5 pmol de fluoróforo/ μ l de muestra (comunicación personal).

Para probar el efecto de modificar los varios parámetros descritos, se tomó una muestra de virus de influenza H3N2 porcino crecido en embrión de pollo (etiquetada como muestra 51, ver más adelante) proporcionada por el Dr. Iván Sánchez de la Facultad de Medicina Veterinaria y Zootecnia. El RNA de este virus fue amplificado por RT-PCR e hibridado utilizando el protocolo original. En estas condiciones se observó una hibridación clara con los oligos específicos, con valores de lectura muy altos.

Las variaciones antes mencionadas, tanto en el volumen de muestra utilizado para extraer el material genético, como en el número de lavados de las columnas durante la purificación del DNA no arrojaron resultados diferentes a los iniciales (datos no mostrados). Sin embargo, sí logramos mejorar la incorporación de fluoróforo en el DNA amplificado, obteniendo cambios significativos en los valores de lectura de las intensidades de los oligos cuando utilizamos el protocolo modificado en comparación con el protocolo original.

En la literatura hay diversas sugerencias sobre la cantidad de fluoróforo que debe hibridarse por cada microarreglo, así como sobre la relación óptima de nucleótido modificado (aa-dUTP) y nucleótido no modificado (dTTP) [75-78]. La relación de aa-dUTP:dTTP incorporada en el DNA amplificado final puede modificarse al alterar la

concentración relativa de estos dos nucleótidos durante la síntesis de DNA. En nuestro ensayo original utilizamos una relación aminoalil-dUTP/dTTP de 2:3, sin embargo, aunque ésta es la proporción más utilizada en diversos protocolos en la literatura también se ha reportado el uso de proporciones de aminoalil-dUTP/dTTP de 3:2, 4:1 y 7:3. El objetivo es tener cuando menos un nucleótido modificado incorporado por cada 60 bases de DNA amplificado. Dado que teóricamente todos los aa-dUTPs incorporados reaccionarán con una molécula de cianina durante el paso de marcaje del DNA, es importante que los nucleótidos incorporados tampoco estén demasiado cerca uno de otro; por ejemplo, si tenemos 1 aa-dUTP por cada 10 dNTPs incorporados, se presenta el fenómeno de apagamiento (“quenching”) de la fluorescencia [75].

En relación a la cantidad de fluoróforo que debe hibridarse por cada microarreglo, se recomienda utilizar entre 20 y más de 200 pmol de fluoróforo incorporado en el DNA amplificado por cada microarreglo. Al revisar la cantidad de fluoróforo que estábamos hibridando en los experimentos anteriores, y el número de bases modificadas que obteníamos (Tabla 4), encontramos que estábamos por debajo de las recomendaciones de la literatura, o en el límite inferior. La razón por la que se obtuvieron resultados positivos en las hibridaciones previas, aún con rendimientos tan bajos de incorporación, podía deberse a que se trataba de virus crecidos en cultivo, por lo que la cantidad de material genético disponible en esas muestras era muy alta.

Como una emulación de muestras con bajas concentraciones de ácidos nucleicos, y para evaluar si las modificaciones a nuestro protocolo mejoraban la señal de hibridación, y por tanto la sensibilidad del ensayo, se hicieron diluciones 1:10, 1:100, 1:1000, 1:10000 del material genético extraído inicialmente de la muestra porcina 51. El material diluido se amplificó por las 3 rondas descritas previamente (Figura 4). El único cambio que se hizo en la ronda C fue probar dos relaciones de aa-dUTP:dTTP, la original 2:3 y una nueva 7:3. Posteriormente se marcó el DNA con Cy3 y se cuantificó la cantidad de DNA y de fluoróforo incorporado en cada una de las diluciones probadas, obteniéndose los resultados que se muestran en la Tabla 8.

Tabla 8. Cuantificación de fluoróforo incorporado en el DNA sintetizado.

Muestra	ng DNA/ μ l >RondaC	ng DNA/ μ l >acoplar	DNArecuperado >acoplar(%)	Cy3 pmol/ μ l	pmol Cy3/ ng DNA	pmol bases (nt) incorporadas/ μ l*	pmol Cy3/ pmol nt incorporado
1/1-aa2:3	105	36	34%	1.02	0.0283	109.09	1/109
1/10-aa2:3	66	25	38%	0.38	0.0152	75.76	1/200
1/100-aa2:3	84	33	39%	0.98	0.0297	100.00	1/102
1/1000-aa2:3	37	15	41%	0.31	0.0207	45.45	1/145
1/10000-aa2:3	86	34	40%	0.99	0.0291	103.03	1/104
1/1-aa7:3	103	39	38%	2.06	0.0528	118.18	1/57
1/10-aa7:3	42	17	40%	0.68	0.0400	51.52	1/76
1/100-aa7:3	85	26	31%	1.73	0.0665	78.79	1/46
1/1000-aa7:3	36	14	39%	1.06	0.0757	42.42	1/40
1/10000-aa7:3	80	32	40%	1.49	0.0466	96.97	1/65

1/1= muestra de virus H3N2 porcino número 51

aa2:3 relación molar 2:3 aa-dUTP/dTTP

aa7:3 relación molar 7:3 aa-dUTP/dTTP

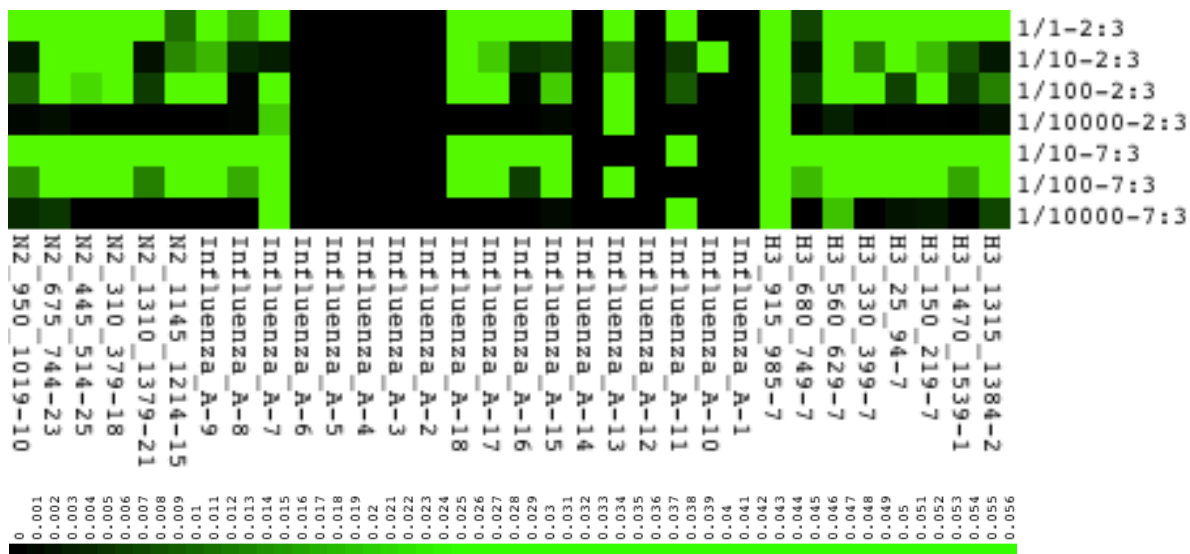
*pmol bases (nt) incorporadas/ μ l = pmol de nt incorporados en dcDNA y recuperados por μ l, después de acoplar Cy3 [33 ng de nt (MW=333) = 100 pmol]

Como se puede ver en la Tabla 8, la proporción de fluoróforo incorporado en el DNA (pmoles de Cy3, reflejo de la cantidad de aa-dUTP incorporado) en relación al DNA sintetizado (pmoles de nucleótidos incorporados) fue de 1:102 a 1:200 cuando se utilizó la relación de aa-dUTP:dTTP 2:3, mientras que con la relación 7:3 se obtuvo 1 pmol de fluoróforo incorporado por cada 40 a 76 pmoles de nucleótidos incorporados. Igualmente, la cantidad de Cy3 incorporado (pmoles de Cy3/ μ l) fue aproximadamente el doble (en las diluciones correspondientes utilizadas) en la relación nucleótido modificado: nucleótido no modificado de 7:3 en comparación con la relación 2:3.

El DNA amplificado y marcado correspondiente a las diluciones 1:10, 1:100, 1:10000 para ambas proporciones de aa-dUTP/dTTP se utilizó para hibridar en los microarreglos. Con los resultados obtenidos se hizo un análisis de agrupamiento (Figura 9), con las intensidades de la señal normalizada, como se describió previamente.

Figura 9. Análisis de agrupamiento de las ampliaciones de muestras diluidas con diferentes proporciones de nucleótido modificado.

Los valores graficados son las intensidades normalizadas de los oligos mostrados en el eje horizontal, para cada una de las diluciones indicadas en el eje vertical, con su correspondiente proporción de aminoalil-dUTP/dTTP (2:3 ó 7:3). Los datos fueron redondeados a su diezmilésimo más cercano, analizados con el programa Cluster 3.0 y visualizados con JavaTreeView. La escala de intensidades se encuentra en la parte inferior



Como podemos observar en la Figura 9, las intensidades de la señal de los oligos van disminuyendo conforme a las diluciones, dentro de una misma proporción de nucleótido marcado hasta ser casi indetectable en la última dilución. Si se comparan las diluciones 1:100 de ambas proporciones de amino-alil se ve claramente que la señal es mucho mayor para la proporción 7:3 que para 2:3. La intensidad de la dilución 1/100 con la proporción 7:3 pudiera incluso ser equiparable a la hibridación de la muestra original 1/1 con la proporción 2:3 del nucleótido modificado. Haciendo una comparación con los valores de las intensidades se puede notar un incremento en la señal de hasta 100 veces (datos no mostrados).

Por lo anterior podemos concluir que el cambio de la relación 2:3 aa-dUTP/dTTP a 7:3 aa-dUTP/dTTP incrementó la sensibilidad del microarreglo alrededor de 100 veces para muestras de virus crecidos en embrión de pollo. Con esta modificación al protocolo se retomaron las muestras clínicas mencionadas previamente, esperando poder detectar en esta ocasión a los virus presentes que se encontraron con el ensayo diagnóstico comercial de RT-PCR de la compañía Seegene.

