



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

BIOMEDICINA

**DESARROLLO DE ESTRATEGIAS MOLECULARES EMPLEANDO
SECUENCIACIÓN DE NUEVA GENERACIÓN Y SU APLICACIÓN EN EL
DIAGNÓSTICO DE NUEVOS PATÓGENOS RELACIONADOS CON
ENFERMEDADES RESPIRATORIAS AGUDAS**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:

VICTOR ALBERTO HERNÁNDEZ HERNÁNDEZ

TUTOR PRINCIPAL DE TESIS: DR. JOEL ARMANDO VÁZQUEZ PÉREZ

FACULTAD DE MEDICINA, INER.

COMITÉ TUTOR: DRA. ROSA ELENA SARMIENTO SILVA

FACULTAD DE MEDICINA, UNAM.

DR. EDGAR ENRIQUE SEVILLA REYES

FACULTAD DE MEDICINA, INER.

MÉXICO, D.F. NOVIEMBRE, 2015



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.



Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted que el Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, en su sesión ordinaria del día 31 de agosto de 2015, aprobó el jurado para la presentación de su examen para obtener el grado de **MAESTRO EN CIENCIAS BIOLÓGICAS** del alumno **HERNÁNDEZ HERNÁNDEZ VÍCTOR ALBERTO** con número de cuenta 302244243, con la tesis titulada **"DESARROLLO DE ESTRATEGIAS MOLECULARES EMPLEANDO SECUENCIACIÓN DE NUEVA GENERACIÓN Y SU APLICACIÓN EN EL DIAGNÓSTICO DE NUEVOS PATÓGENOS RELACIONADOS CON ENFERMEDADES RESPIRATORIAS AGUDAS"**, realizada bajo la dirección del **DR. JOEL ARMANDO VÁZQUEZ PÉREZ**:

Presidente: DR. ARTURO CARLOS II BECERRA BRACHO
Vocal: DR. SANTIAGO ÁVILA RÍOS
Secretario: DRA. ROSA ELENA SARMIENTO SILVA
Suplente: DR. EZEQUIEL MOISÉS FUENTES PANANÁ
Suplente: DR. EDGAR ENRIQUE SEVILLA REYES

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Cd. Universitaria, D.F., a 13 de octubre de 2015

M. del Coro Arizmendi

DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA



AGRADECIMIENTOS

Al programa de Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México,
por todo el aprendizaje invaluable.

Al Consejo Nacional de Ciencia y Tecnología por su apoyo con la beca: CVU 587984 y número de registro 307257. A la Secretaría de Ciencia y Tecnología en Innovación del D.F. (SECITI DF) proyecto PICSA 12-164 y a la Fundación Miguel Alemán por el apoyo financiero para llevar a cabo este trabajo de tesis

A los miembros del comité tutor Dr. Joel Armando Vázquez Pérez, Dra. Rosa Elena Sarmiento Silva y al Dr. Edgar Enrique Sevilla Reyes, por su apoyo académico, sus consejos, tiempo y esfuerzo para enriquecer mi trabajo durante mis estudios de posgrado.

**Primero que nada le agradezco a mi mamá, mi tía y mi abuela por todo lo que soy, saben que sin ustedes nada de esto hubiese sido posible,
¡Gracias!**

A mi esposa y a mi hija por darme todo el apoyo, sus cuidados y por generar mi alegría. A pesar de que no fue un camino fácil ustedes me dieron las fuerzas necesarias para seguir y superarme día a día, las amo.

A toda mi familia por su apoyo, amistad y por estar siempre ahí cuando más se les necesita Isaac, Reina, Arturo, Antonia y Luis.

A mi segunda familia, les agradezco mucho su confianza y su apoyo incondicional Sra Silvia, Sr Fernando, Ara y a las pequeñas Ale y Fer.

A mis amigos Alfonso, Pablo, Roque, Román, Antonio, Daniel y Rogelio, por todo su apoyo y por impulsar todas mis locuras y estar ahí para escucharme.

A mis amigas Carmen, Elvira, Loana, Janet, Laura, Liliana, Cecilia, Amanda y Mariana porque me han ayudado, enseñado y apoyado en tantas cosas que no hubiese sido posible llegar hasta aquí sin su apoyo.

A mis colegas del laboratorio de Investigación en Virología y Micología del INER. M. en C. Fidencio, Q.F.B. Pedro, M. en C. Fernando, Dr Carlos y al Dr Joaquín, por aportar su granito de arena en este trabajo, ya que sin la disponibilidad de sus laboratorios y sin toda la retroalimentación en mi trabajo no hubiese sido posible terminarlo.

A todos los amigos de Liver Aguilar, Ponce, Bety, Sandra, Pablo, Isabel, Claudia, Alicia y Angelina.

Agradezco el apoyo de la Dra Ongay de Fisiología Celular en la UNAM, por dejarnos utilizar su secuenciador para realizar parte del procesamiento.

Finalmente agradezco a todos los pacientes que amablemente me proporcionaron su consentimiento para llevar a cabo este estudio, sin sus muestras nada de esto hubiese sido posible.

Les dedico este trabajo a mi hija

a mi esposa,

y mi mamá.

Saben que son lo más importante en mi vida

y espero que este esfuerzo recompense un poco de todo

lo que ustedes me dan.

ÍNDICE	
ÍNDICE GENERAL.....	I
ÍNDICE DE FIGURAS.....	IV
ÍNDICE DE TABLAS.....	V
ÍNDICE DE GRÁFICAS.....	VI
ÍNDICE DE DIAGRAMAS	VII
ÍNDICE DE MATERIAL SUPLEMENTARIO.....	VIII
ABREVIATURAS.....	IX
RESUMEN.....	1
ASBTRACT.....	2
I. INTRODUCCIÓN.....	3
1.1.-Enfermedad respiratoria aguda.....	3
1.2.-Agentes etiológicos.....	3
1.3. Patogenicidad.....	6
1.3.1 Factores del hospedero.....	6
1.3.2 Factores de los agentes etiológicos.....	7
1.4.-Técnicas de diagnóstico.....	8
1.5.-No moleculares.....	8
1.6.-Moleculares.....	9
1.7.- Técnicas moleculares comerciales.....	10
1.8.- Secuenciación masiva.....	11
1.9.- Secuenciación masiva basada en síntesis (Illumina).....	13
1.9.1.- Fundamento de la metodología de Illumina MiSeq	13
II. ANTECEDENTES.....	16
III. JUSTIFICACIÓN.....	24
IV. HIPÓTESIS.....	25
V. OBJETIVOS.....	26
VI. METODOLOGÍA.....	27

6.1.- Criterios de inclusión.....	29
6.2.- Toma de muestra.....	29
6.3.- Extracción y purificación de ácidos nucleicos.....	30
6.3.1.- Alícuota para bacterias.....	30
6.3.1.1.-Amplicones de la región hipervariable V1-V3.....	30
6.3.1.2.- Cuantificación de ácidos nucleicos para 16S.....	31
6.3.2.- Alícuota para detección de virus.....	31
6.3.2.1.- Detección de virus y clasificación de muestras	32
6.3.3.- Alícuota para estrategias de secuenciación (novo o shoot-gun).....	33
6.3.3.1.- RT-PCR para síntesis de cDNA.....	34
6.3.3.2.- PCR para síntesis de dsDNA.....	36
6.3.3.3.- PCR para amplificación con oligonucleótidos M13.....	37
6.3.3.4.- Cuantificación de los Productos de PCR.....	38
6.3.3.5.- Estimación del tamaño de los fragmentos con equipo bioanalyzer.....	39
6.4.- Secuenciación masiva con plataforma Illumina.....	39
6.4.1.- Construcción de bibliotecas con oligonucleótidos 16S ribosomal.....	39
6.4.2.- Construcción de bibliotecas de novo.....	40
6.5.- Análisis bioinformático.....	41
6.5.1.- Análisis con Galaxy y Metavir.....	41
6.5.2.- Análisis con Ubuntu Biolinux.....	42
6.5.2.1.- Bases de datos.....	43
6.5.2.2.- Filtrado por calidad.....	43
6.5.2.2.1.- Proceso bioinformático A.....	44
6.5.2.2.2.- Proceso bioinformático B.....	45
6.5.2.2.3.- Proceso bioinformático C.....	46
6.5.2.2.3.1.- Ensamble de lecturas.....	47
6.5.3.- Análisis de 16S ribosomal.....	47
VII. RESULTADOS.....	48
7.1.- Número de pacientes y muestras analizadas.....	48
7.2.- Detección de virus con equipo BioMark.....	48
7.3.- Clasificación de muestras.....	49
7.4.- Cuantificación de los Productos de PCR.....	51
7.5.- Estimación de peso molecular.....	52
7.6.- Resultados de secuenciación en equipo Illumina.....	55
7.7.- Comparación de análisis bioinformáticos para viroma.....	56
7.7.1.- Análisis con Galaxy y Metavir.....	56
7.7.2.- Procesos en Biolinux.....	58
7.7.3.- Proceso bioinformático A, B y C.....	58
7.7.4.- Elección del procesamiento a utilizar.....	61
7.8.- Viroma y Bacterioma del pool 1 o adultos con agente etiológico viral.....	63
7.9.- Viroma y Bacterioma del pool 2 o adultos sin agente etiológico viral.....	65
7.10.- Viroma y Bacterioma del pool 3 o niños con agente etiológico viral.....	66

7.11.- Viroma y Bacterioma del pool 4 o niños sin agente etiológico viral.....	68
7.12.- Ensamblado de novo.....	70
VIII. DISCUSIÓN.....	71
8.1.- Toma de muestra y prevalencia de virus respiratorios.....	71
8.2.- Preparación de muestras para estrategia de secuenciación shot gun.....	73
8.3.- Comparación de los métodos bioinformáticos.....	75
8.4.- Comparación entre qPCR en Biomark y NGS.....	77
8.4.1.- Adultos y niños con presencia de virus con Biomark.....	78
8.5.- Comparación de Bacterias y fagos en todos los grupos de pacientes.....	81
8.5.1.- Adultos con presencia de virus.....	81
8.5.2.- Adultos sin presencia de virus.....	82
8.5.3.- Niños con presencia de virus.....	83
8.5.4.- Niños sin presencia de virus.....	84
8.6.- Comparación con otros estudios de la microbiota en IRA.....	86
8.7.- Importancia del estudio de la microbiota en humanos.....	87
IX. CONCLUSIONES.....	90
X. PERSPECTIVAS.....	92
XI. LITERATURA CITADA.....	94
XII. ANEXOS.....	108

ÍNDICE DE FIGURAS

Figura 1.- Flujo de trabajo desde bibliotecas hasta lectura en equipo Illumina	15
Figura 2.- Representación esquemática de los oligonucleótidos hexámeros aleatorios y hexámeros aleatorios M13.....	35

ÍNDICE DE TABLAS

Tabla 1.- Características de virus relacionados en IRA	4
Tabla 2.- Características de las bacterias implicadas en IRA	5
Tabla 3.- Estudios del microbioma respiratorio humano	17
Tabla 4.- Agrupamiento de muestras por característica del paciente	32
Tabla 5.- Detección de virus en el total de las muestras	49
Tabla 6.- Prevalencia de virus respiratorios en las muestras seleccionadas para su agrupamiento y secuenciación	50
Tabla 7.- Concentración de los productos de PCR de cada uno de los pools leídos con Nanodrop	51
Tabla 8.- Concentración de los productos para 16S de cada uno de los pools con Qubit ...	52
Tabla 9.- Estimación de la concentración y del tamaño de los productos para secuenciación de <i>novο</i>	54
Tabla 10.- Resultados obtenidos de las rondas de secuenciación	55
Tabla 11.- Lecturas obtenidas mediante el proceso de análisis bioinformático C	62
Tabla 12.- Lecturas obtenidas mediante el proceso de análisis de la aplicación 16S Metagenomics de Base Space	63

ÍNDICE DE GRÁFICAS

Grafica 1.- Productos obtenidos del grupo de adultos con virus y niños sin virus	53
Grafica 2.- Diversidad viral del grupo de niños con virus	57
Grafica 3.- Proporción de virus con procesos bioinformáticos A, B y C del archivo R2 del grupo de niños con presencia de virus	60
Gráfica 4.- Viroma y Bacterioma del grupo de adultos con virus	64
Gráfica 5.- Viroma y Bacterioma del grupo de adultos sin virus	65
Gráfica 6.- Viroma y Bacterioma del grupo de niños con virus	67
Gráfica 7.- Viroma y Bacterioma del grupo de niños sin virus	69

ÍNDICE DE DIAGRAMAS

Diagrama 1.- Flujo de trabajo para procesamiento de muestras	27
Diagrama 2.- Flujo de trabajo para análisis bioinformático	28

ÍNDICE DE MATERIAL SUPLEMENTARIO

Tabla A1. Mezcla de reacción 1, para hexámeros (izquierda) y hexámeros unidos a M13 (derecha)	108
Tabla A2.- Mezcla de reacción 2, donde se observa la concentración y volumen de cada componente de la reacción para realizar la retro transcripción	109
Tabla A3.- Mezcla de reacción para síntesis de dsDNA con fragmento klenow	110
Tabla A4. Mezcla de reacción para la amplificación de las secuencias M13	110
Tabla A5.- Mezcla de reacción para preparar productos de amplificación de 16S ribosomal	111

ABREVIATURAS

IRA	Infección Respiratoria Aguda
EPOC	Enfermedad Pulmonar Obstructiva Crónica
SSA	Secretaría de Salud
WHO	World Health Organization (Organización Mundial de la Salud)
Nm	nanómetros
Kb	kilobases
A.N.	Ácidos Nucleicos
µm	micrómetros
Mb	Megabaes
DNA	Ácido Desoxirribonucleico
RNA	Ácido Ribonucleico
RT	Retro transcripción
PCR	Reacción en Cadena de la Polimerasa, en ingles
Tm	Temperatura de Desnaturalización (temperature melting)
INER	Instituto Nacional de Enfermedades Respiratorias
BWA	Burrows Wheeler Aligner
GAAS	Genome Relative Abundance and Average Size
Sam	Sequence Aligment Map
BLAST	Basic Local Aligmen Search Tool
EDV68	Enterovirus D68
SARS	Síndrome Respiratorio Agudo Severo
MERS-CoV	Síndrome Respiratorio por Coronavirus de Oriente Medio
NGS	Secuenciación de nueva generación o secuenciación masiva

RESUMEN

La infección respiratoria aguda (IRA) es la enfermedad más recurrente a lo largo de la vida del ser humano. A pesar de contar con técnicas para el diagnóstico avaladas por organizaciones de salud, los agentes emergentes y reemergentes hacen necesario tener plataformas de secuenciación de siguiente generación ó secuenciación masiva (NGS) para la detección de virus y bacterias, de forma simultánea y total sin importar la variabilidad de su genoma.