13. Re-análisis de muestras clínicas con el protocolo con sensibilidad mejorada

Sólo se repitieron las amplificaciones e hibridaciones con las dos muestras que presentaron coinfecciones, la número 10 (positiva para rinovirus y virus sincicial respiratorio) y la número 15 (positiva para influenza B y virus sincicial respiratorio). Los resultados se muestran en la Tabla 9. En el panel A y C están los primeros 40 oligos ordenados de manera descendente conforme a la intensidad cruda de la señal de las muestras. Posteriormente se eliminaron todos aquellos oligos identificados como “sucios” (definidos previamente por el método de RP), quedando únicamente los mostrados en el panel B y D, que corresponden a los diseñados para detectar rinovirus e influenza B, respectivamente.

Tabla 9. Análisis de muestras clínicas por producto de rangos. Los paneles A y B corresponden a la muestra clínica marcada con el número 10. Los oligos diseñados para la identificación de rinovirus están marcados en rojo, los de RSV en azul y los de otras especies en negro. En A se listan los primeros cuarenta oligos ordenados de manera descendente conforme a los valores de intensidad (F532Median), así como su valor normalizado. En B están los resultados después de eliminar todos aquellos setentámeros considerados como inespecíficos de acuerdo al análisis de RP (producto de rangos).

A

Name	F532Median	Valor Normalizado
61098_13-Human-rhinovirus-2-Picornaviridae	65535	0.01368
4731888_25-Human-rhinovirus-sp.-Picornavir	63774	0.01337
9629176_21-Human-rhinovirus-16-Picornavir	55118	0.01149
Human_rhinovirus_8_FJ445113_Polymerase_I	26010	0.00530
268_Rhinovirus_8_Capsid_AY3551891_Rhino	22188	0.00450
9626735_18-Human-rhinovirus-B-Picornavirid	18038	0.00349
12060914_11-Human-respiratory-syncytial-vi	12097	0.00245
453208_13-Human-rhinovirus-sp.-Picornavirid	11961	0.00242
12060914_9-Human-respiratory-syncytial-vir	9397	0.00186
RMPlate4_H_7-RM1-NP_MUMPS_120_189_rc	7559	0.00140
RMPlate5_K_15-RM1-GlicoC_CvHV-1_1525_1	6880	0.00126
260_Sapovirus_Pol-A_-DQ366344_Sapovirus	6796	0.00123
745_Rhinovirus_36_Capsid_AY3552171_Rhino	6472	0.00117
75261032_nt296-Human-respiratory-syncytia	6400	0.00114
781_Rhinovirus_89_Capsid_AY3552701_Rhino	6052	0.00111
17939994_192-Bovine-respiratory-syncytial-v	5798	0.00100
61098_16-Human-rhinovirus-2-Picornaviridae	5414	0.00090
9629176_9-Human-rhinovirus-16-Picornavirid	5335	0.00091
75261032_nt281-Human-respiratory-syncytia	4790	0.00088
248_Sapovirus_Pol-A_-DQ366344_Sapovirus	4508	0.00083
9629176_19-Human-rhinovirus-16-Picornavir	4305	0.00073
976_Enterovirus_Human_Coxsackievirus_A20	4265	0.00075
648_Rhinovirus_95_Capsid_AY3552761_Rhino	4091	0.00064
9629176_16-Human-rhinovirus-16-Picornavir	4080	0.00060
1510_Enterovirus_B_Pol_AY8967601_RS1 (re	3866	0.00057
9626451_16-Human-rhinovirus-1B-Picornavir	3846	0.00047
9627730_25-Human-rhinovirus-A-Picornavirid	3816	0.00056
RMPlate3_O_4-RM1-GLYCO_AV_METAPNEUMOC	3795	0.00047
94_Enterovirus_Human_Echovirus_13_Capsid	3754	0.00059
12060914_7-Human-respiratory-syncytial-vir	3712	0.00059
453204_10-Human-rhinovirus-sp.-Picornavirid	3521	0.00054
RMPlate5_K_14-RM1-GlicoC_CvHV-1_1505_1	3329	0.00056
450_Human_Metapneumovirus_NL17_-_AY29	3173	0.00054
449_Human_adenovirus_D09_-_AJ8544861H	3127	0.00045
458_Human_adenovirus_D09_-_AJ8544861H	3095	0.00038
1670_HTCV-UC1_-RS1 (reversed)	3092	0.00038
254_HTCV-UC2_Human_TMEV-like_cardioviru	3058	0.00055
926_Enterovirus_71_Pol_FJ1586011_RS1 (rev	2908	0.00047
407680_9-Ovine-respiratory-syncytial-virus-P	2781	0.00038

B

Name	F532Median	Valor Normalizado
61098_13-Human-rhinovirus-2-Picornaviridae	65535	0.01368
4731888_25-Human-rhinovirus-sp.-Picornavir	63774	0.01337
9629176_21-Human-rhinovirus-16-Picornavir	55118	0.01149
Human_rhinovirus_8_FJ445113_Polymerase_I	26010	0.00530
268_Rhinovirus_8_Capsid_AY3551891_Rhino	22188	0.00450
9626735_18-Human-rhinovirus-B-Picornavirid	18038	0.00349
453208_13-Human-rhinovirus-sp.-Picornavirid	11961	0.00242
745_Rhinovirus_36_Capsid_AY3552171_Rhino	6472	0.00117
61098_16-Human-rhinovirus-2-Picornaviridae	5414	0.00090
9629176_9-Human-rhinovirus-16-Picornavirid	5335	0.00091
9629176_19-Human-rhinovirus-16-Picornavir	4305	0.00073
648_Rhinovirus_95_Capsid_AY3552761_Rhino	4091	0.00064
9629176_16-Human-rhinovirus-16-Picornavir	4080	0.00060
9626451_16-Human-rhinovirus-1B-Picornavir	3846	0.00047
453204_10-Human-rhinovirus-sp.-Picornavirid	3521	0.00054
458_Human_adenovirus_D09_-_AJ8544861H	3095	0.00038
707_Enterovirus_Human_Echovirus_24_Capsid	2541	0.00032
Neisseria_Gonorrhoeae_Genus_igaProtease_N	2533	0.00030
Actinomyces_Israelii_Genus_16SrRNA_AF479	2509	0.00021
4731874_35-Human-rhinovirus-sp.-Picornavir	2508	0.00038
199_HTCV-UC2_Human_TMEV-like_cardioviru	2498	0.00019
440_Enterovirus_Human_Coxsackievirus_B4	2455	0.00029
RMPlate3_K_23-RM1-N9_NEURO_155_224-4	2430	0.00026
9627719_18-Human-enterovirus-A-Picornavir	2422	0.00025
638_Human_astrovirus_Pol-A_-_Z25771_Hun	2395	0.00026
1597_TMEV-GDVII_TMEV-GDVII_complete_ge	2350	0.00029
RMPlate3_O_5-RM1-GLYCO_AV_METAPNEUMOC	2275	0.00027
RMPlate5_O_3-RM1-POL_BHV-4_95_164_rc	2216	0.00023
RMPlate3_P_10-RM1-NP_CAN_DIS_1305_137	2198	0.00017
RMPlate1_L_24-RM1-GLYCO_MOBALA_1375_1	2197	0.00019
9626433_17-Human-enterovirus-C-Picornavir	2173	0.00026
RMPlate1_K_5-RM1-NP_LASSA_425_494_rc	2170	0.00024
Human_rhinovirus_32_FJ445127_Polymerase_I	2170	0.00020
252_Escherichia_Coli_bfp_EPEC_Factor_Stool	2166	0.00021
2082_Norovirus_Pol-1_-_NC_001959_Norwalk	2123	0.00019
1255648_73-Human-parainfluenza-virus-4a-P	2073	0.00015
602_Enterovirus_Human_Echovirus_12_Capsid	2057	0.00026
988_Enterovirus_Human_Coxsackievirus_A20	2008	0.00029
75261032_nt371-Human-respiratory-syncytia	1988	0.00015
162_Escherichia_Coli_stx2_EHEC_Factor_Stoc	1958	0.00021

Tabla 9 (continuación). Análisis de muestras clínicas por producto de rangos. Los paneles C y D corresponden a la muestra clínica marcada con el número 15. Los oligos diseñados para la identificación de influenza B están marcados en rojo, los de RSV en azul y los de otras especies en negro. En C se listan los primeros cuarenta oligos ordenados de manera descendente conforme a los valores de intensidad (F532Median), así como su valor normalizado. En D están los resultados después de eliminar todos aquellos setentámeros considerados como inespecíficos de acuerdo al análisis de RP (producto de rangos).

C

Name	F532Median	Valor Normalizado
12060914_11-Human-respiratory-syncytial-virus-P	8020	0.00368
Acinetobacter_spY8_Genus_16SrRNA_HighBacteria	7950	0.00363
745_Rhinovirus_36_Capsid_AY3552171_Rhinovirus	6745	0.00308
RMPlate5_K_15-RM1-Glicoc_CvHV-1_1525_1594_r	6709	0.00304
75261032_nt296-Human-respiratory-syncytial-viru	6275	0.00278
RMPlate4_H_7-RM1-NP_MUMPS_120_189_rc	6254	0.00286
260_Sapovirus_Pol-A_-DQ366344_Sapovirus_HuE	6228	0.00279
12060914_9-Human-respiratory-syncytial-virus-Pai	5402	0.00242
781_Rhinovirus_89_Capsid_AY3552701_Rhinovirus	5004	0.00222
75261032_nt281-Human-respiratory-syncytial-viru	4957	0.00220
248_Sapovirus_Pol-A_-DQ366344_Sapovirus_HuE	4919	0.00216
976_Enterovirus_Human_Coxsackievirus_A20_Pol_	3945	0.00170
17939994_192-Bovine-respiratory-syncytial-virus-/	3616	0.00156
RMPlate5_K_14-RM1-Glicoc_CvHV-1_1505_1574_r	3402	0.00144
325251_nt0796-Influenza-B-virus-Orthomyxovirida	3198	0.00134
94_Enterovirus_Human_Echovirus_13_Capsid_AB1	3194	0.00131
cds4_Influenza_B_NA_-EU515987_Influenza_B_v	3127	0.00131
449_Human_adenovirus_D09_-AJ8544861Human	2630	0.00106
50059411_nt0451-Influenza-B-virus-(B/Hong-Kong	2608	0.00105
9627730_25-Human-rhinovirus-A-Picornaviridae-V4	2430	0.00101
12060914_7-Human-respiratory-syncytial-virus-Pai	2369	0.00095
RMPlate3_C_23-RM1-Influenza_B_virus-7_rc	2365	0.00093
926_Enterovirus_71_Pol_FJ1586011_RS1 (reverse	2353	0.00097
RMPlate3_D_3-RM1-Influenza_B_virus-11_rc	2294	0.00092
1041_Sapovirus_G3_-AF182760_Sapovirus_G3_-	2257	0.00090
50059411_nt1096-Influenza-B-virus-(B/Hong-Kong	2256	0.00092
8486152_37-Influenza-B-virus-Orthomyxoviridae-V	2241	0.00091
1510_Enterovirus_B_Pol_AY8967601_RS1 (reverse	2196	0.00090
50059411_nt0976-Influenza-B-virus-(B/Hong-Kong	1946	0.00076
407680_9-Ovine-respiratory-syncytial-virus-Param	1884	0.00070
RMPlate3_D_1-RM1-Influenza_B_virus-9_rc	1867	0.00067
Human_rhinovirus_3_DQ473485_Polymerase_Resp	1837	0.00071
566_Human_Cosavirus_Pol-D1_HCoSV-D1_-extra	1742	0.00062
458_Human_adenovirus_D09_-AJ8544861Human	1695	0.00066
707_Enterovirus_Human_Echovirus_24_Capsid_AB	1684	0.00061
450_Human_Metapneumovirus_N117_-AY296021	1658	0.00062
440_Enterovirus_Human_Coxsackievirus_B4_Capsi	1652	0.00060
254_HTCV-UC2_Human_TMEV-like_cardiovirus_iso	1643	0.00064
RMPlate3_O_5-RM1-GLYCO_AV_METAPNEUMO_80C	1624	0.00057
254_Influenza_B_HA_-FJ225980_Influenza_B_vir	1562	0.00055

D

Name	F532Median	Valor Normalizado
Acinetobacter_spY8_Genus_16SrRNA_HighBacteria	7950	0.00363
325251_nt0796-Influenza-B-virus-Orthomyxovirida	3198	0.00134
cds4_Influenza_B_NA_-EU515987_Influenza_B_v	3127	0.00131
50059411_nt0451-Influenza-B-virus-(B/Hong-Kong	2608	0.00105
RMPlate3_C_23-RM1-Influenza_B_virus-7_rc	2365	0.00093
RMPlate3_D_3-RM1-Influenza_B_virus-11_rc	2294	0.00092
50059411_nt1096-Influenza-B-virus-(B/Hong-Kong	2256	0.00092
8486152_37-Influenza-B-virus-Orthomyxoviridae-V	2241	0.00091
50059411_nt0976-Influenza-B-virus-(B/Hong-Kong	1946	0.00076
RMPlate3_D_1-RM1-Influenza_B_virus-9_rc	1867	0.00067
458_Human_adenovirus_D09_-AJ8544861Human	1695	0.00066
707_Enterovirus_Human_Echovirus_24_Capsid_AB	1684	0.00061
440_Enterovirus_Human_Coxsackievirus_B4_Capsi	1652	0.00060
RMPlate3_O_5-RM1-GLYCO_AV_METAPNEUMO_80C	1624	0.00057
254_Influenza_B_HA_-FJ225980_Influenza_B_vir	1562	0.00055
RMPlate1_F_22-RM1-HEXON_HU_ADENO_E_1370_	1559	0.00056
950_Enterovirus_71_Pol_AF3029961_RS1 (reverse	1532	0.00055
RMPlate3_O_4-RM1-GLYCO_AV_METAPNEUMO_66C	1513	0.00055
602_Enterovirus_Human_Echovirus_12_Capsid_AB	1512	0.00055
546_Human_Metapneumovirus_BR2_-AY296014_	1509	0.00050
8486154_59-Influenza-B-virus-Orthomyxoviridae-V	1477	0.00051
Acinetobacter_spY8_Genus_16SrRNA_HighBacteria	1465	0.00052
638_Human_astrovirus_Pol-A_-Z25771_Human_	1451	0.00050
61098_13-Human-rhinovirus-2-Picornaviridae-V4 (1445	0.00049
1670_HTCV-UC1_-RS1 (reversed)	1436	0.00052
252_Escherichia_Coli_bfp_EPEC_Factor_Stool-RS1	1392	0.00048
RMPlate3_O_3-RM1-GLYCO_AV_METAPNEUMO_54C	1340	0.00044
RMPlate1_L_24-RM1-GLYCO_MOBALA_1375_1444	1338	0.00044
RMPlate3_K_23-RM1-N9_NEURO_155_224-4_rc	1315	0.00042
Neisseria_Gonorrhoeae_Genus_igaProtease_NGO0	1311	0.00043
cds2_Influenza_B_NA_-EU515987_Influenza_B_v	1306	0.00045
9629176_9-Human-rhinovirus-16-Picornaviridae-V4	1302	0.00043
Actinomyces_Israelii_Genus_16SrRNA_AF479270_	1283	0.00038
RMPlate2_K_9-RM1-SS_TGV_385_454_rc	1276	0.00039
162_Escherichia_Coli_stx2_EHEC_Factor_Stool-RS	1269	0.00040
16923729_3-Human-callicivirus-SLV/Hou7-1181-Ca	1263	0.00041
4731888_25-Human-rhinovirus-sp-Picornaviridae-V	1251	0.00042
9629176_21-Human-rhinovirus-16-Picornaviridae-V	1232	0.00044
55_Enterovirus_Human_Coxsackievirus_A18_Capsi	1221	0.00039

Dentro de los 40 oligos con intensidades más altas, se encontraron oligos para rinovirus y RSV (muestra 10) o para influenza B y RSV (muestra 15). Sin embargo, al quitar los oligos que dan señal inespecífica, en ambas muestras se eliminaron todos los oligos para la detección de RSV quedando únicamente los de rinovirus o influenza tipo B en las muestras analizadas.

Es de llamar la atención que todos los oligos diseñados para la detección de RSV hayan caído en la lista de aquellos a eliminar. Sin embargo, es interesante que todos los oligos que dieron positivo para RSV, y que posteriormente se eliminaron, fueron diseñados por la UCSF. Ninguno de los oligos diseñados por nosotros para detectar RSV dio señal positiva con las muestras probadas. Los resultados negativos con nuestros oligos pudieran

explicarse por una baja cantidad de RSV presente en las muestras, lo cual queda todavía por debajo de nuestro límite de detección. Por otro lado, la aparición y posterior eliminación por inespecíficos de los oligos de RSV de UCSF pudiera deberse, como se comentó anteriormente, a la estrategia de diseño que ellos emplearon, en la cual maximizaban la detección cruzada de virus relacionados y por lo tanto son menos específicos.

A pesar de no poder detectar la presencia de RSV en ninguna de las dos muestras utilizando el protocolo modificado, sí obtuvimos una señal más intensa y confiable para la muestra positiva para influenza B, y logramos detectar con claridad el rinovirus humano, lo cual no había sido posible utilizando el protocolo inicial. Sin embargo, es necesario evaluar el desempeño del microarreglo (utilizando el procedimiento optimizado) con un mayor número de muestras clínicas para poder determinar su especificidad y sensibilidad en relación al método comercial múltiplex de la compañía Seegene.

Por todo lo anterior podemos establecer que el aumento de la proporción del nucleótido 5-(3-aminoalil)-2'-deoxiuridina 5'-trifosfato (aa-dUTP), en relación al nucleótido no modificado dTTP, es un factor importante para la mejoría de la sensibilidad del microarreglo, especialmente para las muestras clínicas, por lo que debiera adoptarse de manera rutinaria para futuros experimentos. Por otra parte, el análisis de productos de rangos permitió identificar los oligos promiscuos que, si se eliminan de los resultados al analizar muestras clínicas, puede ayudar a identificar los oligos que dan una señal significativa y específica, sin la necesidad de hacer réplicas del experimento.

14. Análisis de muestras de campo de origen porcino

En el Departamento de Producción Porcina de la Facultad de Medicina Veterinaria y Zootecnia de la UNAM se está haciendo un estudio de vigilancia epidemiológica sobre las cepas porcinas de influenza tipo A que circulan en el país. Entre noviembre de 2009 y febrero de 2011 colectaron más de 1200 biopsias de pulmón de cerdos que se encontraban en rastros en diferentes estados del país y que presentaban secreción nasal, tos y/o estornudo, lesión apical pulmonar (Dr. Iván Sánchez, comunicación personal).

Todas las muestras fueron analizadas en la FMVZ por RT-PCR en tiempo real para detectar al gen M de influenza A. Obtuvieron un total de 34 muestras positivas para

influenza, las cuales inocularon en embrión de pollo para su aislamiento, para así aumentar los títulos de los mismos, enriqueciendo las muestras y por tanto caracterizarlos mejor.

El Dr. Iván Sánchez, de la FMVZ, nos suministró 13 de los virus adaptados a crecer en embrión de pollo, los cuáles tenían títulos de hemaglutinación variables, desde menos de 1:16 hasta 1:2,048. Las 13 muestras fueron analizadas en nuestro laboratorio por RT-PCR para detectar el gen M de influenza (ver Materiales y Métodos). Seis de ellas fueron claramente positivas en este ensayo, cinco dieron un resultado no claro (amplificación de bandas débiles que no correspondían al tamaño esperado del amplicón, indicado como +/- en la tabla), y dos resultaron negativas (Tabla 10). Las 7 muestras que no dieron un resultado claro, o que fueron negativas, tenían títulos de hemaglutinación de 1:16 o menores. Las 13 muestras se analizaron con el microarreglo descrito en este trabajo. Sólo en las 6 muestras que tenían títulos de hemaglutinación altos (1:1,024 a 1:2,048), y que habían sido positivas para el gen M por RT-PCR, pudimos detectar el virus de influenza y determinar su subtipo. En este sentido es importante mencionar que las cepas estacionales de influenza que infectan más frecuentemente a cerdos pertenecen a los subtipos H1N1 y H3N2 (que a pesar de pertenecer a los mismos subtipos de las cepas estacionales humanas, tienen diferencias antigénicas considerables).