Si se emplea NGS en condiciones estandarizadas, esta nos permitirá la detección y caracterización de cualquier patógeno presente en las IRAs sin importar su variabilidad genética, entonces podrá ser empleada para la detección de agentes etiológicos que no son detectados mediante metodologías de diagnóstico molecular.

El objetivo del estudio es desarrollar estrategias moleculares empleando NGS, a partir de muestras de hisopado nasofaríngeo, para la detección y caracterización genética de los agentes etiológicos que causan IRAs y su aplicación en el diagnóstico y vigilancia epidemiológica.

En 650 pacientes que requirieron de hospitalización por IRAs de octubre de 2013 a junio de 2014, se les realizó la detección de virus respiratorios con qPCR en Biomark. De acuerdo a lo anterior se escogieron 116 muestras para clasificarlas en adultos y niños y por la presencia o ausencia de virus. A estos cuatro grupos se les realizó la preparación de bibliotecas tipo shotgun para detectar al viroma y bibliotecas del amplicon 16S ribosomal para el bacterioma, siguiendo las indicaciones del protocolo Nextera XT de Illumina. El análisis de las lecturas obtenidas se realizó con herramientas bioinformáticas.

En el total de los pacientes incluidos en el estudio se utilizó la técnica de qPCR, con esta metodología se encontró que en el 35% había la presencia de virus respiratorios, los más frecuentes fueron influenza con 41%, rinovirus-enterovirus un 30% y virus sincicial respiratorio 11%. Del análisis metagenómico de pacientes pediátricos se pudo obtener una nueva cepa de rinovirus C, la presencia de virus reemergentes como Enterovirus D68 y Parvovirus B19, en la temporada invernal evaluada. En adultos y niños sin presencia de virus tenemos mayor presencia de bacterias residentes de la cavidad bucal, mientras que en adultos y niños con virus, encontramos mas bacterias residentes y patógenas del tracto respiratorio ó aquellas que se adquieren de forma nosocomial.

La NGS es una plataforma capaz de aportar información que las técnicas de diagnóstico moleculares no proveen. Con lo cual se obtuvo una nueva cepa de rinovirus C, así como agentes re-emergentes, como es el caso de Enterovirus D68 y Parvovirus B19 estos dos últimos provocaron recientemente brotes importantes en E.U.A. y Japón respectivamente. También se conoció la diversidad de virus, fagos y bacterias presentes en IRAs y aunque algunos de los microorganismos no se les asocia con la enfermedad, el papel de estos en los procesos infecciosos aún queda por estudiarse.

ASBTRACT

Acute respiratory infection (ARI) is the most recurrent disease throughout the life of the human being. Despite having diagnostic techniques supported by health organizations, emerging and re-emerging agents they do need to have platforms next generation sequencing or mass sequencing (NGS) for detecting viruses and bacteria, simultaneous and complete regardless of the variability of its genome.

If NGS is used in standardized conditions, this will allow us to detect and characterize any pathogens present in the IRAs regardless of their genetic variability, then can be used for the detection of etiologic agents that are not detected by methods of molecular diagnosis.

The aim of the study is to develop molecular strategies using NGS, from nasopharyngeal swabs samples for the detection and genetic characterization of the etiological agents that cause ARIs and its application in the diagnosis and surveillance.

In 650 patients requiring hospitalization for IRAs of October 2013 to June 2014 they were performed the detection of respiratory viruses with qPCR of Biomark. According to the above they were selected 116 samples to classify them in adults and children and the presence or absence of virus. These four groups underwent the preparation of such shotgun libraries to detect the viroma and amplicon 16S ribosomal libraries for bacterioma, following the indications of Nextera XT Illumina protocol. The analysis of the readings was performed with bioinformatics tools.

In the total of patients included in the study was used qPCR technique, with this methodology found that 35% had the presence of respiratory viruses, the most common with 41% were influenza, rhinovirus-enterovirus 30% and RSV 11%. Metagenomic analysis of pediatric patients could obtain a new strain of rhinovirus C, the presence of virus reemerging as enteroviruses and parvovirus B19 D68, evaluated in the winter season. In adults and children without the presence of virus we have a greater presence of resident bacteria of the oral cavity, whereas in adults and children with viruses, bacteria find more residents and respiratory tract pathogenic or those that are acquired nosocomial

The NGS is a platform able to provide information that molecular diagnostic techniques do not provide. Whereby a new strain of rhinoviruses C and re-emerging agents was obtained, as in the case of parvovirus B19 and Enterovirus D68 latter two recently caused serious outbreaks in Japan and USA respectively. Diversity of viruses, phages and bacteria present in ARIs is also known and although some microorganisms do not associated with disease, the role of these in infectious processes yet to be studied.

I. INTRODUCCIÓN

1.1.- Enfermedad respiratoria aguda

La infección respiratoria aguda (IRA), es la enfermedad más recurrente a lo largo de la vida del ser humano y su evolución es menor a 15 días en promedio (Eiros et al. 2009; “SSA” 2015). Se estima que cada año fallecen 3.9 millones de personas por esta causa (WHO 2015a). Las IRAs presentan diferentes sintomatologías, debido al agente etiológico que se encuentra presente durante la enfermedad, así como de factores propios del hospedero y del medio ambiente (“SSA” 2015).

Cuando las IRAs afectan al tracto respiratorio alto pueden causar otitis aguda, otitis media aguda, sinusitis, rinitis, faringoamigdalitis, laringotraqueitis y traqueítis, al afectar el tracto respiratorio bajo provoca bronquiolitis, neumonía y neumonía adquirida en la comunidad (Cosio Villegas et al. 2006).

Hay enfermedades crónicas degenerativas como el cáncer de pulmón, bronquios y tráquea, así como la enfermedad pulmonar obstructiva crónica (EPOC), asma y aquellos afectados por la tuberculosis, donde las IRAs pueden exacerbar su condición poniendo en riesgo la vida de quienes las padecen (WHO 2015b).

En México las IRAs constituyen un importante problema de salud pública, siendo la primera causa de morbilidad en el país y es responsable de la mayor cantidad de consultas en los centros de salud (“SSA” 2015).

1.2.- Agentes etiológicos

Los agentes que están asociados a las IRAs tienen principalmente etiología viral, bacteriana o en ocasiones puede ser causadas por coinfecciones (Eiros et al. 2009).

Se han identificado más de 200 virus diferentes distribuidos en seis familias, *Orthomyxoviridae*, *Paramyxoviridae*, *Picornaviridae*, *Coronaviridae*, *Adenoviridae* y *Parvoviridae*. En la parte de bacterias podemos encontrar a los géneros: *Streptococcus*, *Staphylococcus*, *Haemophilus*, *Legionella*, *Mycoplasma*, *Chlamydomphilia*, *Moraxella*, *Coxiella*, *Klebsiella* y *Bordetella* (Roth et al. 2004; van den Bergh et al. 2012). En las tablas 1 y 2 se muestran las principales características de los virus y bacterias implicados en IRAs respectivamente.

Tabla 1.- Características de virus relacionados en IRA.

Virus	Familia	Envoltura	Tamaño (nm) ^a	Material Genético	Longitud de an ^d (kb) ^e	Fuente
Adenovirus	<i>Adenoviridae</i>	No	65 a 80	dsDNA, linear ^b	23 a 46	(Ghebremedhin 2014)
Bocavirus	<i>Parvoviridae</i>	No	18 a 26	ssDNA, linear ^c	5 a 5.3	(Gurda et al. 2010; Lüsebrink et al. 2009; Schildgen 2013)
Coronavirus	<i>Coronaviridae</i>	No	120 a 160	ssRNA, linear polaridad positiva	26.4 a 31.7	(te Velthuis et al. 2010; Woo et al. 2010)
Enterovirus	<i>Picornaviridae</i>	No	30	ssRNA, linear polaridad positiva	7 a 8.5	(Tuthill et al. 2010; van der Schaar et al. 2013)
Rinovirus	<i>Picornaviridae</i>	No	30	ssRNA, linear polaridad positiva	7 a 8.5	(Tuthill et al. 2010)
Parainfluenza	<i>Paramyxoviridae</i>	Si	150 a 300	ssRNA, linear de polaridad negativa	15.5 a 17.4	(Schomacker et al. 2012; Kiss et al. 2014)
Influenza	<i>Orthomyxoviridae</i>	Si	100 a 300	ssRNA, linear de polaridad negativa	12 a 15	(Arias et al. 2009)
Virus Sincial Respiratorio	<i>Paramyxoviridae</i>	Si	100 a 1000	ssRNA, linear de polaridad negativa	15.2	(Liljeroos et al. 2013; Schomacker et al. 2012)
Metapneumo virus	<i>Paramyxoviridae</i>	Si	150 a 600	ssRNA, linear de polaridad negativa	13 a 20	(Panda et al. 2014)

- a) Nanómetros: nm
- b) Doble cadena: ds
- c) Una sola cadena: ss
- d) Ácidos nucleicos: an
- e) Kilobases: kb

Tabla 2.- Características de las bacterias implicadas en IRA.

Género de bacteria	Filo	Forma y tamaño (μm) ^a	Tipos de respiración	Genoma (Mb) ^b	Fuente
<i>Streptococcus</i>	<i>Firmicutes</i>	Forma oval y el extremo distal lanceolado, de 1.2 a 1.8	Anaerobio facultativo	2.04	(S. NCBI 2015; CDC 2015; Hoskins et al. 2001)
<i>Staphylococcus</i>	<i>Firmicutes</i>	Forma esférica agrupado en racimos, alrededor 1	Anaerobio facultativo	2.8	(«Staphylococcus.pdf»2015; Holden et al. 2004)
<i>Haemophilus</i>	<i>Proteobacteria</i>	Cocobacilos, de 1 de ancho y entre 0.5 y 2 de largo	Aerobios y anaerobios facultativos	2.2	(Sanger Institute 2015a; Dajani, Asmar, y Thirumoorthi 1979)
<i>Legionella</i>	<i>Proteobacteria</i>	Bacilos, de 0.5 a 0.7 de ancho y 2 de longitud	Aerobios	3.9	(Bender et al. 1990; Sánchez y Calderon Vega 2013)
<i>Mycoplasma</i>	<i>Tenericutes</i>	Redonda con una extensión puntiaguda sobresaliente, 0.3 a 0.8 de diámetro	Aerobios y anaerobios facultativos	0.81	(M. NCBI 2015; Dewall y Cheng 2011)
<i>Chlamydomphila</i>	<i>Chlamydiae</i>	Redondas u ovoides, 0.2 a 0.4	Aerobios o anaerobios facultativos	1.23	(Chlamydia NCBI 2015; Ostos Ortiz y Mélida Sánchez 2003)
<i>Moraxella</i>	<i>Proteobacteria</i>	Diplococos, de 0.6 a 1.0	Aerobios	1.7	(NCBI 2015; PHAC 2015; Nguyen, Hansen, y Farinha 1999)
<i>Coxiella</i>	<i>Proteobacteria</i>	Pleomórfico, de 0.2 a 0.7	Aerobios o anaerobios facultativos	2	(Coxiella NCBI 2015; Roca 2007)
<i>Klebsiella</i>	<i>Proteobacteria</i>	Pleomórfico, de 0.3 a 1 de ancho y de 0.6 a 6 de largo	Anaerobios facultativos	5.33	(K. NCBI 2015; K. PHAC 2015)
<i>Bordetella</i>	<i>Proteobacteria</i>	Cocobacilo, de 0.5 a 1	Aerobios	4.09	(NBCI 2015; B. PHAC 2015)

a) Micrómetros: μm

b) Megabases: Mb

En la tabla 1 podemos observar varias particularidades de los virus: su tamaño no sobrepasa un micrómetro, su material genético puede ser RNA de polaridad negativa o positiva, pero

pueden ser también de DNA de cadena sencilla o doble y la longitud de su genoma es menor a los 46 kb.

Al observar la tabla 2 podemos resaltar algunas de las características de las bacterias implicadas en las IRAs: en su mayoría pertenecen al filo *Proteobacteria*, todas son gram negativas excepto *Streptococcus* y *Staphylococcus*, miden desde 0.2 hasta los 6 micrómetros, la longitud de su genoma es de 0.81 hasta 5.33 Mb y este es de DNA, pero adicionalmente pueden tener plásmidos.

1.3.- Patogenicidad

El cuerpo humano alberga más 100 billones de microbios, una gran variedad de ellos se sugiere pueden contribuir a la nutrición del hospedero, a la regulación del desarrollo, a la protección contra patógenos y al desarrollo de la respuesta inmunológica (Johnson y Versalovic 2012).

La patogenicidad es un proceso que consiste en una serie de pasos que ocurren cuando un virus o bacteria infecta a su hospedero, resultando en eventos estocásticos que inducen la enfermedad y que en algunos individuos es aparentemente inofensiva (asintomáticos) (Fields 2013; Finlay y McFadden 2006).