Dos de las muestras caracterizadas fueron de subtipo H3N2, una H1N? (el subtipo de N no se pudo determinar) e interesantemente, las otras tres resultaron A/H1N1/ 2009, esto es, la cepa que causó la pandemia de 2009-2010 (Tabla 10). Sólo hay un reporte previo [79] que ha descrito en nuestro país la presencia del virus pandémico en cerdos, aunque esto sí ha sido reportado en otros países.

Resulta también interesante que estas muestras habían sido previamente analizadas en la FMVZ usando un RT-PCR diseñado para determinar los subtipos H1, H3, N1 y N2 de virus de influenza A porcinos. Utilizando este ensayo sólo habían podido determinar el subtipo N de dos muestras, una que resultó N2, y la otra que dio señal para los dos subtipo de neuraminidasa probados, N1 y N2 . Con los oligos empleados no pudieron determinar el subtipo H de ninguna de estas 6 muestras (comunicación personal de Iván Sánchez, Tabla 10).

Tabla 10. Muestras colectadas y su análisis para estudio epidemiológico de cepas de influenza A circulantes en cerdos en México.

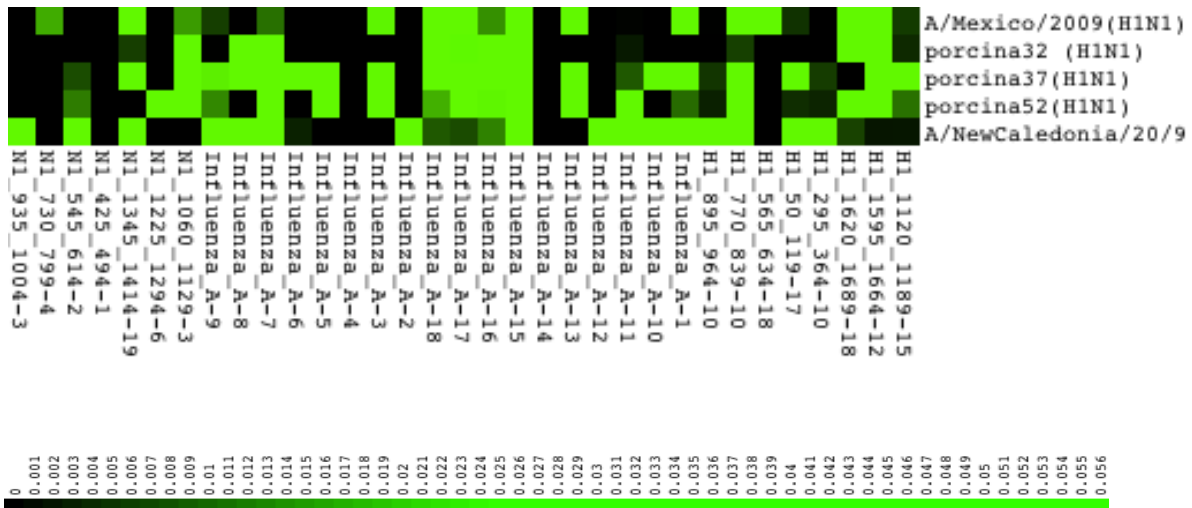
Número de muestra	Origen de colección	Titulo hemoaglutinante	RT-PCR FMVZ				RT-PCR IBt	Microarreglos
			H1	N1	H3	N2	M	
4	Tamaulipas	<1:16	-	+	-	-	+/-	-
6	Tamaulipas	<1:16	-	-	-	-	-	-
7	Tamaulipas	<1:16	-	+	-	-	-	-
19	Edo. Mex	1:16	-	+	-	-	+/-	-
23	Edo. Mex	<1:16	-	-	-	-	+/-	-
26	Jalisco	<1:16	-	-	-	-	+/-	-
29	Veracruz	1:2,048	-	-	-	-	+	Influenza A H1
31	Veracruz	1:16	-	+	-	-	+/-	-
32	Querétaro	<1:16	-	-	-	-	+	Influenza A H1N1 2009
37	Veracruz	1:1,024	-	-	-	-	+	Influenza A H1N1 2009
50	Edo. Mex	1:1,024	-	+	-	+	+	Influenza A H3N2
51	Edo. Mex	1:1,024	-	-	-	+	+	Influenza A H3N2
52	Edo. Mex	1:2,048	-	-	-	-	+	Influenza A H1N1 2009

Cabe mencionar que el diseño de los setentámeros fue previo a la pandemia y que por tanto el objetivo de su selección no fue diferenciar a la cepa causante de la pandemia de la cepa estacional que circulaba previamente, sino únicamente identificar el subtipo H1. Sin embargo, analizando los patrones de hibridación de los oligos de H1 cuando se caracterizan cepas estacionales humanas y la cepa pandémica aislada en nuestro laboratorio, observamos ciertas diferencias. Tres de los ocho oligos diseñados para este subtipo dan una señal más intensa al hibridar el material genético del virus de influenza pandémico, el cual se mantiene circulando hasta la fecha. Para la cepa estacional humana que probamos (A/NewCaledonia/20/99), los oligos que presentan una mayor intensidad son los otros cinco restantes. Estos patrones se pueden observar en la Figura 10 donde se muestran los análisis de agrupamiento de las intensidades de los oligos de H1N1, resultado de las hibridaciones de la cepa pandémica (A/Mexico/IBt-1/2009(H1N1)), la estacional (A/NewCaledonia/20/99) humana, y las muestras porcinas positivas para H1N1. Es claro que el patrón de las intensidades más altas en los oligos de H1 para las cepas porcinas es más similar al patrón observado para la cepa pandémica humana de 2009 que para la estacional. Lo cual nos hizo sospechar que en las muestras porcinas estaba presente un virus de influenza cuya hemaglutinina era similar a la de la cepa pandémica de humanos.

Sin embargo esta diferencia no se presentó en los oligos de la neuraminidasa como se muestra en la misma figura. Por lo tanto era necesario corroborar estos datos con una metodología diseñada expresamente para diferenciar entre cepas H1N1 estacionales y la cepa H1N1 pandémica.

En nuestro laboratorio se tiene montada una técnica de RT-PCR para discernir entre virus pandémicos y estacionales basado en la detección del gen de la neuraminidasa. Las muestras de cerdos se analizaron bajo esta metodología confirmando que, en efecto, las tres muestras que resultaron positivas por microarreglo para el virus H1N1 pandémico mostraban una neuraminidasa N1 pandémica por RT-PCR, lo cual indicaba que las cepas porcinas analizadas eran cepas derivadas realmente del virus pandémico A/H1N1/2009. Este análisis complementó las observaciones de los patrones de hibridación de microarreglos, así como también permitió confirmar el hecho de que las diferencias en las intensidades de la señal de los oligos hibridados pudiera ser un parámetro que sugiera fuertemente si se trata de una cepa pandémica 2009 o una estacional.

Figura 10. Análisis de agrupamiento de cepas porcinas de campo, estacional y pandémica humana. Los valores graficados son las intensidades normalizadas de los oligos mostrados en el eje horizontal, para cada una de las muestras, indicadas en el eje vertical. Los datos fueron redondeados a su cienmilésimo más cercano, analizados con el programa Cluster 3.0 y visualizados con JavaTreeView. La escala de intensidades se muestra en la parte inferior.



Finalmente, se puede decir que el microarreglo es capaz de detectar y subtipificar muestras de campo aisladas en embrión de pollo que por técnicas convencionales muy específicas no pudieron ser analizadas, mostrando su capacidad para detectar ciertos virus que puedan tener algunos cambios en su genoma que, en el caso de una RT-PCR clásica, anulan su capacidad para detectar al virus o al gen que se quiere identificar.

VII. CONCLUSIONES Y PERSPECTIVAS

En este trabajo se diseñó e implementó una nueva metodología para el diseño de un microarreglo capaz de detectar patógenos virales asociados a enfermedades respiratorias, así como de subtipificar al virus influenza tipo A basado en sus proteínas de superficie H y N. Esta técnica hace uso de varias herramientas bioinformáticas que le dan robustez a la misma. En principio utiliza filogenias para la selección de datos no redundantes, pero que a su vez reflejan la variabilidad de las secuencias reportadas. Posteriormente hace un análisis de especificidad a nivel de secuencia a través de BLAST. Finalmente, usa parámetros termodinámicos que considera desapareamientos en el alineamiento, además de ser sensible a la posición de éstos en el oligo. La combinación de estas herramientas, hasta donde nosotros sabemos, no ha sido utilizada en reportes anteriores. Es precisamente este acoplamiento de diversos programas bioinformáticos lo que le da ventajas a nuestro diseño del microarreglo sobre otras estrategias previamente reportadas. Lo anterior se debe a que la mayoría de los trabajos previos seleccionan los oligonucleótidos tomando en cuenta los resultados de una sola de las varias herramientas aquí empleadas. El programa más utilizado es BLAST, el cual es necesario más no suficiente para tener predicciones confiables de hibridación entre los oligonucleótidos de un microarreglo y una muestra dada. Existen también reportes en los que el diseño de los oligos está basado en comparaciones de patrones de hibridaciones teóricas, sin embargo los resultados obtenidos bajo este enfoque han mostrado ser relativamente inespecíficos para la detección concreta de los virus a nivel de especie.

Para la validación del microarreglo diseñado en este trabajo se utilizaron varias cepas de referencia. El análisis de los datos obtenidos a partir de las hibridaciones de estas cepas se llevó a cabo analizando el algoritmo de producto de rangos, el cual a partir de réplicas experimentales, define de manera rápida, sencilla pero estadísticamente confiable aquéllos que hayan presentado una mayor señal en todas las hibridaciones. Esta metodología no ha sido utilizada previamente para microarreglos de detección, siendo éste el primer trabajo que lo propone. Dicho análisis permitió determinar la reproducibilidad de los resultados al hibridar material genético de virus crecidos en cultivo, aún cuando la

metodología de microarreglos presenta un alto nivel de ruido inherente, así como identificar correctamente las cepas de referencia híbridadas.

Por otra parte, el método de rangos permitió también determinar aquellos oligos (dirigidos contra diversos virus, bacterias y hongos) que daban una señal intensa y reproducible en diversos experimentos, independientemente del virus de referencia hibridado, o de la muestra estudiada y que por tanto presentan una hibridación cruzada inespecífica.

Sin embargo es necesario un análisis distinto para aquellas muestras en las que no se puedan hacer réplicas del experimento de hibridación. Para ello pudiera ser útil el cálculo de un valor de probabilidad (p-value) asumiendo una distribución hipergeométrica. Con esta aproximación pudiera determinarse de manera significativa la población de oligos “sucios” y “limpios” dentro de un mismo experimento. Lo anterior basado en la agrupación del total de los oligos impresos dependiendo del virus que reconocen. De esta manera se pueden encontrar aquellos grupos mayormente representados dentro de cada experimento, es decir, aquel o aquellos virus para los cuales haya un mayor número de oligos con una señal alta, dado el número de oligos impresos en el microarreglo para detectar ese virus en particular. Se espera que el número de oligos “limpios” sea estadísticamente más significativo que el número de oligos “sucios” y que por tanto éstos últimos puedan ser descartados con una mayor confianza. Este análisis pudiera arrojar también un estudio detallado de la variabilidad de esos oligos “sucios” con lo que se definiría si en efecto la señal se debe a que el virus está presente en la muestra, o a un mal diseño de los mismos.

A pesar de ser necesaria la implementación de una nueva metodología de análisis de datos, el microarreglo pudo detectar de manera eficiente, con los parámetros hasta ahora descritos, al virus del SARS, parainfluenza tipo 2 y 3, adenovirus humano tipo 6, influenza B y sarampión, así como al virus de influenza tipo A con sus 16 variantes de hemaglutinina y 9 de neuraminidasa, incluyendo cepas estacionales y pandémicas humanas.

Se logró la identificación de virus en muestras clínicas, cuyo resultado fue corroborado por metodologías comerciales de RT-PCR, con un protocolo modificado que aumentó la sensibilidad del microarreglo. A pesar de la mejoría en la sensibilidad del ensayo lograda en este estudio, es necesario trabajar más en optimizarla. Lo anterior, con la finalidad de hacer al microarreglo más confiable para detectar virus en muestras clínicas.

Este punto es de mayor importancia para la implementación de la metodología como una herramienta básica en el diagnóstico de virus asociados a enfermedades respiratorias.

Será importante validar el funcionamiento de este microarreglo con un número mayor de muestras clínicas en comparación con métodos diagnósticos establecidos, como es el ensayo de RT-PCR de la compañía Seegen. Ya se cuenta con una colección de más de 100 muestras aisladas de niños con padecimientos respiratorios, que se analizará en colaboración con el Dr. José Ignacio Santos de la Facultad de Medicina de la UNAM.

El microarreglo se utilizó también para caracterizar exitosamente el subtipo de H y N del virus de influenza A de origen porcino de muestras colectadas de animales enfermos en diferentes estados de la República, las cuales no habían podido ser subtipificadas por métodos convencionales de RT-PCR, demostrando ser una herramienta útil cuando las metodologías convencionales fallen.

A pesar de no haber sido diseñado para ello, encontramos que el microarreglo puede distinguir entre diferentes cepas pertenecientes al mismo subtipo de H y N. Así permitió diferenciar un virus A/H1N1/2009 de origen pandémico, de una cepa H1N1 estacional con base en el patrón de hibridación de los oligos.

Este trabajo presenta también las bases para el desarrollo de nuevos microarreglos diseñados para la identificación de patógenos virales asociados a otro tipo de padecimientos, así como para determinar el origen filogenético de los ocho segmentos del genoma del virus de influenza. Este último punto proveerá un instrumento importante para la vigilancia epidemiológica de las cepas circulantes del virus de influenza tipo A en las poblaciones, de lo cual se tiene poca información en México. El estudio podrá hacerse no sólo en infecciones humanas sino también en diversos hospederos como son los cerdos, que han mostrado ser un punto clave para el surgimiento de cepas rearreglantes que pudieran ser causa de futuras epidemias.

En resumen, el microarreglo diseñado, implementado y validado en este proyecto es una valiosa herramienta de identificación, caracterización y vigilancia epidemiológica de virus asociados a enfermedades respiratorias, que no sólo incluye a los patógenos clásicos, lo que permite una exploración más amplia de los agentes etiológicos causantes de estos padecimientos. Si bien su uso como método de diagnóstico rutinario pudiera verse lejano, sí

constituye un instrumento útil para la caracterización de aquellas muestras que por los métodos clásicos de detección hayan sido negativas, así como para ampliar el conocimiento sobre las especies que constituyen el viroma o la flora viral normal del aparato respiratorio. Todo lo anterior en un mismo ensayo que tarda 2 días en arrojar resultados, y que por técnicas convencionales sería muy complicado, costoso y tardado igualar.

VIII. REFERENCIAS

1. Restrepo MA, Bedoya VI, Restrepo MI, Robledo J. Enfermedades infecciosas . Corporación para Investigaciones Biológicas 2003; sexta edición, México.
2. Organización Mundial de la Salud : Eliminar Obstáculos del Desarrollo Sustentable. www.who.int/infectious-disease-report/idr99-spanish/pages/ch2text.html
3. Dessimon YH Tai, SARS: How to Manage Future Outbreaks? *Annals Academy of Medicine* 2006; 35:368-373.
4. <http://www.ncbi.nlm.nih.gov/sites/entrez?db=Genome&itool=toolbar>
5. International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome. *Nature* 2004; 431:931-945.
6. Elnifro EM, Ashshi AM, Cooper RJ, Klapper PE. Multiplex PCR: optimization and application in diagnostic virology. *Clinical Microbiology Reviews* 2000, 13:559-570.
7. van Elden LJ, Nijhuis M, Schipper P, Schuurman R, van Loon AM. Simultaneous detection of influenza viruses A and B using real-time quantitative PCR. *Journal of Clinical Microbiology* 2001, 39:196–200.
8. Kuypers J, Wright N, Ferrenberg J, Huang ML, Cent A, Corey L, Morrow R. Comparison of real-time PCR assays with fluorescent antibody assays for diagnosis of respiratory virus infections in children. *Journal of Clinical Microbiology* 2006, 44:2382–2388.
9. Pehler-Harrington K, Khanna M, Waters CR, Hendrickson KJ. Rapid detection and identification of human adenovirus species by Adenoplex, a multiplex PCR-enzyme hybridization assay. *Journal of Clinical Microbiology* 2004, 42:4072–4076.
10. Ramírez J, Chávez L, Santillán JL, et al. Microarreglos de DNA. Mensaje Bioquímico XXVII. Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México 2003:97-120.
11. 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.
12. Wang D, Urisman A, Liu YT, et al. Viral discovery and sequence recovery using DNA microarrays. *PLoS Biology* 2003; 1:257-260.
13. Palacios G, Phenix-Lan Q, Jabado OJ, et al. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerging Infectious Diseases* 2007; 13:73-81.
14. Phenix-Lan Q, Palacios G, Jabado OJ, et al. Detection of a Respiratory viruses and subtype identification of Influenza A viruses by GreeneChipResp oligonucleotide microarray. *Journal of Clinical Microbiology* 2007; 45:2359-2364.
15. Wang D, Coscoy L, Zylberberg M, et al. Microarray-based detection and genotyping of viral pathogens. *Proceedings of the National Academy of Sciences USA* 2002; 99: 15687-15692.

16. Sengupta S, Onodera K, et al. Molecular detection and identification of influenza viruses by oligonucleotide microarray hybridization. *Journal of Clinical Microbiology* 2003; 41:4542-4550.
17. Mehlmann M, Dawson ED, et al. Robust sequence selection method used to develop the FluChip diagnostic microarray for influenza virus. *Journal of Clinical Microbiology* 2006; 44:2857-2862.
18. Chiu CY, Rouskin S, et al. Microarray detection of human parainfluenzavirus 4 infection associated with respiratory failure in an immunocompetent adult. *Clinical Infectious Diseases* 2006; 43:e71-e76.
19. Chiu CY, Urisman A, et al. Utility of DNA microarrays for detection of viruses in acute respiratory tract infections in children. *The Journal of Pediatrics* 2008; 153:76-83.
20. Xueqing H, Xiangmei L, et al. Simultaneously subtyping of all influenza A viruses using DNA microarrays. *Journal of Virological Methods* 2008; 152:117-121-
21. Ryabnin VA, Kostina EV, et al. Universal oligonucleotide microarray for subtyping of influenza A virus. *PLOS ONE* 2011; 6:e17529.
22. Heron M, Tejada-Vera B. Deaths:Leading Causes for 2005. *National Vital Statistics Reports* 2009; Vol. 58 No. 8.
23. Armstong GL, Pinner RW. Outpatient visits for infectious diseases in the United States, 1980 through 1996. *Archives of Internal Medicine* 1999; 159:2531-2536.
24. Anuarios de morbilidad. Dirección General de epidemiología. www.dgepi.salud.gob.mx/anuario/html/anuarios.html (accesado junio 2011)
25. Baron S. *Medical Microbiology*. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996, Capítulo 93: Tabla 93-1.
26. Rosete Olvera D, Archundia Sánchez F, Cabello Gutiérrez C, Manjarrez Zavala ME. Patogenia de las infecciones respiratorias por virus. *Revista Instituto Nacional de Enfermedades Respiratorias* 2002, 15-4: 239-254.
27. Talis A, McIntosh. Respiratory syncytial virus. Ch 83. In: *Manual of Clinical Microbiology*. 5th ed. Balows A et al (eds). ASM, Washington DC p 883.
28. McIntosh K and Chanock RM. Respiratory syncytial virus. Ch. 38. In: *Fields Virology*. Fields BN and Knipe DM et al (eds): Raven Press New York. 1990, 1:1045.
29. Gardner PS and McQuillin J. Application of immunofluorescent antibody technique in rapid diagnosis of respiratory syncytial virus infection. *British Medical Journal* 1968; 3:340-343.
30. Johnson KM, Chanock RM, et al. Respiratory syncytial virus iv. Correlation on virus shedding, serologic responses and illness in adult volunteers. *Journal of the American Medical Association* 1961; 176:663-667.
31. McIntosh K et al. Enzyme linked immunosorbent assay for detection of respiratory syncytial virus infection: application to clinical samples. *Journal of Clinical Microbiology*. 16:329-333.