1.3.1.- Factores del hospedero

Es importante resaltar que el hospedero tiene un papel fundamental en la evolución de la enfermedad (Tregoning y Schwarze 2010). A lo largo de la vida del ser humano las IRAs son recurrentes, pero los niños y adultos mayores la presentan con mayor frecuencia. Otros factores importantes son la nutrición de los individuos, la actividad física y comorbilidades como EPOC, fibrosis quística, asma y obesidad (Tregoning y Schwarze 2010; J. M. Beck, Young, y Huffnagle 2012).

A lo anterior se suman componentes ambientales que son: hacinamiento humano, densidad poblacional, temporada invernal y contacto con especies animales sean domésticas o silvestres (Zambon 2014; Tregoning y Schwarze 2010). Este último elemento es de suma importancia, ya que recientemente agentes reemergentes y emergentes de origen zoonótico como el Síndrome Respiratorio Agudo Grave (SARS), Síndrome respiratorio por coronavirus de Oriente Medio (MERS-CoV) y varios subtipos de influenza A: H10N8, H7N9, H6N1, H5N6, H5N1, H1N1 (cepa 2009), han causado en humanos desde casos esporádicos hasta la primera pandemia del siglo XXI. Alrededor del 75% de los nuevos patógenos descritos son virus de RNA, muchos de los cuales se transmiten a través de mucosas o vías respiratorias (Payne y Bellamy 2014; Hui y Zumla 2015).

1.3.2.- Factores de los agentes etiológicos

Las IRAs comienzan cuando un virus infecta a un hospedero susceptible y hay un súbito aumento en la carga viral acompañado de una respuesta inflamatoria significativa. La duración del periodo de incubación y manifestación de los síntomas varía según el agente etiológico y la condición del hospedero, pero en los pacientes con comorbilidades, inmunodeprimidos y sumando a los componentes medioambientales, en conjunto son factores que pueden aumentar la severidad de la enfermedad y prolongar la recuperación (“SSA” 2015; Hui y Zumla 2015; Tregoning y Schwarze 2010).

Otro factor importante es la región geográfica en donde se evalúen a las IRAs y el método de detección que permite conocer la causa, en estos casos entre un 60 a un 90% son causadas por virus (Berry, Gamielien, y Fielding 2015; Heikkinen y Järvinen 2003; Tran et al. 2012; Galindo-Fraga et al. 2013). Estos microorganismos pueden dañar el epitelio del pulmón, modificando las barreras físicas y alterando al sistema inmunológico, como consecuencia aumenta la susceptibilidad de una coinfección bacteriana.

Algunos ejemplos de las estrategias que utilizan los virus para contrarrestar la respuesta inmunológica del hospedero son:

La evasión de los anticuerpo de reconocimiento, la alteración de la señalización de las células del sistema inmunológico, sus mecanismos de latencia, entre otros (Finlay y McFadden 2006; Tregoning y Schwarze 2010). La microbiota del hospedero está compuesta por varias bacterias como las descritas en la tabla 2, por ejemplo *Streptococcus*, *Staphylococcus*, *Moraxella* y *Bordetella*. Las bacterias residentes en el humano cuentan con estrategias propias para evadir la respuesta inmunológica, por ejemplo: Inhibir la fagocitosis y la fusión del fagolisosoma, expresión de superantígenos y el bloqueo de la presentación de antígenos, entre otros. Todas las estrategias le permiten a las bacterias establecer una invasión en los tejidos susceptibles (Tregoning y Schwarze 2010; Johnson y Versalovic 2012; Finlay y McFadden 2006).

1.4.- Técnicas de diagnóstico

La necesidad de conocer el agente etiológico que causa IRA, es fundamental para permitir un adecuado tratamiento y manejo del paciente oportunamente, así como en el control de brotes epidémicos, el descubrimiento agentes patógenos o en casos de sintomatología donde no se conoce la causa (Zambon 2014; Payne y Bellamy 2014; Hui y Zumla 2015).

Las técnicas de diagnóstico se dividen básicamente entre las pruebas moleculares y las no moleculares, a continuación se describirán las no moleculares (Eiros et al. 2009; E. T. Beck y Henrickson 2010).

1.5.- No moleculares

El estándar de oro en el diagnóstico de la identificación de virus y bacterias es el cultivo, posteriormente se realiza la identificación mediante hemaglutinación, hemadsorción o técnicas de inmunofluorescencia para el caso de los virus y para las bacterias son una serie de pruebas bioquímicas en conjunto con la microscopía.

Estas técnicas son relativamente más baratas comparadas con las pruebas moleculares, pero tienen el inconveniente de tener una limitada sensibilidad, aunado a esto existen virus y bacterias que no crecen en los cultivos o líneas celulares y se necesitan de 3 a 14 días aproximadamente para obtener resultados (Eiros et al. 2009; E. T. Beck y Henrickson 2010; Epstein 2013).

La detección de antígenos se realiza mediante inmunofluorescencia, inmunoensayos o inmunocromatografía, los resultados pueden ser esperados en cuestión de minutos. Pero la sensibilidad y la especificidad de estas técnicas resultan bajas en comparación con las técnicas moleculares a pesar de contar con muestras de buena calidad (Eiros et al. 2009; E. T. Beck y Henrickson 2010).

1.6.- Moleculares

Las técnicas moleculares se basan fundamentalmente en la Reacción en Cadena de la Polimerasa (PCR) que tiene una gran gama de variaciones. La retrotranscripción (RT) para aquellos virus de RNA, nos permite construir la cadena complementaria de DNA y a partir de esta continuar con la PCR. La variación de ensayos múltiples nos permite dirigir la reacción a varios blancos al mismo tiempo, pero varios iniciadores dentro de la reacción pueden causar amplificaciones inespecíficas. El PCR degenerado se basa en la amplificación de muchos blancos utilizando un solo par de iniciadores, permite la detección de nuevos subtipos, pero el diseño de los iniciadores resulta difícil (Eiros et al. 2009; Wang y Taubenberger 2010).

La PCR en tiempo real es una variante de la PCR en la cual se hace uso de sondas que permiten el monitoreo de la amplificación de las moléculas blanco mediante la fluorescencia que estas sondas emiten. La ventaja es que nos permite cuantificar, hacer

detecciones múltiples y ahorrar tiempo en el análisis. Pero esta técnica necesita de equipos costosos ya que cuentan con un termociclador acoplado a un sistema óptico de detección de la fluorescencia emitida por las sondas (Eiros et al. 2009; Wang y Taubenberger 2010).

Una de las limitantes de la PCR en tiempo real es el número de ensayos que se podrían realizar dentro de la misma reacción, pues el número de fluorocromos que actualmente se pueden detectar es de hasta cinco en los equipos mas sofisticados, lo cual obliga a utilizar volúmenes de reactivos significativos al realizar el análisis de una muestra frente a un panel de reacciones distintas, usando los formatos y tubos de reacción convencionales.

1.7.- Técnicas moleculares comerciales

Comercialmente hay diversas variantes de la PCR que se encuentran disponibles para el diagnóstico de múltiples infecciones con la ventaja de ya haber sido validadas, entre las que podemos mencionar:

El ensayo xMAP®, nos permiten evaluar hasta 14 virus respiratorios; el panel viral respiratorio (Respiratory Viral Panel, RVP), usando Luminex xTAG, que puede detectar hasta 20 virus respiratorios; Allplex™ Respiratory Full Panel Assay que está diseñado para la detección de 19 virus respiratorios y 7 bacterias patógenas; RespiFinder® SMART 22 FAST es una técnica de RT-PCR múltiple en tiempo real, con lo cual podemos identificar 18 virus y 4 bacterias en un solo ensayo (Eiros et al. 2009; Kehl et al. 2001; Xiaoyan et al. 2014). Cabe señalar que todos estos ensayos en realidad detectan apenas una docena de patógenos en el mejor de los casos y el resto trata de tipos, subtipos y variantes.

Las técnicas anteriores necesitan de un personal capacitado para la extracción de los ácidos nucleicos, la preparación de cada una de las reacciones y la interpretación de los resultados.

Pero también hay sistemas automatizados como el FilmArray® en el cual solo tenemos que agregar una solución de hibridación, la muestra e insertar en el equipo, con esta tecnología se pueden identificar hasta 18 virus y 3 bacterias (BioFire Diagnostics 2015; Hammond et al. 2012).

Los microarreglos son otra variante de las técnicas moleculares que se basa en el uso de un substrato sólido que puede ser de vidrio o de silicón, al cual se le agregan miles de sondas de oligonucleótidos, estos son construidos a partir de secuencias conservadas, de virus, bacterias o genes.

Una vez fijadas las secuencias que van desde 25-mer hasta 70-mer, se pueden evaluar a las muestras marcadas con fluorescencia y que hibridan a los microarreglos, los patrones de fluorescencia son analizados y podemos detectar nuevos virus, por hibridación cruzada. Comercialmente se encuentran disponibles varios microarreglos, por ejemplo ViroChip y GreeneChip (Chiu 2013; Donatin y Drancourt 2012).

1.8.- Secuenciación masiva

Las técnicas comerciales para el diagnóstico están destinadas para utilizarse como apoyo en el diagnóstico en muchos hospitales en el mundo, pero los agentes emergentes como MERS-CoV, influenza A H10N8, H7N9, rara vez son detectados por que las técnicas están limitadas a la detección de secuencias virales y bacterianas preponderantes, altamente conservadas o previamente caracterizadas (Payne y Bellamy 2014; Hui y Zumla 2015; Chiu 2013)

A pesar de contar con una gran cantidad de pruebas comerciales de diagnóstico, hay casos en los que se desconoce la etiología que causa la enfermedad, lo cual dificulta un buen manejo del paciente, el control de brotes epidémicos e incluso la prescripción innecesaria de antibióticos (Tran et al. 2012; Galindo-Fraga et al. 2013).

El conocer el agente causal de las IRAs es de suma importancia y recientemente las plataformas de secuenciación masiva, se perfilan como una alternativa de detección para los agentes etiológicos, ya sean virus o bacterias presentes en muestras respiratorias de forma simultánea y total (Mardis 2013; Quiñones-Mateu et al. 2014; E. T. Beck y Henrickson 2010) .

Debido a que la NGS no depende de la secuencia específica como las técnicas de PCR y microarreglos, su empleo en muestras clínicas comienza a ser más frecuente para identificar y caracterizar no sólo nuevos patógenos, sino también la microbiota de diferentes muestras sin el sesgo inherente en técnicas de cultivo microbiano o en el uso de técnicas de PCR o microarreglos (Mardis 2013; Quiñones-Mateu et al. 2014; E. T. Beck y Henrickson 2010; Chiu 2013).

La secuenciación por Sanger es la metodología que se basa en la terminación de la cadena de DNA mediante la incorporación de dideoxynucleotidos, siendo esta la plataforma de primera generación. Las tecnologías de secuenciación masiva son las plataformas de segunda generación como 454 de Roche que se basa en la pirosecuenciación y Ion Torrent que se basa en la detección de protones liberados durante la polimerización del DNA.

Con las tecnologías de secuenciación masiva se han realizado diversos estudios del microbioma en océanos, animales y en el ser humano. En recientes años se ha incrementado la cantidad de estudios de la microbiota, mayormente relacionados con el tracto digestivo y aquellos donde se hace secuenciación shotgun, que es una técnica en la cual se hacen pequeños fragmentos de DNA que sirven para reconstruir genomas, sin el sesgo de amplificar amplicones como es el caso de 16S ribosomal (Rinke et al. 2013; Johnson y Versalovic 2012; Chiu 2013).

En la tabla 3 se muestran estudios realizados en el tracto respiratorio del ser humano en los últimos 6 años utilizando diversas plataformas. Destaca que los trabajos realizados por diversos grupos de investigación alrededor del mundo, principalmente se enfocan al estudio de 16S ribosomal, en pacientes con enfermedades asociadas como son la fibrosis quística,

EPOC o en pacientes fumadores y utilizando plataformas que van desde los microarreglos, secuenciación con sanger y hasta equipos de secuenciación masiva como son 454 GS-FLX o Illumina GAIIx.

1.9.- Secuenciación masiva basada en síntesis (Illumina)

La plataforma de Illumina MiSeq presenta la menor tasa de error $\sim 0.1\%$, puede cubrir una longitud de lectura de hasta 600 pares de bases actualmante y es la segunda plataforma con mayor cantidad de publicaciones relacionadas con virus, solo por debajo de la plataforma de 454, pero esta última tiene como desventajas una mayor tasa de error y un costo elevado (Ledergerber y Dessimoz 2011; Liu et al. 2012; Quiñones-Mateu et al. 2014)

1.9.1.- Fundamento de la metodología de Illumina MiSeq

Primero se realiza una fragmentación del material genético que poseé la muestra y que servirá para construir las bibliotecas que posteriormente se inyectarán en la celda de flujo o flow cell, en condiciones desnaturalizantes para tener cadenas sencillas de DNA, y que bajo un gradiente de temperatura permite la hibridación de los adaptadores complementarios y de los barcodes añadidos.

Los complementos de los adaptadores se encuentran en la superficie de la flow cell y permiten la unión de una sola molécula de DNA de forma equidistante gracias a la proporción de moléculas añadidas, por eso es de gran importancia seguir las condiciones del protocolo Nextera XT para asegurar la concentración necesaria para tal motivo (Ledergerber y Dessimoz 2011; Liu et al. 2012; Mardis 2008; Mardis 2013).

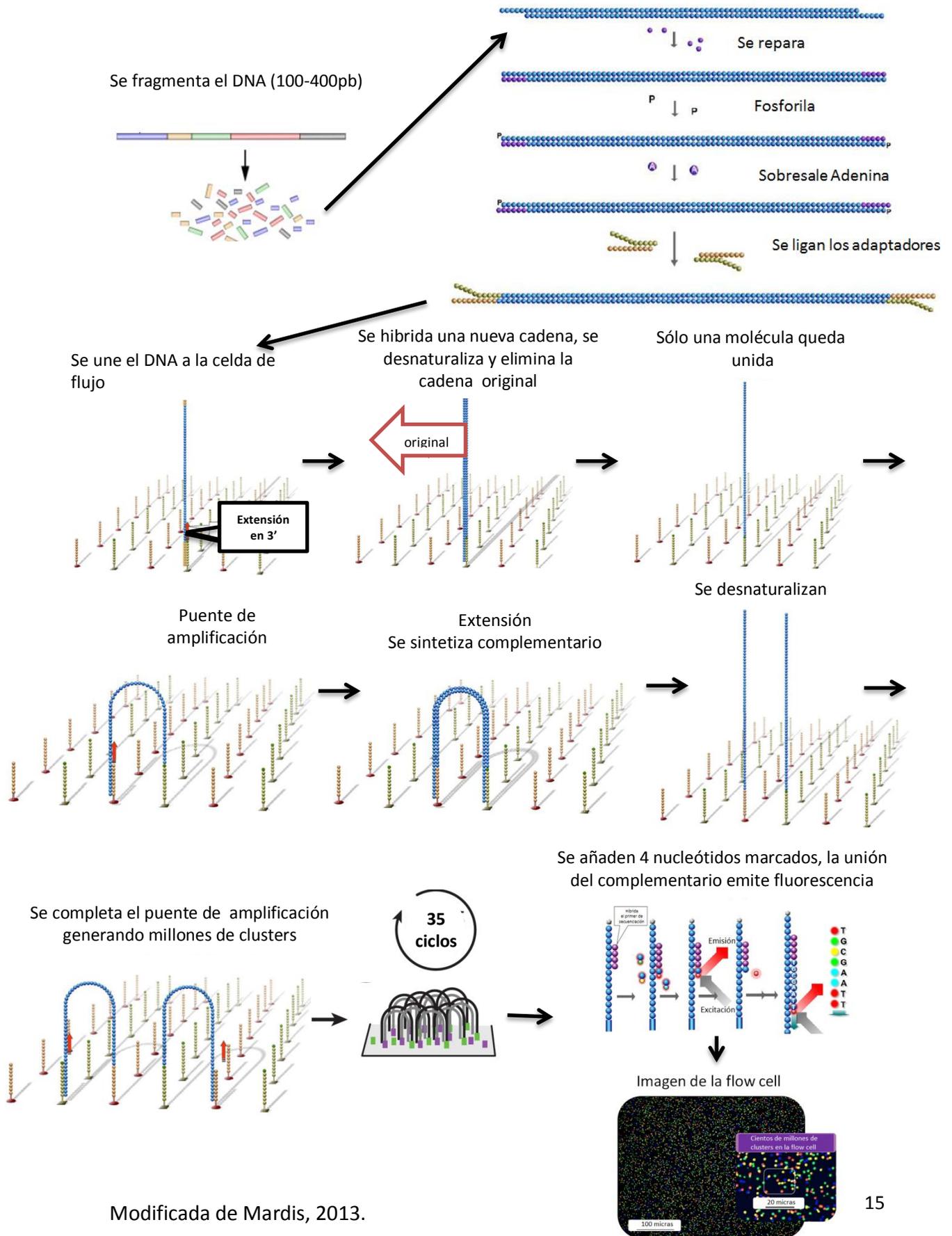
Una vez unidas las moléculas se hibrida la cadena complementaria, y nuevamente se genera una condición desnaturalizante para descartar la molécula original (Ledergerber y Dessimoz 2011; Liu et al. 2012; Mardis 2008; Mardis 2013).

A partir de la molécula que se queda unida se generan ciclos de alineamiento o puentes de amplificación, extensión, y desnaturalización, para completar millones de grupos o clusters provenientes de una sola cadena. La última desnaturalización deja a las moléculas de forma lineal y se agregan bloqueadores en los extremos 3' (Ledergerber y Dessimoz 2011; Liu et al. 2012; Mardis 2008; Mardis 2013).

Cuatro nucleótidos marcados con un fluoroforo distinto y una DNA polimerasa de alta fidelidad permiten en un solo ciclo agregar el nucleótido complementario a la secuencia correspondiente en cada uno de los clusters generados. Mediante la excitación de los fluoroforos con un láser se capta al mismo tiempo con los 4 detectores del equipo la incorporación del nucleótido según la secuencia y se captura mediante una fotografía en cada ciclo (Ledergerber y Dessimoz 2011; Liu et al. 2012; Mardis 2008; Mardis 2013).

Cada ciclo corresponde a un nucleótido añadido, por lo tanto la secuencia de cada molécula se construye a partir de cada ciclo, este se repite de acuerdo a la longitud de la secuencia deseada, que para nuestro caso fueron 2 rondas de 250 ciclos (pair end), debido a este par de ciclos se generaron dos archivos, R1 para el primer ciclo y R2 para el segundo. Los archivos que se obtuvieron por el equipo MiSeq se encuentran en formato fastq, los cuales servirán posteriormente para su análisis (Ledergerber y Dessimoz 2011; Liu et al. 2012; Mardis 2008; Mardis 2013). En la figura 1 se muestra de forma esquemática el proceso de secuenciación desde la generación de las bibliotecas.

Figura 1.- Preparación de bibliotecas y proceso de lectura en equipo Illumina



Modificada de Mardis, 2013.

II. ANTECEDENTES

El concepto microbiota es usado para referirse a los microbios, mientras que microbioma se usa para referirse a los microorganismos, genes e interacciones ambientales de comunidades de bacterias, arqueas, virus, fagos, hongos y otros microbios eucariontes en undeterminado nicho ecológico. A su vez el viroma humano es definido por los componentes virales y el bacterioma humano esta definido por las bacterias, por lo tanto ambos componentes pueden conformar la microbiota de un cierto nicho anatomico (Cox, Cookson, y Moffatt 2013; Lecuit y Eloit 2013).

Si se observan los trabajos realizados a partir de 2014 en la tabla 4, el uso de los equipos de secuenciación masiva es más común, debido principalmente a un costo cada vez más reducido para el uso de estas técnicas.

Particularmente en México Cobian y colabs estudiaron pacientes con y sin síndrome respiratorio severo de la epidemia de influenza de 2009 en el D.F. y Taboada y colabs comenzaron a realizar estudios de secuenciación masiva en muestras respiratorias de 2010-2011 y 2011-2012, de pacientes pediátricos con padecimientos respiratorios que acudieron a un hospital de Veracruz.

Los resultados de estos trabajos y los demás mencionados en la tabla 4 son diversos y en general no hay evidencia del descubrimiento de un virus nuevo asociado a IRA, excepto en el trabajo de descrito Zaki y colabs donde se secuenció el genoma de MERS-CoV mediante 454 GS-FLX, lo cual demuestra la capacidad de las técnicas de secuenciación masiva para el descubrimiento de nuevos patógenos.

Tabla 3.- Estudios del microbioma respiratorio humano.

Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Detección metagenómica de patógenos virales en muestras nasales y fecales	454 GS-FLX	Shotgun	(Nakamura et al. 2009)
Análisis de las comunidades virales de DNA en personas con y sin fibrosis quística	454 GS-FLX	Shotgun	(Willner et al. 2009)
Comparar las características del microbioma de sujetos sanos, asma y EPOC	Sanger	16S	(Hilty et al. 2010)
Estudiar integralmente el perfil bacteriano del tracto respiratorio de pacientes con EPOC	PhyloChip	16S	(Huang et al. 2010)
Análisis comparativo de la microbiota bacteriana de las fosas nasales y orofaringe de adultos sanos	PhyloChip y Sanger	16S	(Lemon et al. 2010)
Caracterizar las enfermedades similares a influenza en muestras nasales y de garganta de niños de Nicaragua de 2007 a 2008	Illumina GAIIX	Amplicones	(Yozwiak et al. 2010)
Evaluar individuos sanos y enfermos, colonizados con S. aureus	Sanger y 454 GS-FLX	16S	(Frank et al. 2010)
Análisis metagenómico de la influenza A pandémica y comparación entre virochip y NGS.	Illumina GAIIX	Amplicones	(Greninger et al. 2010)

Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Analizar la microbiota de la narina derecha e izquierda de la nasofaringe y la orofaringe entre fumadores y no fumadores	454 GS-FLX	16S	(Charlson et al. 2010)
Determinar si las diferencias geográficas influyen en la composición de las comunidades bacterianas, presentes en vías respiratorias de pacientes con fibrosis quística	T-RFLP	16S	(Stressmann et al. 2011)
Evaluar la comunidad bacteriana entre individuos sanos, fumadores y con EPOC	454 GS-FLX	16S	(Erb-Downward et al. 2011)
Estudiar la variabilidad y diversidad de la microbiota de la nasofaringe de niños sanos	454 GS-FLX	16S	(Bogaert et al. 2011)
Utilizar métodos moleculares para definir la microbiota bacteriana presente en los pulmones de individuos sanos y evaluar su relación con las poblaciones de las vías respiratorias superiores	454 GS-FLX	16S	(Charlson et al. 2011)
Secuenciar y caracterizar el viroma de niños y adultos hospitalizados por severas infecciones del tracto respiratorio bajo	454 GS20 Y GS-FLX	16S y shotgun	(Lysholm et al. 2012)
Investigar la dinámica de la comunidad microbiana en seis pacientes con fibrosis quística a lo largo de 8 a 9 años	454 GS-FLX	16S	(Zhao et al. 2012)

Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Caracterizar la comunidad bacteriana de pacientes no fumadores, fumadores, fibrosis quística y EPOC severo	454 GS-FLX	16S	(Sze et al. 2012)
Análisis del viroma en niños febriles y afebriles	454 GS-FLX e Illumina GAIIX	Shotgun	(Wylie et al. 2012)
Comparar los taxa de vías respiratorias superiores y bajas en pacientes con diferentes estadios de la fibrosis quística	454 GS-FLX	16S	(Goddard et al. 2012)
Estudiar el microbioma respiratorio e intestinal en lactantes en desarrollo y con fibrosis quística, con el propósito de promover y mantener un microbioma resistente a colonizaciones por bacterias patógenas	454 GS-FLX	16S	(Madan et al. 2012)
Evaluar las poblaciones bacterianas en el pulmón, mediante análisis repetidos de pacientes sanos de vías respiratorias altas y bajas	454 GS-FLX	16S	(Charlson et al. 2012)
El papel del microbioma en pacientes con EPOC moderado y severo, así como en pacientes sanos	454 GS-FLX	16S	(Pragman et al. 2012)
Aislar nuevo Coronavirus de un hombre con neumonía en Arabia Saudita	454 GS-FLX	Shotgun	(Zaki et al. 2012)
Identificar los patógenos virales y bacterianos presentes en muestras obtenidas de pacientes con síndrome respiratorio severo positivo y negativo a influenza	Illumina GAIIX	Amplicones, Shotgun y 16S	(Cobian Guemes y Arias-Ortiz 2013)
Determinar la relación entre la vía aérea, inflamación y	454 GS-FLX	16S	(Zemanick et al. 2013)

función del pulmón en pacientes con fibrosis quística tratados por exacerbación			
Comparar el microbioma del tracto respiratorio superior y bajo de pacientes no fumadores y fumadores	454 GS-FLX	16S	(Morris et al. 2013)
Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Determinar si la inmunodeficiencia causada por VIH resulta en la alteración del microbioma de pulmón	Illumina HiSeq	16S	(Lozupone et al. 2013)
Los cambios de la microbiota de la vía respiratoria de pacientes con fibrosis quística durante una exacerbación	454 GS-FLX	16S	(Carmody et al. 2013)
El propósito de este estudio fue investigar la biología microbiana de la laringe y el análisis de la biodiversidad de la garganta en pacientes con carcinoma de laringe y pacientes sanos	454 GS-FLX	16S	(Gong et al. 2013)
Examinar las bacterias presentes en la microbiota en lavados bronquio alveolares de pacientes que sufrieron lesión traumática	454 GS-FLX	16S	(Huebinger et al. 2013)
Caracterizar el microbioma del tracto respiratorio alto utilizando el blanco universal cpn60 en pacientes con influenza A H1N1 pandémica	454 GS-FLX	Amplicón	(Chaban et al. 2013)
Analizar detalladamente la estructura de la comunidad microbiana del pulmón de individuos sanos y con EPOC	Sanger	16S	(Zakharkina et al. 2013)