32. Waner JL. Parainfluenza viruses. Capítulo 82. En: Manual of clinical microbiology. Fifth edition. Balows A et al (eds.) ASM, Washington DC.
33. Hierholzer J. Adenoviruses. Capítulo 86. En: Manual of clinical microbiology, Fifth edition. Balows A et al. (eds.) ASM, Washington DC.
34. DeJong J, Wigand R, et al. Candidate adenoviruses 40 and 41: fastidious adenoviruses from human infant stool. *Journal of Medical Virology* 1983; 11:215-231.
35. Gary G, Hierholzer J and Black R. Characteristics of a non cultivatable adenovirus associated with diarrhea in infants: a new subgroup of human adenoviruses. *Journal of Clinical Microbiology* 1979; 10:96-103.
36. Van den Hoogen BG, DeJong JC, et al. A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nature Medicine* 2001; 7:719-724.
37. Boivin G, De Serres G, et al. Human metapneumovirus infections in hospitalized children. *Emerging Infectious Diseases Journal* 2003; 9:634-640.
38. Viazov S, Ratjen F, et al. High prevalence of human metapneumovirus infection in young children and genetic heterogeneity of the viral isolates. *Journal of Clinical Microbiology* 2003; 41:3043-3045.
39. Mullins JA, Erdman DD, et al. Human metapneumovirus infection among children hospitalized with acute respiratory illness. *Emerging Infectious Diseases Journal* 2004; 10:700-705.
40. Peiris JS, Lai ST, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 2003; 361:1319-1325.
41. Ksiazek TG, Erdman D, et al. A novel coronavirus associated with severe acute respiratory syndrome. *New England Journal of Medicine* 2003; 348:1953-1966.
42. Drosten C, Gunther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New England Journal of Medicine* 2003; 348:1953-1966.
43. World Health Organization. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. www.who.int/csr/sars/country/table2004_04_21/en/. (accesado junio de 2011)
44. Arias CF, Escalera-Zamudio M, et al. Molecular anatomy of 2009 influenza virus A (H1N1). *Archives of Medical Research* 2009; 40:643-654.
45. Drake JC. Rates of spontaneous mutation among RNA virus. *Proceedings of the National Academy of Sciences USA* 1993; 90:4171-4175.
46. Franco PC, Téllez I, Santos-Preciado JI. Pandemia de influenza: posible impacto de la influenza aviaria, *Salud Pública de México* 2005; 47:107-108.
47. Wright P, Neumann G, Kawaoka Y. Orthomyxoviruses. In: D Knipe; P Howley, editors, translator and editor *Fields Virology*. Fifth Edition edn. Vol. II. Philadelphia: Lippincott Williams & Wilkins; 2006; p. 1691-740.

48. Strauss J, Strauss E. Minus-strand RNA viruses. In, translator and editor *Viruses and human disease*. Second edn. Vol. 4. Oxford, UK: Elsevier; 2008; 137-75.
49. Smith GJ, Bahl J, Vijaykrishna D et al. Dating the emergence of pandemic influenza viruses. *Proceedings of the National Academy of Sciences USA* 2009; 106:11709-12.
50. Subbarao K, Klimov A, Katz J, et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science* 1998; 279:393-396.
51. Ungchaisak K, Auewarakul P, Dowell SF, et al. Probable person-to-person transmission of avian influenza A (H5N1). *New England Journal of Medicine* 2005; 352:333-340.
52. European Center for Disease Prevention and Control. Origin of the 2009 pandemic influenza A(H1N1) virus. www.ecdc.europa.eu/EN/HEALTHTOPICS/H1N1/BASIC_FACTS/Pages/origin_of_virus.aspx (accesado junio de 2011).
53. Garten RJ, Davis CT, Russell CA et al. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 2009; 325:197-201.
54. Trifonov V, Khiabani H, Rabadan R. Geographic dependence, surveillance, and origins of the 2009 influenza A (H1N1) virus. *New England Journal of Medicine* 2009; 361:115-9.
55. Fraser C, Donnelly CA, Cauchemez S et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 2009; 324:1557-61.
56. Brown IH. The epidemiology and evolution of influenza viruses in pigs. *Veterinary Microbiology* 2000; 74:29-46.
57. Butler D. How severe will the flu outbreak be? 2009. *Nature* 2009; 14:14-15.
58. World Health Organization. Pandemic (H1N1) 2009 update 112. www.who.int/int/csr/don/2010_08_06/en/index.html (accesado junio 2011).
59. World Health Organization. www.who.int/csr/don/2010_09_10/en/index.html
60. Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ. Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences* 1998; 10:403-405.
61. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in bioinformatics* 2004 5:150-163.
62. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 1987; 4:406-425.
63. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 1993; 10:512-526.
64. Altschul SF, Warren G, Webb M, Myers WE, Lipman JD. Basic local alignment search tool. *Journal of Molecular Biology* 1990, 215:403-410.

65. SantaLucia Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Science* 1998; 95:1460-1465.
66. Soto del Río, María de los Dolores. (2008) *Implementación y validación de una metodología para la detección de patógenos asociados a enfermedades respiratorias por medio de un microarreglo: subtipificación del virus influenza tipo a* (Tesis de Licenciatura – Universidad Nacional Autónoma de México)
67. Ginocchio C, Zhang F, et al. Evaluation of multiple test methods for the detection of the novel 2009 influenza A (H1N1) during the New York City outbreak. *Journal of Clinical Virology* 2009; 45:191-195.
68. Breitling R, Armengaud P, et al. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS letter* 2004; 573:83-92.
69. Eisen MB, Spellman PT, Brown PO, et al. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Science* 1998; 95:14863–14868.
70. Van Regenmortel MHV, Fauquete CM, et al. *Virus Taxonomy Seventh report of the International Committee on Taxonomy of Viruses*. Academic Press 2000. USA.
71. Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of influenza evolution. *Nature* 2003; 422:428-433.
72. Smith DJ, Lapedes AS, de Jong JC, et al. Mapping the antigenic and genetic evolution of influenza virus. *Science* 2004; 305:371-376.
73. Fernández, Selene. (2007) *Diseño, implementación y validación de un microarreglo piloto para la identificación de patógenos asociados a enfermedades respiratorias y gastrointestinales* (Tesis de Licenciatura – Universidad Nacional Autónoma de México)
74. Sánchez Quinto, Federico Andrés (2008) *Diseño, implementación y validación de un microarreglo piloto para identificar patógenos respiratorios y gastrointestinales* (Tesis de Licenciatura – Universidad Nacional Autónoma de México)
75. www.pangloss.com/seidel/Protocols/probecheck.html
76. www.riedlab.nci.nih.gov/publications/cDNA%20indirect%20%20labelling.pdf
77. www.biotechniques.com/multimedia/archive/00055/BTN_A_000113228_O_55829a.pdf
78. www.cme.msu.edu/tiedjelab/protocols-microarray2.doc
79. Escalera-Zamudio M, Cobián-Güemes G, et al. Identification of an Influenza Virus Strain that Represents a Sister Lineage to the A/H1N1/2009 Pandemic Clade". *PLOS One*, manuscrito enviado.

Anexo I. Familias, géneros y especies virales incluidos en el microarreglo.

---- ADENOVIRIDAE ----

*** ATADENOVIRUS ***

- BOVINE_ADENOVIRUS_D
- BOVINE_ADENOVIRUS_E
- BOVINE_ADENOVIRUS_F
- DUCK_ADENOVIRUS_A
- OVINE_ADENOVIRUS_D
- POSSUM_ADENOVIRUS

*** AVIANADENOVIRUS ***

- FOWL_ADENOVIRUS_A
- FOWL_ADENOVIRUS_B
- FOWL_ADENOVIRUS_C
- FOWL_ADENOVIRUS_D
- FOWL_ADENOVIRUS_E

*** MASTADENOVIRUS ***

- BOVINE_ADENOVIRUS_B
- BOVINE_ADENOVIRUS_C
- CANINE_ADENOVIRUS
- EQUINE_ADENOVIRUS_A
- HUMAN_ADENOVIRUS_A
- HUMAN_ADENOVIRUS_B
- HUMAN_ADENOVIRUS_C
- HUMAN_ADENOVIRUS_E
- HUMAN_ADENOVIRUS_F
- HUMAN_ADENOVIRUS_G
- OVINE_ADENOVIRUS_A
- OVINE_ADENOVIRUS_B
- OVINE_ADENOVIRUS_C
- PORCINE_ADENOVIRUS_A
- PORCINE_ADENOVIRUS_B
- PORCINE_ADENOVIRUS_C
- SIMIAN_ADENOVIRUS

---- ARENAVIRIDAE ----

*** ARENAVIRUS ***

- GUANARITO_VIRUS
- IPPY_VIRUS
- JUNIN_VIRUS
- LASSA_VIRUS
- LYMPHOCYTIC_CHORIOMENINGITIS_VIRUS
- MACHUPO_VIRUS
- MOBALA_VIRUS