Evaluar el riesgo de microaspiración de microorganismos orales de individuos no fumadores, fumadores y ex fumadores	454 GS-FLX	16S	(Segal et al. 2013)
Investigar el efecto de la exacerbación por rinovirus en pacientes con EPOC y la modificación en el microbioma	454 Junior	16S	(Molyneaux et al. 2013)
Identificar bacterias asociadas a pacientes con tuberculosis que tienen diferentes resultados en su tratamiento	454 GS-FLX	16S	(Wu et al. 2013)
Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Analizar la prevalencia de las bacterias en pacientes con y sin fibrosis quística	454 GS-FLX	16S	(Twomey et al. 2013)
Estudiar las vías respiratorias medias de pacientes sanos para ver la riqueza y diversidad de su microbioma	454 GS-FLX	16S	(Ramakrishnan et al. 2013)
Genómica comparativa de aislados de Pseudomonas aureginosa en pacientes con fibrosis quística	454 GS-FLX	Shotgun	(Jeukens et al. 2014)
Evaluar la utilidad de la secuenciación del genoma, aplicado a pacientes con fibrosis quística	Illumina TruSeq	16S	(Hauser et al. 2014)
Utilizar RT-PCR y NGS para caracterizar la muestra de un adulto con enfermedad similar a influenza	454 GS-FLX	Shotgun	(Bialasiewicz et al. 2014)
Comprender la composición y dinámica de la microbiota al inocular rinovirus en voluntarios sanos	454 GS-FLX	16S	(Allen et al. 2014)
Hacer una ultrarápida identificación de patógenos basados en su secuencia con muestras diversas: cáncer de	Illumina TruSeq	Shotgun	(Naccache et al. 2014)

próstata y colon, hepatitis, diarrea, sangre, neumonía e influenza			
Investigar la dinámica de la microbioma bacteriana de las vías respiratorias, en pacientes con EPOC al inicio y después de una exacerbación	PhyloChip e Illumina MiSeq	16S	(Huang et al. 2014)
Comparación entre RT-PCR y NGS para la detección de virus respiratorios	454 Junior y GS-FLX	Shotgun	(Prachayangprecha et al. 2014)
Buscar nuevos patógenos virales en pediátricos con infecciones del tracto respiratorio	Illumina GAIIX	Shotgun	(Taboada et al. 2014)
Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Caracterizar y comparar las comunidades bacterianas en pacientes con lesiones bucales benignas	454 Junior	16S	(Hanshew, Jetté, y Thibeault 2014)
Describir y comparar la comunidad microbiana en vías respiratorias bajas de niños con y sin fibrosis quística	PhyloChip	16S	(Renwick et al. 2014)
Comparar la composición de la microbiota en adultos sanos que adquirieron naturalmente o no neumococos y si la microbioma nasofaríngeo es perturbado por la exposición a S. pneumoniae	454 GS-FLX	16S	(Cremers et al. 2014)
Caracterizar el microbioma de pulmón de individuos sanos y comparar con datos del Proyecto del Microbioma Humano	454 Junior	16S	(Venkataraman et al. 2015)
Aproximación metagenómica para caracterizar las poblaciones de bacteriófagos en pacientes con y sin	Illumina MiSeq	Shotgun	(Tariq et al. 2015)

fibrosis quística

Utilizar secuenciación para determinar el viroma de muestras clínicas para el diagnóstico clínico de infecciones virales.	Ion Torrent PGM	Shotgun	(Zoll et al. 2015)
Analizar en un sujeto las mucosas de boca y nariz como fuente de las comunidades del microbioma de estómago y pulmón	454 GS-FLX	16S	(Bassis et al. 2015)
Examinar la diferencia entre la comunidad bacteriana de niños y adultos mayores con y sin fibrosis quística	Illumina GAIIX	16S	(Coburn et al. 2015)
Objetivo del estudio	Plataforma de análisis	Estrategia	Autor y año
Revisión de la diversidad de patógenos microbianos y sus genes de resistencia a antibióticos presentes en la India	Illumina HiSeq	Shotgun	(Jalali et al. 2015)

III. JUSTIFICACIÓN

La falta de metodologías capaces de detectar agentes emergentes, así como la variedad de virus y bacterias presentes en muestras respiratorias de forma simultánea y total, hacen que la NGS pueda ser una alternativa en el diagnóstico de las IRAs, ya que estas plataformas no dependen de la secuencia y de la variabilidad del genoma de los microorganismos. Las plataformas están diseñadas para secuenciar estos genomas a un costo y tiempo considerablemente menor comparado con la secuenciación por Sanger.

Por lo tanto con este proyecto se desea conocer a los patógenos presentes en pacientes con enfermedad respiratoria aguda y con base en este conocimiento proponer estrategias de secuenciación, proporcionar bases de datos de microorganismos, en un corto tiempo y a bajo costo en comparación con la técnica de Sanger. De esta forma podremos detectar y caracterizar virus y bacterias presentes en muestras respiratorias, aspecto indispensable para confrontar una nueva epidemia con más herramientas y que nos acerca al ideal de contar con toda una gama de pruebas para gérmenes respiratorios.

IV. HIPÓTESIS

Si se emplea la Secuenciación de Nueva Generación en condiciones estandarizadas, esta nos permitirá la detección y caracterización de cualquier patógeno presente en las IRAs sin importar su variabilidad genética, entonces podrá ser empleada para la detección de agentes etiológicos que no son detectados mediante metodologías de diagnóstico molecular.

V. OBJETIVOS

General:

Estandarizar una serie de metodologías de extracción de ácidos nucleicos, producción de bibliotecas y secuenciación de nueva generación, a partir de muestras de hisopado nasofaríngeo, para detectar de manera simultánea cualquier tipo de virus y bacterias respiratorias. Este desarrollo tecnológico servirá para implementar metodologías modernas de diagnóstico y de vigilancia epidemiológica, así como detectar y caracterizar genéticamente nuevos patógenos asociados a IRAs.

Específicos:

Optimizar los métodos de extracción de RNA y DNA a partir de muestras respiratorias de hisopado nasofaríngeo de pacientes pediátricos y adultos con enfermedad respiratoria aguda.

Construir bibliotecas tipo shotgun de cDNA y dsDNA provenientes de virus con genoma de RNA y DNA respectivamente, así como bibliotecas de amplicones del gen 16S a partir de DNA genómico bacteriano, así como la estandarización de las condiciones óptimas para la realización de la NGS.

Identificar los patógenos presentes en las muestras respiratorias por medio de recursos bioinformáticos y asociarlos con presencia/ausencia o grado de severidad de la enfermedad respiratoria aguda.

VI. METODOLOGÍA

Diagrama 1.-Flujo de trabajo para procesamiento de muestras.

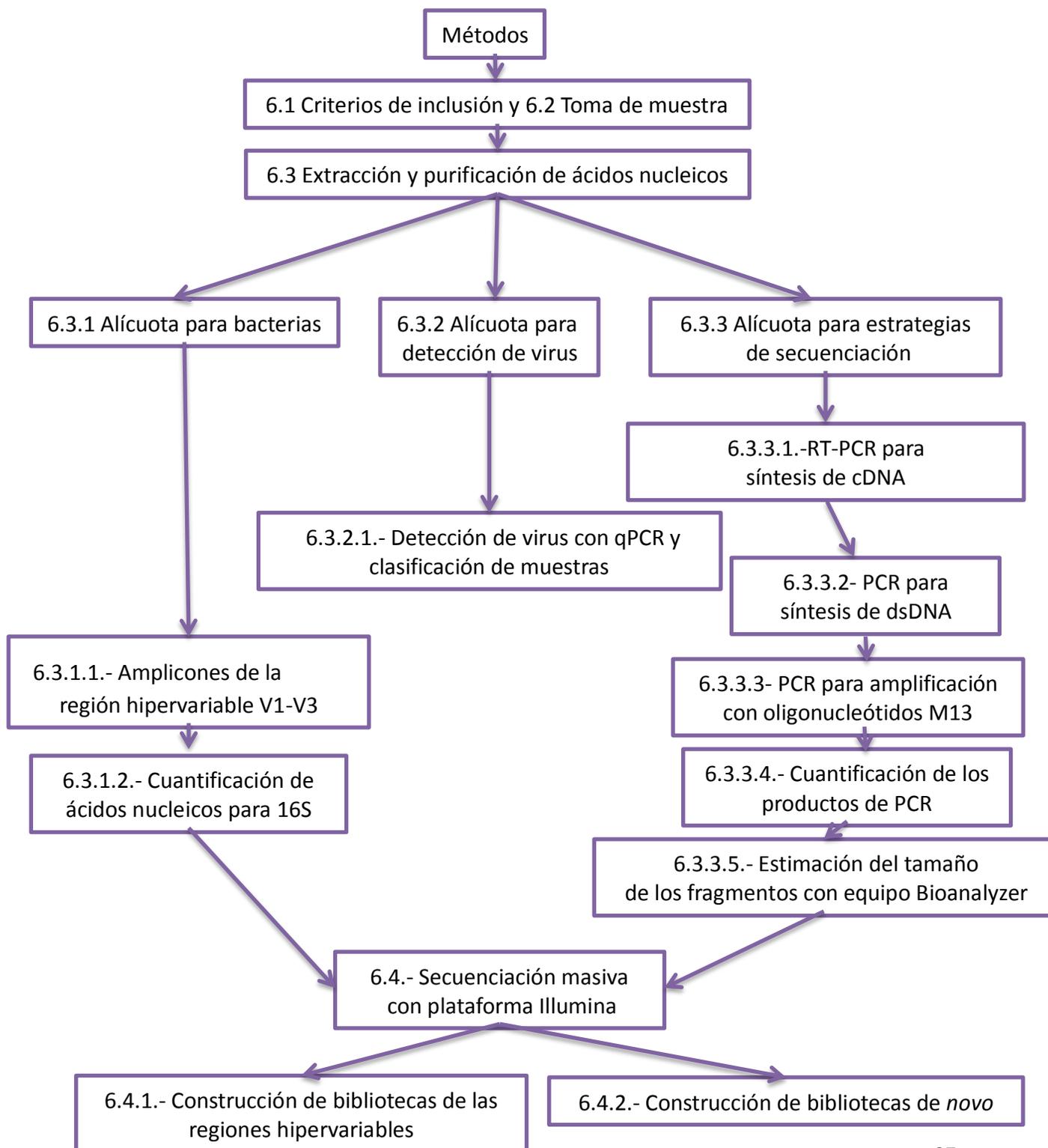
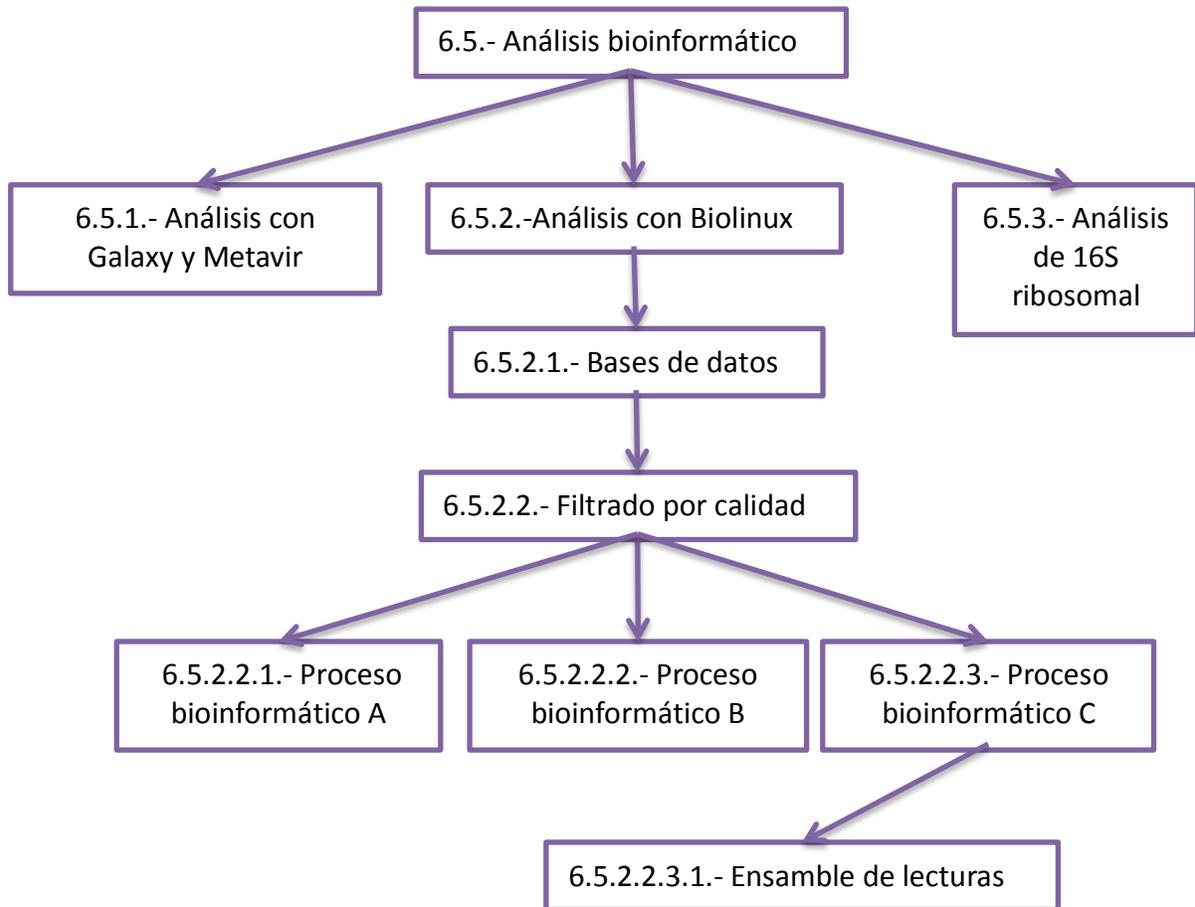


Diagrama 2.- Flujo de trabajo para el análisis bioinformático.



6.1.- Criterios de inclusión

Este estudio tuvo un diseño transversal y prospectivo, ya que se utilizaron muestras de pacientes de la temporada invernal 2013-2014, que hubieran aceptado mediante consentimiento informado participar en el protocolo de investigación (código B26-13). El presente proyecto fue aprobado por el Comité de Ética en Investigación del Instituto Nacional de Enfermedades Respiratorias. Se consideró dentro del grupo de los adultos a aquellas personas mayores de 15 años y dentro del grupo de niños a los que tenían edad menor a los 15 años, de acuerdo al criterio de admisión del Instituto. Para el caso de los menores de edad el consentimiento informado fue otorgado por sus padres o tutores.

6.2.- Toma de muestra

Debido que el periodo de excreción viral es, en general muy corto, la toma de muestra fue entre las primeras 48 a 72 horas después del comienzo de los síntomas, para asegurar obtener la mayor cantidad de virus posible (Eiros et al. 2009).

El procedimiento de la toma de muestras fue el siguiente: al paciente sentado o recostado, se le elevó un poco su cabeza. Suavemente el hisopo estéril con mango flexible de plástico y punta de dacron o rayón, se introdujo por la nariz paralelo al paladar, casi en su totalidad hasta llegar a la nasofaringe (aproximadamente 2.5 cm en adulto y un poco menos en niños); para llegar hasta la parte trasera donde se encuentran los cornetes. Una vez ahí se rotó suavemente el hisopo para frotar las paredes internas y obtener una mayor cantidad de células infectadas con el virus. Posteriormente se retiró cuidadosamente el hisopo sin dejar de rotar («SINAVE.pdf» 2015; «Copan» 2015) .

Para mantener la integridad de los virus lo mejor posible, se depositó la punta del hisopo en un tubo con medio de transporte viral, posteriormente el resto del mango se cortó y se cerró el tubo, conservándolo a una temperatura de 4°C y de no procesarse la muestra en las horas posteriores a la toma, ésta se congeló a -80°C (“SINAVE.pdf” 2015).

6.3.- Extracción y purificación de ácidos nucleicos.

Del total de la muestra obtenida por cada paciente, se separaron tres alícuotas:

6.3.1.- Alícuota para bacterias.

La primera alícuota se guardó a -80 °C hasta que se realizó la extracción de los ácidos nucleicos de las bacterias presentes en cada una de las muestras. La extracción de ácidos nucleicos bacterianos se realizó con QIAamp ® Cador ® Pathogen Mini Kit (QIAGEN 2015), En condiciones altamente desnaturizantes. A cada muestra se le adicionó el buffer ACB para ajustar las condiciones para la purificación de los ácidos nucleicos. Se agregó todo este lisado a las columnas y se sometió a una centrifugación. Los ácidos nucleicos son absorbidos sobre las membranas de sílice mientras que los contaminantes pasan a través de éstas. Se realizaron dos pasos de lavados para eliminar eficientemente los contaminantes restantes e inhibidores de enzimas. Finalmente los ácidos nucleicos fueron eluidos en un buffer de baja fuerza iónica. Los ácidos nucleicos obtenidos fueron usados posteriormente para construir bibliotecas de 16S ribosomal.

6.3.1.1.- Amplicones de la región hipervariable V1-V3.

Se realizó la preparación de las regiones hipervariables de 16S ribosomal, de acuerdo al protocolo 16S Metagenomic Sequencing Library Preparation de illumina (Illumina 2015), la enzima recomendada se cambió por Platinum Taq Polymerase (Thermo Fisher Scientific 2015d) y los iniciadores utilizados se sintetizaron de acuerdo a las instrucciones del mismo protocolo, la amplificación de la región hipervariable V1-V3 (460pb), incluyendo el

respectivo adaptador que recomienda el mismo protocolo. La mezcla de reacción se muestra en la tabla A5 de anexos.

Para esta amplificación se utilizó el termociclador Veriti (Thermo Fisher Scientific 2015a) de acuerdo a las siguientes condiciones: desnaturalización inicial 94°C durante 2 minutos; amplificación de 25 ciclos: 94°C durante 30 segundos, 55°C durante 30 segundos, 72°C durante 30 segundos; elongación 72°C durante 5 minutos y enfriamiento a 4°C (Illumina 2015).

6.3.1.2.- Cuantificación de ácidos nucleicos para 16S

Se utilizó Qubit® 2.0 Fluorometer (Thermo Fisher Scientific 2015f). Es un fluorómetro que utiliza fluorocromos específicos para medir ácidos nucleicos, permite la cuantificación del RNA del 16S ribosomal de forma específica, por lo tanto las mediciones con este equipo son mucho más precisas.

6.3.2.- Alícuota para detección de virus.

La segunda alícuota se utilizó durante las horas inmediatamente posteriores a la toma de muestra o después de su almacenamiento a -80 °C, para su procesamiento en el equipo automatizado MagNA Pure LC 2.0 Instrument (La Roche 2015a).

El procedimiento de aislamiento se basó en la tecnología de perlas magnéticas. Brevemente las muestras se lisaron por incubación con un buffer que contiene una sal caotrópica y proteinasa K. Posteriormente se agregaron partículas magnéticas de cristal que permiten la

adhesión de los ácidos nucleicos en su superficie. Los restos celulares y sustancias que no se adhirieron se eliminaron por varios procesos de lavados. Finalmente se obtuvo RNA y DNA purificados que se diluyeron con un buffer de baja fuerza iónica. Posteriormente con estos ácidos nucleicos se realizó un tamizado con arreglo dinámico de nanofluidos, en el equipo BioMark\u2122 HD System de Fluidigm (ver sección 6.3.2.1).

6.3.2.1.- Detección de virus con qPCR y clasificación de muestras

Con el equipo de Fluidigm de Biomark se evaluaron 48 blancos distintos de las seis principales familias de virus, de acuerdo a la metodología descrita por (Orozco Sanchez y Vázquez Pérez 2012). Del total de las muestras analizadas se decidió clasificar a las muestras entre positivas ó negativas a virus y de acuerdo a la edad, es decir mayores o menores de 15 años, como se muestra en la tabla 4.

Tabla 4.- Agrupamiento de muestras por característica del paciente.

Sector de la población	Presencia de agente etiológico viral	Grupo o pool
Adultos	Si	1
	No	2
Niños	Si	3
	No	4

La clasificación anterior sirvió para escoger entre el total de pacientes reclutados a los cuatro grupos que se utilizaron para realizar la secuenciación de *novo* y para 16S ribosomal para conocer el viroma y el bacterioma respectivamente. La principal finalidad de hacer

grupos fue tratar de sumar los ácidos nucleicos presentes y tener mayores posibilidades de secuenciar estas moléculas presentes, debido a que se tiene el antecedente en muchos de los estudios de la tabla 3 realizados por shot gun, que la cantidad lecturas obtenidas al final de la secuenciación son muy pocas (Lysholm et al. 2012; Taboada et al. 2014; Zoll et al. 2015).

Cada grupo se propuso realizarlo con 10 pacientes como mínimo y con un máximo de 30, el primer grupo ó de adultos con detección de virus, el grupo 2 ó adultos sin virus detectado, el grupo 3 ó niños con detección de virus y el grupo 4 ó niños sin detección de virus. Recordando que la detección de virus se realizó con qPCR en equipo Biomark, esto con el propósito de tener las dos condiciones, pacientes con y sin virus detectados y a su vez adultos y niños. Esto se realizó tomando como referencia el trabajo de Taboada y colabs y con la misma finalidad de saber si hay virus que no se detectan mediante técnicas de diagnóstico molecular, otro propósito fue la detección y caracterización de nuevos patógenos asociados a IRAs, para este fin se buscó principalmente en pacientes con detección negativa a virus por qPCR.

6.3.3.-Alicuota para estrategias de secuenciación (de *novo* o shot-gun)

La tercera alícuota se utilizó enseguida de la toma de muestra para extraer y purificar ácidos nucleicos de virus. Se realizó una centrifugación a baja velocidad a 5000 rpm durante 10 minutos, se tomó el sobrenadante y se filtró por membranas de 0.20 μm (Lysholm et al. 2012). A este sobrenadante se le realizó tratamiento con DNAsas (Thermo Fisher Scientific 2015b), la cual es una endonucleasa que digiere cadenas simples y dobles de DNA con el propósito de eliminar DNA humano y dejar todo aquello que este protegido por lipoproteínas como es el caso de los viriones.

Las partículas virales se procesaron con PureLink® Viral RNA/DNA Mini Kit (Thermo Fisher Scientific 2015e). Brevemente las muestras se lisaron usando proteinasa K y el buffer L22, cuya mezcla de buffers formulados específicamente para permitir la lisis de las diferentes partículas de virus. Conteniendo además, 5 µL de acrilamida (1U/µL) a 56°C, en lugar de RNA acarreador. La acrilamida se utilizó como un acarreador neutro y eficiente para cantidades de hasta picogramos de ácidos nucleicos de interés. (Lysholm et al. 2012; «Applied Biosystems» 2015).

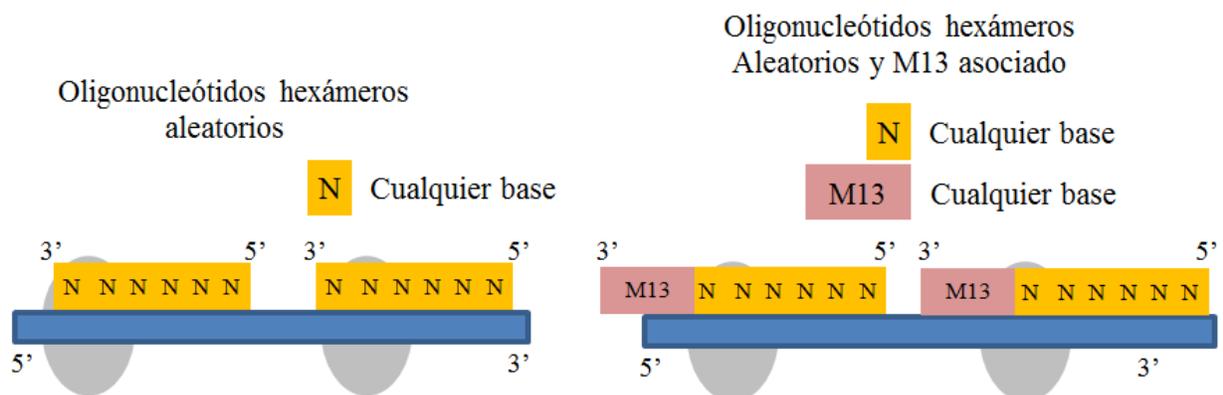
Posteriormente se añadió etanol a temperatura ambiente y el total del lisado se pasó por una columna de sílice. Las moléculas de RNA y DNA se unen a la base de sílice y mediante centrifugación y lavados, las impurezas como proteínas y nucleasas son eliminados. Finalmente los ácidos nucleicos se diluyeron con agua libre de RNAsas (Thermo Fisher Scientific 2015e). El producto obtenido de este paso se dividió a su vez en dos partes, la primera para procesar virus de RNA como se muestra en la sección 6.3.3.1 y para procesar virus de DNA como se describe en la sección 6.3.3.2.

6.3.3.1.-RT-PCR para síntesis de cDNA

En el caso de los virus con genoma de RNA presentes en cada grupo, se sintetizó DNA complementario utilizando el estuche First Strand cDNA Synthesis Kit (La Roche 2015b). En este paso se realizó la síntesis del cDNA y se aprovechó para utilizar oligonucleótidos aleatorios de seis nucleótidos o hexámeros aleatorios propios del ensayo comercial, así como con oligonucleótidos hexámeros aleatorios unidos a una secuencia de bacteriófago M13. El propósito de utilizar oligonucleótidos hexámeros aleatorios y hexámeros aleatorios unidos a M13 es aumentar la cantidad de moléculas de virus de RNA ya que en estudios anteriores se ha descrito que la concentración de estas moléculas es muy baja (Lysholm et al. 2012; Taboada et al. 2014; Zoll et al. 2015). Esta estrategia de amplificación también busca evitar sesgos al hacer amplificaciones blanco específico y por eso se utilizan los

oligonucleótidos hexámeros aleatorios (Lysholm et al. 2012). En la figura 2 se esquematiza la naturaleza de los oligonucleótidos.

Figura 2.- Representación esquemática de los oligonucleótidos hexámeros aleatorios y hexámeros aleatorios M13.



A la izquierda se muestra una representación de los oligonucleótidos del ensayo comercial y a la derecha los oligonucleótidos con la secuencia de M13 asociada.

A partir de una transcriptasa reversa recombinante expresada en *Escherichia coli* (incluida en el ensayo comercial), se generaron secuencias al azar. En la tabla A1 de anexos se muestra la mezcla de reacción tanto para hexámeros aleatorios como para hexámeros aleatorios unidos a M13.

Se utilizó un termociclador Veriti (Thermo Fisher Scientific 2015a) para la incubación a 65°C durante 10 minutos (para cada una de estas mezclas de reacción por separado), este paso permite la alineación de los oligonucleótidos. Posteriormente se detuvo la reacción en

hielo y se le agregó a cada una de las mezclas anteriores, la mezcla 2 de reacción para realizar la retro transcripción, como se muestra en la tabla A2 de anexos (La Roche 2015b).

Para llevar a cabo la retrotranscripción se utilizó un termociclador Veriti (Thermo Fisher Scientific 2015a) de acuerdo a las siguientes condiciones: ya que es una longitud menor a 4kb se utilizó una temperatura de 25°C durante 10 minutos seguido de 55°C durante 30 minutos, finalmente se inactivó la enzima a 85°C durante 5 minutos y enfriamiento a 4°C (La Roche 2015b).

Una vez que se terminaron los pasos de la termociclación, se formó cDNA a partir de las moléculas de RNA presentes en los virus que poseen este material genético. Estos cDNAs son formados al azar ya que los oligonucleótidos utilizados son secuencias de seis nucleótidos que se alinean azarosamente en los pasos de alineamiento de la termociclación (ver figura 2).

6.3.3.2.- PCR para síntesis de dsDNA

Se construyó la doble cadena de DNA (dsDNA) a partir de las moléculas de cDNA formadas en la sección 6.3.3.1. Empleando nuevamente oligonucleótidos hexámeros aleatorios M13 y fragmento klenow (Epicentre 2015). En el caso de los virus de DNA, se tomó parte sobrante del paso 6.3.3 y también se sintetizó una dsDNA con hexámeros aleatorio M13.

Con este paso se logró sintetizar dsDNA a partir de cDNA y DNA, de virus de RNA y DNA respectivamente. El uso del fragmento de klenow aseguró el marcaje de los ácidos nucleicos con el oligonucleótido aleatorio de la secuencia de M13. A pesar de haberla añadido en la sección 6.3.3.1 para las moléculas de RNA, se volvió a procesar con esta

estrategia de PCR con la finalidad de marcar aquellas moléculas que no fueron susceptibles al procesamiento en dicha sección. Posterior a este paso de amplificación nos aseguramos de tener secuencia M13 en todas las moléculas tanto las que provienen de RNA y de DNA. En la tabla A3 de anexos se muestra la mezcla de reacción para llevar a cabo la síntesis de dsDNA.

Posteriormente se utilizó un termociclador Veriti (Thermo Fisher Scientific 2015a) para la incubación a 94°C durante 3 minutos, pasando a continuación a hielo durante 2 minutos. Durante este tiempo se agregó 0.25 µL de polimerasa klenow en cada tubo de reacción (10U/µL). Nuevamente utilizando el termociclador se incubó a 37°C durante una hora, seguido de 94°C durante 3 minutos, terminado este tiempo se pasó a hielo durante 2 minutos. Y se vuelve a repetir el agregarle 0.25 µL de polimerasa klenow en cada tubo de reacción (10U/µL). Otra vez utilizando el termociclador se incubó a 37°C durante una hora, seguido de 75°C durante 10 minutos (Epicentre 2015).