---- BUNYAVIRIDAE ----

*** HANTAVIRUS ***

- ANDES_VIRUS
- BAYOU_VIRUS
- BLACK_CREEK_CANAL_VIRUS
- JUQUITIBA
- LAGUNA_NEGRA_VIRUS
- LECHIGUANAS_VIRUS
- MONONGAHELA_VIRUS
- NEW_YORK_VIRUS
- ORAN_VIRUS
- PUUMALA_VIRUS
- SIN_NOMBRE_VIRUS

```

*** NAIROVIRUS ***
  - CRIMEAN_CONGO_HEMORRHAGIC_FEVER_VIRUS
  - DUGBE_VIRUS
*** PHLEBOVIRUS ***
  - RIFT_VALLEY_FEVER_VIRUS
---- CALICIVIRIDAE ----
*** LAGOVIRUS ***
  - EUROPEAN_BROWN_HARE
  - RABBIT_HEMORRHAGIC_DISEASE_VIRUS
*** NOROVIRUS ***
  - NORWALKVIRUS
*** SAPOVIRUS ***
  - SAPPORO_VIRUS
*** VESIVIRUS ***
  - FELINE_CALICIVIRUS
  - RABBIT_VESIVIRUS
---- CORONAVIRIDAE ----
*** CORONAVIRUS ***
  - CANINE_CORONAVIRUS
  - EQUINE_CORONAVIRUS
  - FELINE_CORONAVIRUS
  - HUMAN_CORONAVIRUS_229E
  - HUMAN_CORONAVIRUS_HKU1
  - HUMAN_CORONAVIRUS_NL63
  - HUMAN_CORONAVIRUS_OC43
  - HUMAN_ENTERIC_CORONAVIRUS
  - INFECTIOUS_BRONCHITIS_VIRUS
  - MURINE_HEPATITIS_VIRUS
  - PHEASANT_CORONAVIRUS
  - PORCINE_EPIDEMIC_DIARRHEA_VIRUS
  - SARS
  - TRANSMISSIBLE_GASTROENTERITIS_VIRUS
  - TURKEY_CORONAVIRUS
*** TOROVIRUS ***
  - BOVINE_TOROVIRUS
  - EQUINE_TOROVIRUS
  - HUMAN_TOROVIRUS
  - PORCINE_TOROVIRUS
---- FLAVIVIRIDAE ----
*** FLAVIVIRUS ***
  - DENGUE_VIRUS_TYPE_1
  - DENGUE_VIRUS_TYPE_2
  - DENGUE_VIRUS_TYPE_3
  - DENGUE_VIRUS_TYPE_4
  - ILHEUS_VIRUS
  - JAPANESE_ENCEPHALITIS_VIRUS
  - LOUPING-ILL_VIRUS
  - NTAYA_VIRUS
  - POWASSAN_VIRUS
  - ST_LOUIS_ENCEPHALITIS_VIRUS
  - WESSELSBRON_VIRUS
  - WEST_NILE_VIRUS
  - YELLOW_FEVER_VIRUS
  - ZIKA_VIRUS
*** HEPACIVIRUS ***
  - HEPATITIS_C_VIRUS
*** PESTIVIRUS ***
  - BORDER_DISEASE_VIRUS
  - BOVINE_VIRAL_DIARRHEA_VIRUS_1

```


- BOVINE_VIRAL_DIARRHEA_VIRUS_2
- CLASSICAL_SWINE_FEVER_VIRUS
- HERPESVIRIDAE ----
- *** ALPHAHERPESVIRINAE ***
- **ILTOVIRUS**
- Gallid herpesvirus 1
- **SIMPLEX_VIRUS**
- Cercopithecine herpesvirus 1
- Cercopithecine herpesvirus 2
- Cercopithecine herpesvirus 16
- Human herpesvirus 1
- Human herpesvirus 2
- Macropodid herpesvirus 1
- Macropodid herpesvirus 2
- Saimiriine herpesvirus 1
- **VARICELLOVIRUS**
- BOVINE HERPESVIRUS 1
- BUBALINE HERPESVIRUS 1
- CANID HERPESVIRUS 1
- CAPRINE HERPESVIRUS 1
- CERCOPITHECINE HERPESVIRUS 9
- CERVID HERPESVIRUS 1
- CERVID HERPESVIRUS 2
- EQUID HERPESVIRUS 1
- EQUID HERPESVIRUS 4
- FELID HERPESVIRUS 1
- HUMAN HERPESVIRUS 3
- SUID HERPESVIRUS 1
- PHOCID HERPESVIRUS 1
- *** BETAHERPESVIRINAE ***
- ** CYTOMEGALOVIRUS **
- HUMAN HERPESVIRUS 5
- ** ROSEOLOVIRUS **
- HUMAN HERPESVIRUS 6
- HUMAN HERPESVIRUS 7
- *** GAMMAHERPESVIRINAE ***
- ** LYMPHOCRIPTOVIRUS **
- HUMAN HERPESVIRUS 4
- ** RHADINOVIRUS **
- ALCELAPHINE HERPESVIRUS 1
- ALCELAPHINE HERPESVIRUS 2
- BOVINE HERPESVIRUS 4
- EQUID HERPESVIRUS 2
- EQUID HERPESVIRUS 5
- MURID HERPESVIRUS 4
- OVINE HERPESVIRUS 2
- SAIMIRIINE HERPESVIRUS 2
- ORTHOMYXOVIRIDAE ----
- *** INFLUENZA_VIRUS_A ***
- INFLUENZA_A & SUBTYPES
- *** INFLUENZA_VIRUS_B ***
- INFLUENZA_B
- *** INFLUENZA_VIRUS_C ***
- INFLUENZA_C
- PARAMIXOVIRIDAE ----
- *** AVULAVIRUS ***
- AVIAN_PARAMYXOVIRUS_2
- AVIAN_PARAMYXOVIRUS_3
- AVIAN_PARAMYXOVIRUS_4

- AVIAN_PARAMYXOVIRUS_6
- GOOSE_PARAMYXOVIRUS_SF02
- NEWCASTLE_VIRUS
- *** HENIPAVIRUS ***
- HENDRA_VIRUS
- NIPAH_VIRUS
- *** METAPNEUMOVIRUS ***
- AVIAN_METAPNEUMOVIRUS
- HUMAN_METAPNEUMOVIRUS
- *** MORBILLIVIRUS ***
- CANINE_DISTEMPER_VIRUS
- CETACEAN_MORBILLIVIRUS_DOLPHIN
- MEASLES_VIRUS
- PESTE-DES-PESTIS-RUMINANTS_VIRUS
- PHOCINE_DISTEMPER
- RINDERPEST_VIRUS
- *** PNEUMOVIRUS ***
- BOVINE_RESPIRATORY_SYNCYTIAL_VIRUS
- HUMAN_RESPIRATORY_SYNCYTIAL_VIRUS
- MURINE_PNEUMONIA_VIRUS
- *** RESPIROVIRUS ***
- BOVINE_PARAINFLUENZA_3
- HUMAN_PARAINFLUENZA_1
- HUMAN_PARAINFLUENZA_3
- SENDAI_VIRUS
- *** RUBULAVIRUS ***
- HUMAN_PARAINFLUENZA_2
- HUMAN_PARAINFLUENZA_4A
- HUMAN_PARAINFLUENZA_4B
- MAPUERA_VIRUS
- MUMPS
- PORCINE_RUBULAVIRUS
- SIMIAN_VIRUS_41
- SIMIAN_VIRUS_5
- *** UNASSIGNED_PARAMYXOVIRUS ***
- FER-D-LANCE_VIRUS
- MENANGLE_VIRUS
- TIOMAN_VIRUS
- PARVOVIRIDAE ----
- *** PARVOVIRINAE ***
- ** BOCAVIRUS **
- BOVINE PARVOVIRUS
- CANINE MINUTE VIRUS
- HUMAN BOCAVIRUS
- ** PARVOVIRUS **
- CHICKEN PARVOVIRUS
- FELINE PANLEUKOPENIA VIRUS
- MINUTE VIRUS OF MICE
- PORCINE PARVOVIRUS
- ** ERYTHOVIRUS **
- HUMAN PARVOVIRUS B19
- PIG-TAILED MACAQUE PARVOVIRUS
- RHESUS MACAQUE PARVOVIRUS
- SIMIAN PARVOVIRUS
- PICORNAVIRIDAE ----
- *** APHTOVIRUS ***
- EQUINE_RHINITIS_VIRUS

*** ENTEROVIRUS ***
 - BOVINE_ENTEROVIRUS
 - HUMAN_ENTEROVIRUS_A
 - HUMAN_ENTEROVIRUS_B
 - HUMAN_ENTEROVIRUS_C
 - HUMAN_ENTEROVIRUS_D
 - PORCINE_ENTEROVIRUS_A
 - PORCINE_ENTEROVIRUS_B
 - SIMIAN_ENTEROVIRUS_A
 *** ERBOVIRUS ***
 - EQUINE_RHINITIS_B_VIRUS
 *** PARECHOVIRUS ***
 - HUMAN_PARECHOVIRUS
 - LJUNGAN_VIRUS
 *** RHINOVIRUS ***
 - HUMAN_RHINOVIRUS_A
 - HUMAN_RHINOVIRUS_B
 - HUMAN_RHINOVIRUS_C
 ---- POXVIRIDAE ----
 *** LEPORIPOXVIRUS ***
 - MYXOMA_VIRUS
 *** ORTHOPOXVIRUS ***
 - COWPOX_VIRUS
 - MONKEYPOX_VIRUS
 - VACCINIA_VIRUS
 - VARIOLA_VIRUS
 ---- REOVIRIDAE ----
 *** ORBIVIRUS ***
 - AFRICAN_HORSE_SICKNESS_VIRUS
 - EPIZOOTIC_HEMORRHAGIC_DISEASE_VIRUS
 *** ORTHOREOVIRUS ***
 - AVIAN_ORTHOREOVIRUS
 - BABOON_ORTHOREOVIRUS
 - MAMMALIAN_ORTHOREOVIRUS
 - MELAKA_VIRUS
 - NELSON_BAY_ORTHOREOVIRUS
 - REPTILIAN_ORTHOREOVIRUS
 ---- RHABDOVIRIDAE ----
 *** EPHEMEROVIRUS ***
 - BOVINE_EPHEMERAL_FEVER_VIRUS
 ---- UNASSIGNEDFAMILY ----
 *** ANNELOVIRUS ***
 - SEN_VIRUS
 - TORQUE_TENO_LIKE_MINI_VIRUS
 - TORQUE_TENO_VIRUS

Anexo II. Resultados del análisis de producto de rangos (RP) de cepas de referencia

Se muestran los resultados de productos de rangos de tres experimentos de hibridación independientes con las cepa de indicada. Sólo se muestran aquellos oligos con un FDR=0. Los oligos cuyo nombre comienzan con "RMPlate" corresponden a los diseñados en este trabajo, el resto es diseño de UCSF. RP es el valor de producto de rangos, FDR es la tasa de falsos positivos.