Una vez que terminó la amplificación con fragmento klenow, las moléculas de dsDNA resultantes tienen el marcaje de la secuencia de M13. Es muy importante resaltar que la incorporación de la secuencia M13 se realizó con la finalidad de aumentar la cantidad de los ácidos nucleicos de interés y esto no debe interferir con la posterior identificación de los agentes etiológicos de interés, debido a que las secuencias de M13 son distintas a los virus y bacterias que se puedan encontrar presentes en las muestras.

6.3.3.3.- PCR para amplificación con oligonucleótidos M13

Se utilizaron los productos de dsDNA con las secuencias M13 añadidas por el fragmento de klenow y se amplificaron por medio de PCR empleando Platinum Taq (Thermo Fisher

Scientific 2015d) y los oligonucleótidos M13F y M13R. En la tabla A4 de anexos se muestra la mezcla de reacción para cada una de las muestras.

Después se utilizó el termociclador Veriti (Thermo Fisher Scientific 2015a). Se activó a la polimerasa a 95°C durante 2 minutos; posteriormente se dieron 30 ciclos como sigue: desnaturalización a 95°C durante 15 segundos, alineamiento a 50°C durante 30 segundos y extensión a 68°C durante 3 minutos; extensión final a 68°C durante 5 minutos.

Para cada uno de los pools, se obtuvieron 3 productos de PCR, dos proveniente de los virus de RNA y otro de los virus de DNA. Los productos de PCR provenientes de RNA se mezclaron entre si y en total fueron 8 productos que a continuación fueron cuantificados.

6.3.3.4.- Cuantificación de los Productos de PCR

Se realizó la medición espectrofotométricamente con el equipo NanoDrop ND1000 UV-Vis (Thermo Fisher Scientific 2015c). En el cual la muestra sin diluir se colocó directamente en el pedestal (2 µL) y se midió la concentración de ácidos nucleicos. El software instalado relaciona valores de absorbancia 260/280, que deben ser lo más aproximados a un valor de 2 para asegurarse de la pureza del DNA extraído y purificado.

Del resultado obtenido de la cuantificación se procedió a mezclar los productos de la amplificación con oligonucleótidos aleatorios M13 que provenían de RNA y de DNA de cada pool, en concentraciones equimolares y se evaluaron en el equipo Bioanalyzer para estimar el peso molecular.

6.3.3.5.- Estimación del tamaño de los fragmentos con equipo Bioanalyzer

Una vez obtenidas las mezclas de dsDNA de cada uno de los pools, se cuantificaron en el equipo 2100 Bioanalyzer Instruments (Agilent Technologies 2015). Esta es una plataforma basada en microfluidos, que permite el análisis electroforético con tan solo 1 μ L de muestra. Mediante la separación, medición del tamaño y cuantificación de bajas concentraciones de dsDNA se pueden evaluar fragmentos desde 50 hasta 7000 pares de bases, de concentraciones por debajo de 5pg/ μ L. Se evaluaron los 4 grupos de pacientes en el bioanalizador, los resultados obtenidos permitieron estimar el peso molecular y la concentración de cada uno (“2100 Agilent Bioanalyzer” 2015).

6.4.- Secuenciación masiva con plataforma Illumina

6.4.1.- Construcción de bibliotecas con oligonucleótidos 16S ribosomal

La subunidad 16S del RNA ribosómico de bacterias es un componente de la maquinaria transcripcional y está altamente conservado. Es una molécula que secuenciada nos permite conocer la microbiota humana, proporcionando la descripción de la mezcla de bacterias presentes. El gen consiste en regiones variables y conservadas, las primeras nos permiten la discriminación entre los microorganismos, mediante el estudio de la filogenia y taxonomía. Las regiones conservadas nos permiten amplificar el gen 16S utilizando oligonucleótidos universales (Cox, Cookson, y Moffatt 2013).

En el caso de los amplicones de la región hipervariable V1-V3 obtenidos en la sección 6.3.1.1 se procedió a realizar una reparación de los extremos al igual que para las bibliotecas de *novo* (Illumina 2015). Comenzando con la fosforilación de 5' y agregando

una adenina en 3'. La reparación anterior es necesaria para la adición de los adaptadores e índices (barcode) de la plataforma de Illumina. La ligación de los adaptadores y de los índices permite la identificación de las muestras y que sus extremos se puedan unir a la superficie de la flow cell y por lo tanto se puedan generar los cluster necesarios para la lectura en el equipo ("Nextera XT DNA Sample Preparation Guide.pdf" 2015).

Se diluyó a 0.2 ng/ μ L el producto de amplificación para la adición de los índices (barcode) de la plataforma de Illumina. Ya que se utilizaron amplicones no fue necesario hacer la tagmentación, pero sí se llevó a cabo la adición de los índices para que éstos puedan ser reconocidos por la plataforma de Illumina. Posteriormente se llevó a cabo la normalización de la biblioteca como indica el estuche de Nextera XT DNA ("Nextera XT DNA Sample Preparation Guide.pdf" 2015).

6.4.2.- Construcción de bibliotecas de *novo*

La mezcla hecha en la sección 6.3.3.4 se diluyó a 0.2 ng/ μ L de acuerdo a Nextera® XT DNA Sample Preparation Guide. A partir de esta solución de trabajo se realizó la fragmentación (tagmentación), utilizando transposasa Tn5, que es una enzima capaz de modular la longitud de los fragmentos de la biblioteca y capaz de aumentar la eficiencia a partir de picogramos de dsDNA, en presencia de cantidades elevadas de proteínas y de polímeros, que secuestran las moléculas de agua, aumentando el contacto entre la transposasa y el dsDNA. Este proceso generó fragmentos de dsDNA con extremos sobresalientes (Picelli et al. 2014; "Nextera XT DNA Sample Preparation Guide.pdf" 2015).

A continuación se llevó a cabo la normalización de la biblioteca como indica el estuche de Nextera XT DNA (Illumina, 2012). Esta última parte permite la separación, selección y purificación de los fragmentos de aproximadamente 500 pares de bases, mediante el uso de

las perlas AMPure XP. Estas perlas magnéticas tienen la capacidad de unirse a los productos de la amplificación, para separarlos y purificarlos. Según la cantidad de perlas utilizadas es el tamaño de los fragmentos que separan y su eficiencia de separación es superior a la de los métodos tradicionales por filtración («Agencourt AMPure XP.book - techdocs» 2015; «Nextera XT DNA Sample Preparation Guide.pdf» 2015).

Finalmente se realizó la secuenciación en el equip MiSeq como se describió en la sección 1.9.1.

6.5.- Análisis bioinformático

Los análisis bioinformáticos se realizaron de dos formas, en Internet con el software Galaxy-Metavir y en work station Dell precisión T5610 con procesador Intel ®, memoria RAM de 8GB, arquitectura de 64 bits y sistema operativo Ubuntu. En el sistema operativo Ubuntu se usó Biolinux y a su vez en éste se utilizaron tres estrategias de análisis con tres distintas herramientas bioinformaticas para el mapeo de las lecturas obtenidas por secuenciación. Es importante resaltar que estas estrategias son exclusivas para el análisis del viroma.

6.5.1.- Análisis con Galaxy y Metavir

Se utilizó primero Galaxy que es un sitio en Internet, público y gratuito. Con esta herramienta se pueden procesar los archivos de las lecturas obtenidas por secuenciación, con una interfaz gráfica y sin necesidad de saber lenguaje de shell o pearl que usualmente se utilizan para el procesamiento de estos datos de secuenciación masiva. Primero se utilizó Galaxy que se compone de la herraienta de mapeo Bowtie2, con esta se procesaron las

lecturas obtenidas por secuenciación de la sección 6.4.2 y se realizó el mapeo a secuencias genómicas de humano utilizando la configuración por defecto y con la base de datos de humano propia del sitio Galaxy. De las lecturas que no mapearon al genoma humano fueron pareadas y después de eliminar aquéllas carentes de par, se obtuvo un solo archivo que se necesitó modificar con las herramientas FASTX-Toolkit en el mismo sitio, para hacer compatible dicho archivo con la siguiente sitio en línea Metavir (“Galaxy” 2015).

Metavir es un servidor web dedicado al análisis de metagenomas virales. Metavir tiene varias herramientas para el procesamiento de los datos, la primera es GAAS (Genome relative Abundance and Average Size) con la cual se realizó la normalización de la abundancia de las especies virales presentes (Angly et al. 2009). Posteriormente Metavir realizó una filogenia con una base de datos de referencia propia y que sirve de base para hacer una comparación de los virus basada en la similitud. A continuación se utilizó otra herramienta que posee Metavir que es tBLASTx, que es un algoritmo que se basa en la búsqueda de similitudes de proteínas con una base de datos propia del sitio en línea. Finalmente Metavir utilizó curvas de rarefacción con las cuales evaluó la riqueza del viroma y el resultado lo reporta en gráficas que muestran la diversidad y proporción de los virus presentes en la muestra (Roux et al. 2011). Es importante mencionar que todo el procesamiento anterior lo hace de forma automática con la propia configuración por defecto del sitio.

6.5.2.- Análisis con Ubuntu Biolinux

De forma alternativa a Galaxy-Metavir, se utilizó Biolinux 8 en sistema operativo Ubuntu 14.2, realizando comandos mediante lenguaje de Shell. Con el uso de esta plataforma podemos trabajar sobre varios archivos simultáneamente o ejecutar varias tareas en orden,

lo cual lo convierte en una herramienta con gran poder de procesamiento. El objetivo de utilizar estas herramientas es poder hacer el análisis bioinformático de los datos obtenidos mediante secuenciación masiva, primero se necesitó realizar nuestras propias bases de datos de humano y de virus para tomarlas como referencia.

6.5.2.1.- Bases de datos

Para identificar las secuencias que corresponden al hospedero y a los agentes etiológicos, se descargaron secuencias de la base de datos del National Center for Biotechnology Information (NCBI). Para humano, se descargaron transcriptoma, DNA mitocondrial, RNA ribosomal y genómico, este genoma de referencia (GRCh38) fue construido con genomas de voluntarios y es una representación consenso que provee una buena aproximación de un solo individuo (“Reference Genome” 2015). Y para hacer la base de datos de los agentes etiológicos se descargaron secuencias de virus bajo los siguientes parámetros:

Viridae [ORGANISM] NOT “Homo sapiens” NOT “probes” NOT “primers” NOT “sequencing projects” con el propósito de tener una base de datos sin secuencias espurias asociadas a virus.

6.5.2.2.- Filtrado por calidad.

Los archivos fastq que se obtuvieron de la sección 6.4.2, contienen información de la calidad de las lecturas, por lo que se procedió a filtrar aquellas lecturas con una calidad menor de Q20, es decir, aquellas que no tienen el 99.9% de ser correctas en al menos 50%

de su secuencia. Para realizar el proceso de mapeo a genoma humano primero se convirtieron los archivos fastaq a archivos en formato fasta, que es un formato que sólo contiene nombre y secuencia nucleotídica utilizando FASTX-Toolkit (“FASTX-Toolkit” 2015).

Una vez obtenidas las bases de datos de humano y de virus y que los archivos de la secuenciación fueron filtrados por calidad, se realizaron tres estrategias de procesamiento bioinformático, utilizando las herramientas con las condiciones por default. Esto se realizó con el fin de comparar los resultados con diferentes herramientas bioinformáticas y obtener el más adecuado para el procesamiento de los datos obtenidos.

6.5.2.2.1- Proceso bioinformático A

Este proceso se realizó para obtener las secuencias virales de los pools de muestras directamente de los archivos filtrados por calidad Q20 y en formato fasta de acuerdo a la sección 6.5.2.2. Con estos archivos se realizó el mapeo por medio de la herramienta BLAST (Basic Local Aligmen Search Tool). Este es un programa basado en un algoritmo heurístico capaz de comparar secuencias de nucleótidos de manera rápida y con lo cual podemos identificar con buena certidumbre el mejor de los hits, de cada secuencia de las lecturas del archivo a analizar con la base de datos de virus como referencia (Lobo 2015). Esta herramienta se utilizó con el parámetro de valor esperado ó *e-value* modificado de su valor por defecto de 10 a un valor de 1×10^{-30} , este valor fue probado junto con otros, pero con este valor se obtuvieron los mejores resultados (datos no mostrados), este valor modificado nos permite realizar una exploración mas profunda en las lecturas que se obtuvieron por secuenciación. Una vez obtenido el archivo analizado con BLAST, se realizó otro análisis con ayuda de la herramienta GAAS, la cual realizó la normalización de

la abundancia de las especies virales presentes en el archivo (Angly et al. 2009). En GAAS modificamos el porcentaje de la similitud con la cual analiza a los nucleótidos de las lecturas a evaluar con respecto a una base de datos, que por defecto tiene 80%. Nosotros le asignamos un valor de 90% con el propósito de hacer mas especifica la búsqueda y clasificación de las lecturas evaluadas y en consecuencia tener una mejor normalización. Posteriormente se realizó una verificación manual de las secuencias para corroborar su cantidad e identidad, utilizando el archivo terminación tblastx procesado por GAAS y que tienen las identificaciones de las secuencias dadas originalmente por BLAST. Se copió el archivo en una hoja de cálculo de Microsoft Office Excel 2007 y se filtraron las secuencias con identificaciones similares, posteriormente se utilizó el nombre de la lectura para localizar su secuencia en el archivo filtrado Q20 obtenido en la sección 6.5.2.2. Una vez localizada la secuencia se utilizó la herramienta BLAST del sitio NCBI para corroborar la identidad de las mismas, se realizó así la verificación para todos los archivos y para todas las secuencias proporcionadas por GAAS.