virus parainfluenza humano tipo 3			
OLIGO	RP	E-VALUE	FDR
60972_14-Human-parainfluenza-virus-3-Paramyxovi	147.300649	2.15E-11	0
RMPlate4_E_4-RM1-NP_HU_PARAINFL3_870_939_rc	143.063901	7.18E-11	0
RMPlate4_E_3-RM1-NP_HU_PARAINFL3_600_669_rc	138.280339	1.15E-10	0
RMPlate4_E_5-RM1-NP_HU_PARAINFL3_890_959_rc	127.532603	6.89E-10	0
RMPlate4_E_7-RM1-POL_HU_PARAINFL3_5150_5219	84.1532555	2.15E-09	0
RMPlate4_E_8-RM1-POL_HU_PARAINFL3_6490_6559	96.2130522	2.33E-09	0
60972_66-Human-parainfluenza-virus-3-Paramyxovi	90.1835369	3.68E-09	0
RMPlate4_E_2-RM1-HN_HU_PARAINFL3_990_1059	99.0610505	3.73E-09	0
12060914_11-Human-respiratory-syncytial-virus-Par	72.504962	5.92E-09	0
10937870_426-Human-parainfluenza-virus-3-Param	90.7965736	6.68E-09	0
10937870_394-Human-parainfluenza-virus-3-Param	83.181597	7.75E-09	0
12060914_9-Human-respiratory-syncytial-virus-Para	67.2364513	8.44E-09	0
75261032_nt296-Human-respiratory-syncytial-virus-	50.3644302	1.40E-08	0
RMPlate4_E_1-RM1-HN_HU_PARAINFL3_715_784_rc	65.4898062	1.56E-08	0
17939994_192-Bovine-respiratory-syncytial-virus-AT	35.4663828	3.00E-08	0
75261032_nt281-Human-respiratory-syncytial-virus-	26.5028478	4.50E-08	0
254_HTCV-UC2_Human_TMEV-like_cardiovirus_isola	24.967289	4.70E-08	0
RMPlate4_H_7-RM1-NP_MUMPS_120_189_rc	25.9506532	5.05E-08	0
450_Human_Metapneumovirus_NL17_-AY296021_1	27.4210257	5.49E-08	0
RMPlate4_E_6-RM1-POL_HU_PARAINFL3_2795_2864	38.2260401	6.82E-08	0
781_Rhinovirus_89_Capsid_AY3552701_Rhinovirus	18.4626096	9.62E-08	0
12060914_7-Human-respiratory-syncytial-virus-Para	18.6024599	1.19E-07	0
260_Sapovirus_Pol-A_-DQ366344_Sapovirus_HuEh	16.95868	1.29E-07	0
10937870_412-Human-parainfluenza-virus-3-Param	47.0953727	1.39E-07	0
RMPlate3_O_4-RM1-GLYCO_AV_METAPNEUMO_660	17.5840641	1.53E-07	0
RMPlate3_P_10-RM1-NP_CAN_DIS_1305_1374_rc	15.3381903	1.61E-07	0
976_Enterovirus_Human_Coxsackievirus_A20_Pol_D	14.1612414	1.92E-07	0
RMPlate5_K_15-RM1-GlicoC_CvHV-1_1525_1594_rc	14.168815	2.23E-07	0
248_Sapovirus_Pol-A_-DQ366344_Sapovirus_HuEh	13.3625619	2.23E-07	0
RMPlate3_O_5-RM1-GLYCO_AV_METAPNEUMO_800	12.7406756	2.35E-07	0
RMPlate4_E_9-RM1-HN_SENDAI_1525_1594_rc	15.7939982	2.37E-07	0
745_Rhinovirus_36_Capsid_AY3552171_Rhinovirus	13.8291182	2.50E-07	0

virus de sarampión			
OLIGO	RP	E-VALUE	FDR
RMPlate4_A_5-RM1-POL_MEASLES_4125_419	118.056097	2.77E-14	0
12060914_11-Human-respiratory-syncytial-vir	65.9445832	1.80E-13	0
RMPlate4_A_3-RM1-POL_MEASLES_1610_167	86.9177308	1.20E-12	0
12060914_9-Human-respiratory-syncytial-viru	42.6371591	1.74E-12	0
RMPlate4_A_4-RM1-POL_MEASLES_1825_189	95.5603632	1.77E-12	0
RMPlate4_A_2-RM1-POL_MEASLES_105_174	92.7699543	3.20E-12	0
75261032_nt296-Human-respiratory-syncytial	40.6653523	4.57E-12	0
RMPlate4_A_7-RM1-POL_MEASLES_4360_442	64.1016445	5.70E-12	0
RMPlate4_A_1-RM1-HN_MEASLES_750_819_r	78.7870602	7.59E-12	0
RMPlate3_P_21-RM1-HN_MEASLES_1125_119	63.0949382	1.72E-11	0
17939994_192-Bovine-respiratory-syncytial-v	27.5478473	2.29E-11	0
RMPlate4_A_6-RM1-POL_MEASLES_435_504	52.9973761	4.87E-11	0
75261032_nt281-Human-respiratory-syncytial	20.1049308	5.72E-11	0
RMPlate3_P_24-RM1-HN_MEASLES_630_699	30.0578012	1.05E-10	0
12060914_7-Human-respiratory-syncytial-viru	20.119385	1.06E-10	0
RMPlate4_H_7-RM1-NP_MUMPS_120_189_rc	18.5342324	1.13E-10	0
781_Rhinovirus_89_Capsid_AY3552701_Rhinc	17.3709124	1.41E-10	0
RMPlate3_P_22-RM1-HN_MEASLES_1290_135	38.7439638	1.89E-10	0
450_Human_Metapneumovirus_NL17_-AY296	15.4307525	2.30E-10	0
254_HTCV-UC2_Human_TMEV-like_cardiovirus	15.6598386	2.63E-10	0
RMPlate3_P_20-RM1-HN_MEASLES_1000_106	23.452572	3.31E-10	0
RMPlate3_O_4-RM1-GLYCO_AV_METAPNEUMO	19.3532092	3.96E-10	0
260_Sapovirus_Pol-A_-DQ366344_Sapovirus	11.8505461	4.87E-10	0
1041_Sapovirus_G3_-AF182760_Sapovirus	11.7688754	5.16E-10	0
248_Sapovirus_Pol-A_-DQ366344_Sapovirus	10.6745173	5.99E-10	0
RMPlate5_K_15-RM1-GlicoC_CvHV-1_1525_15	11.0788078	6.74E-10	0
745_Rhinovirus_36_Capsid_AY3552171_Rhinc	11.5899087	6.82E-10	0
55_Enterovirus_Human_Coxsackievirus_A18	11.1238325	6.87E-10	0
976_Enterovirus_Human_Coxsackievirus_A20	11.3722613	6.88E-10	0
94_Enterovirus_Human_Echovirus_13_Capsid	10.6868807	7.86E-10	0
1670_HTCV-UC1_-RS1 (reversed)	10.8634884	1.07E-09	0
RMPlate3_O_5-RM1-GLYCO_AV_METAPNEUMO	12.127516	1.08E-09	0
RMPlate3_P_23-RM1-HN_MEASLES_1530_159	17.5735113	1.12E-09	0
407680_9-Ovine-respiratory-syncytial-virus-P	9.31652627	2.28E-09	0
926_Enterovirus_71_Pol_FJ1586011_RS1 (rev	7.42999063	2.70E-09	0
RMPlate4_A_9-RM1-HN_PEST_RUMINANT_100	10.9123256	2.82E-09	0
17939994_201-Bovine-respiratory-syncytial-v	10.1485373	2.91E-09	0
RMPlate1_F_16-RM1-POL_HU_ADENO_C_100	9.68941627	2.93E-09	0
1510_Enterovirus_B_Pol_AY8967601_RS1 (rev	6.92667543	4.30E-09	0
Human_rhinovirus_3_DQ473485_Polymerase	6.48323714	4.43E-09	0
RMPlate1_K_5-RM1-NP_LASSA_425_494_rc	6.17042909	5.73E-09	0
RMPlate5_K_14-RM1-GlicoC_CvHV-1_1505_15	6.1139685	5.76E-09	0
449_Human_adenovirus_D09_-AJ8544861Hu	6.73437615	6.16E-09	0
RMPlate4_F_21-RM1-NP_HU_PARAINFL4B_170	6.89372543	6.28E-09	0
64_Enterovirus_Human_Coxsackievirus_A18	7.70234049	6.70E-09	0
RMPlate4_C_17-RM1-GLYCO_MU_PNMN_310	6.39522042	7.20E-09	0
566_Human_Cosavirus_Pol-D1_HCoSV-D1_-	5.75177922	7.40E-09	0
546_Human_Metapneumovirus_BR2_-AY296	9.04985775	7.56E-09	0
210_Human_Metapneumovirus_BR2_-AY296	6.090716	8.08E-09	0
9627730_25-Human-rhinovirus-A-Picornavirid	5.96108778	8.27E-09	0
RMPlate1_F_22-RM1-HEXON_HU_ADENO_E_1	5.22664448	9.79E-09	0
Actinomyces_Israelii_Genus_16SrRNA_AF4792	5.45159168	1.14E-08	0
16923729_3-Human-calicivirus-SLV/Hou7-118	4.82155595	1.33E-08	0
440_Enterovirus_Human_Coxsackievirus_B4	4.69780987	1.86E-08	0
719_Enterovirus_Human_Echovirus_21_Capsi	4.85382535	1.87E-08	0
707_Enterovirus_Human_Echovirus_24_Capsi	4.58924238	1.92E-08	0
9629096_739-Canine-adenovirus-Adenovirida	5.1729765	1.99E-08	0
458_Human_adenovirus_D09_-AJ8544861Hu	4.4386486	2.27E-08	0
21407997_816-Human-adenovirus-B-Adenovi	5.34842494	2.34E-08	0
RMPlate3_O_3-RM1-GLYCO_AV_METAPNEUMO	6.10222885	2.41E-08	0
RMPlate1_A_19-RM1-HumanGADPH_505_rc	5.39989373	2.46E-08	0