6.5.2.2.2.- Proceso bioinformático B

Este proceso se realizó para eliminar las secuencias humanas utilizando el mapeador BWA (Burrows-Wheeler Aligner) y posteriormente obtener las secuencias virales. El archivo fasta que se obtuvo en la sección 6.4.2 se mapeo contra genoma humano utilizando BWA que es un mapeador diseñado para tener un buen balance entre rendimiento y precisión. Su algoritmo se basa en el mapeo de secuencias con baja divergencia tomando como base un genoma de referencia, para el proceso B se utilizó la base de datos de genoma humano para hacer el mapeo (“Burrows-Wheeler Aligner” 2015). Este mapeador lo utilizamos con los parámetros por defecto y nos permitió la identificación de las lecturas del hospedero presentes en la muestra, resultando en un archivo sam (Sequence Alignment/Map) que es un formato genérico para almacenar alineamientos de las secuencias de nucleótidos y nos dice el nombre de la referencia contra la cual alineó, así como la

posición absoluta del genoma, la calidad del mapeo, entre otras características (“SAMtools” 2015). Utilizando samtools se obtuvo a las lecturas que no mapearon a humano, a partir de este archivo sam se generó un archivo binario bam (“samtools(1) manual page” 2015). Este archivo binario se convirtió a formato fasta y con este último se realizó el proceso de mapeo a secuencias virales mediante la herramienta BLAST con los parámetros modificados como se explicó en el proceso bioinformático A y se generó un archivo txt. Con el propósito de conocer la diversidad de virus en las muestras se realizó un análisis con GAAS con los parámetros modificados y su verificación manual como se explicó en el proceso bioinformático A.

6.5.2.2.3.- Proceso bioinformático C

De igual forma al proceso bioinformático B, se eliminaron las secuencias humanas ahora empleando el mapeador SMALT y posteriormente se obtuvieron las secuencias virales. El archivo fasta obtenido de la sección 6.5.2.2 se utilizó para hacer el mapeo contra genoma humano utilizando SMALT, que es un algoritmo que alinea secuencias cortas contra un genoma de referencia, sus parámetros pueden ser modificados de acuerdo a la sensibilidad y a la velocidad de procesamiento que se deseen. Su algoritmo se basa en la construcción de un índice de palabras cortas (k-mero), los cuales hacen un muestreo en incrementos equidistantes a lo largo de la secuencia de referencia y aquellas lecturas identificadas son alineadas en base al algoritmo de Smith-Waterman. Se basa en la comparación de varias longitudes de segmentos en lugar de una secuencia completa, siendo un algoritmo muy parecido al de BLAST, pero es más sensible por que busca un mayor campo de posibilidades (Sanger Institute 2015b). El parámetro modificado en esta herramienta fue el umbral de identidad, le asignamos un valor del 90% lo cual resulta similar al parámetro modificado en GAAS. Como resultado del análisis con este mapeador se obtuvo un archivo formato bam únicamente con las lecturas no mapeadas a humano. Posteriormente se utilizó FASTX-Toolkit para cambiar este archivo bam a formato fasta, con este archivo fasta se

realizó el mapeo con BLAST utilizando la base de datos de virus y por último se procedió a hacer el análisis con GAAS y su verificación manual con los parámetros modificados como se explico en el proceso bioinformático A.

6.5.2.2.3.1.- Ensamble de lecturas

El ensamblaje de lecturas metagenomas de *novo*, se hace en el caso de que no se tenga un genoma de referencia. Se obtuvo mediante regiones de traslape que construyen nodos, como se describen en las graficas de Bruijn, usadas por el algoritmo de Velvet y Metavelvet. Primero con Velvet se construyó un arreglo de k-meros, en donde se tienen que probar varios tamaños para encontrar el optimo, para los datos obtenidos se utilizaron k-meros de 63, 189, 195, 201, 207, 213 y a partir de estos hizo la construcción de gráficos agrupándolas en nodos. Con los gráficos construidos en Metavelvet se buscó la solución a los diferentes nodos construidos, buscando los nodos quiméricos, los separó y los descompuso para realizar subgráficas. Con estas dos graficas construyó el supercontig o unión de secuencias contiguas. Finalmente se visualizó el resultado en el programa tablet (Zerbino y Birney 2008; Namiki et al. 2012).

6.5.3.- Análisis de 16S ribosomal

Para realizar el análisis de las lecturas obtenidas para la región hipervariable V1-V3 de la sección 6.4.1 se utilizó la aplicación 16S Metagenomics versión 1.0.0.0 de BaseSpace. Esta aplicación realizó la clasificación taxonómica utilizando una versión curada de Green Genes y el algoritmo que utilizó es una implementación de alto rendimiento del proyecto de base de datos de ribosoma (RDP) (“- BaseSpace Application -” 2015).

VII. RESULTADOS

7.1.- Número de pacientes y muestras analizadas

Se logró reclutar a 890 pacientes, pero solamente 655 entraron al protocolo de investigación (código B26-13), al total de los individuos se les realizó la identificación de virus respiratorios y sólo 116 muestras fueron elegidas para ser procesadas para las estrategias de secuenciación.

Es importante resaltar que de las 116 muestras, sólo 36 muestras fueron procesadas en fresco de acuerdo a la sección 6.3.3 y a la par se guardó la alícuota para detección de virus y para la fracción bacteriana de acuerdo a la sección 6.3.1 y 6.3.2. Las otras 80 muestras se tomaron del banco de muestras congeladas a -80°C y se siguieron las mismas condiciones de proceso de la sección 6.3.3. La prevalencia de virus obtenida por qPCR de Biomark se puede observar en la tabla 5 y la clasificación de las 116 muestras para secuenciación en la tabla 6.

7.2.- Detección de virus con equipo BioMark.

Con la alícuota de la sección 6.3.2 en el equipo BioMark se obtuvieron los resultados de la detección de la seis principales familias de virus, en la Tabla 5 se muestran que agentes etiológicos prevalecieron en la temporada invernal 2013-2014.

Tabla 5.- Detección de virus en el total de las muestras.

Adultos/agente etiológico	N^a	%^b	Niños/agente etiológico	N^a	%^b
Influenza A y B	81	16.23	Influenza A y B	16	10.25
Enterovirus/Rinovirus	35	7.01	Enterovirus/Rinovirus	36	23.07
Virus sincial respiratorio A y B	12	2.40	Virus sincial respiratorio A y B	15	9.61
Coronavirus 229E	9	1.80	Coronavirus 229E	3	1.92
Parainfluenza 1,2,3 y 4	8	1.60	Parainfluenza 1,2,3 y 4	4	2.56
Bocavirus	2	0.40	Bocavirus	2	1.28
Metapneumovirus	1	0.20	Metapneumovirus	2	1.28
Adenovirus	0	0	Adenovirus	4	2.56
Negativo	352	70.36	Negativo	75	47.47
Total	499	100	Total	156	100

a) N: Número de virus detectados

b) Proporción de virus dada en porcentaje

De los resultados obtenidos en la tabla 5, podemos observar la detección de virus respiratorios tanto para adulto y para pacientes pediátricos de la temporada invernal 2013 a 2014. Aquellos virus que se encontraron en mayor porcentaje para la población de adultos fueron influenza, seguidos de enterovirus/rinovirus y virus sincial respiratorio. En contraste los virus encontrados en mayor porcentaje para los pacientes pediátricos fueron primero enterovirus/rinovirus, seguidos de influenza y virus sincial respiratorio.

7.3.- Clasificación de muestras

A partir de los resultados obtenidos en 6.3.2 se clasificaron las muestras en los siguientes grupos de pacientes: adultos con virus detectados (pool 1), grupo de adultos sin detección de virus (pool 2). Para el caso de los pacientes pediátricos se formó un grupo con aquellos en los que hubo detección de virus (pool 3) y otro grupo de niños en los que no se detectaron virus (pool 4), la cantidad de muestras que conformaron cada pool, así como los virus que se detectaron en cada grupo se muestra en la tabla 6. En una primera ronda de secuenciación se procesaron los pool 1, 2, 3 y 4, de las 36 muestras procesadas en fresco.

Para la segunda ronda de secuenciación se repitieron los pool 1, 2 y 4, debido a que al analizar los pool 1, 2 y 4 de la primer secuenciación, los resultados no fueron adecuados para obtener datos, es importante recordar las 80 muestras que conformaron a estos grupos que se repitieron provinieron del banco de muestras congeladas a -80°C.

Tabla 6. Prevalencia de virus respiratorios en las muestras seleccionadas para su agrupamiento y secuenciación

Sector de la población	Con agente etiológico viral	1er secuenciación		2 da secuenciación	
		36 ^a	Virus encontrados	80 ^b	Virus encontrados
Adultos	Si	8	5 Inf A H1N1 ^c 2 Inf A H3N2 ^d 1 E/R ^e	25	17 Inf A H1N1 ^c 1 Inf A H3N2 ^d 4 Inf no sub ^h 2 Coro 229E ⁱ 1 E/R ^e
	No	10	NA ^j	25	NA ^j
Niños	Si	9	6 E/R ^e 1 Inf A H1N1 ^c 1 Meta ^f 1 Adeno ^g	NA ^j	NR ^k
	No	9	NA ^j	30	NA ^j

- a) 36: Total de muestras utilizadas en una primera secuenciación y procesadas en fresco
- b) 80: Total de muestras utilizadas en una segunda secuenciación y procesadas de muestras congeladas
- c) Inf A H1N1: Influenza A H1N1 (2009)
- d) Inf A H3N2: Influenza A H3N2
- e) E/R: Enterovirus/Rinovirus
- f) Meta: Metapneumovirus
- g) Adeno: Adenovirus
- h) Inf no sub: Influenza A subtipo no determinado
- i) Coro 229E: Coronavirus 229E
- j) NA: No aplica
- k) NR: No realizado

7.4.- Cuantificación de los Productos de PCR

Se procesaron tanto RNA y de DNA de cada uno de los grupos, como se describió en 6.3.3.1 y 6.3.3.2 respectivamente, finalmente se les realizó PCR de acuerdo a la sección 6.3.3.3. Posteriormente se les midió su concentración (sección 6.3.3.4) y los resultados se muestran en la tabla 7.

Tabla 7.- Concentración de los productos de PCR de cada uno de los pools leídos con Nanodrop

1er secuenciación				2 da secuenciación			
Muestra Molde de RNA	[] ^a	Muestra Molde de DNA	[] ^a	Muestra Molde de RNA	[] ^a	Muestra Molde de DNA	[] ^a
pool 1	10.2	pool 1	14.8	pool 1	8.95	pool 1	8.8
pool 3	10.3	pool 2	6.9	pool 3	14.52	pool 3	21.6
pool 4	11.6	pool 4	8.26	No aplica			
pool 5	23.6	pool 5	13.3	pool 5	10.4	pool 5	18.1

a) []: Concentración en ng/μL

Con estos resultados se procedió a realizar el mezclado de las muestras de forma equimolar, de cada fracción proveniente de RNA y DNA para cada uno de los 4 grupo estudiados, posteriormente se realizó la preparación de las bibliotecas y finalmente su secuenciación para conocer el viroma.

Por otra parte se procesó la alícuota para bacterias donde se obtuvieron amplicones de acuerdo a la sección 6.3.1.1 y a continuación se procedió a hacer la cuantificación de los ácidos nucleicos según la sección 6.3.1.2, los resultados se muestran en la tabla 8.

Tabla 8.- Concentración de los productos para 16S de cada uno de los pools con Qubit

Muestra	[] ^a
pool 1	1.4
pool 3	1.15
pool 4	0.414
pool 5	1.45

a) []: Concentración en ng/μL

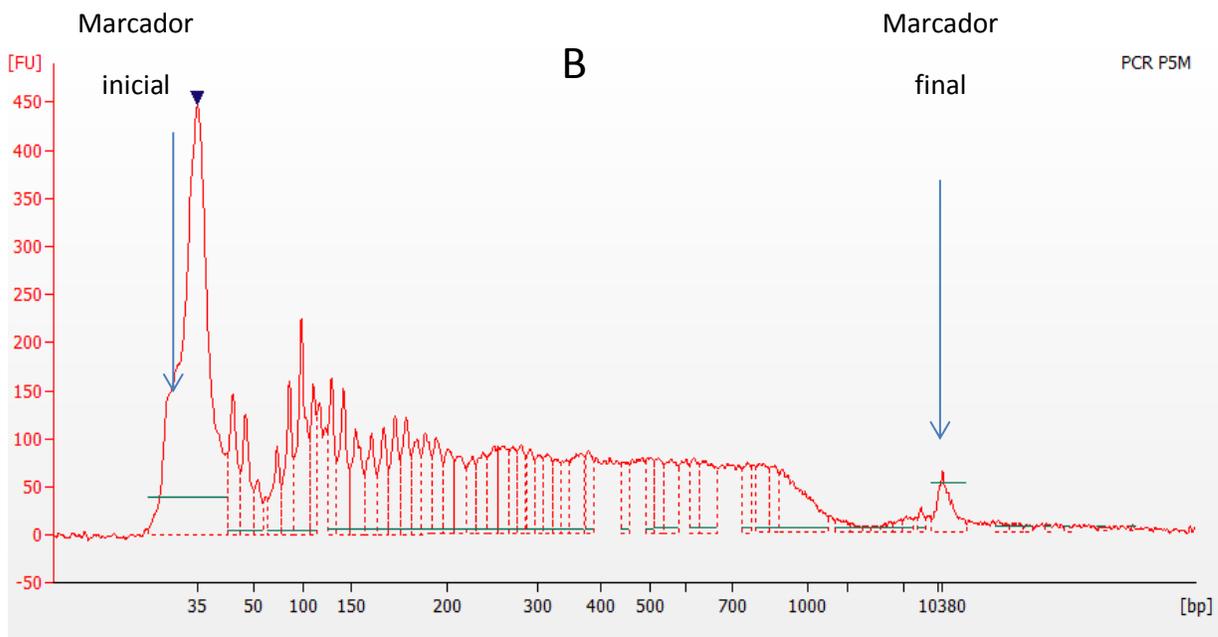
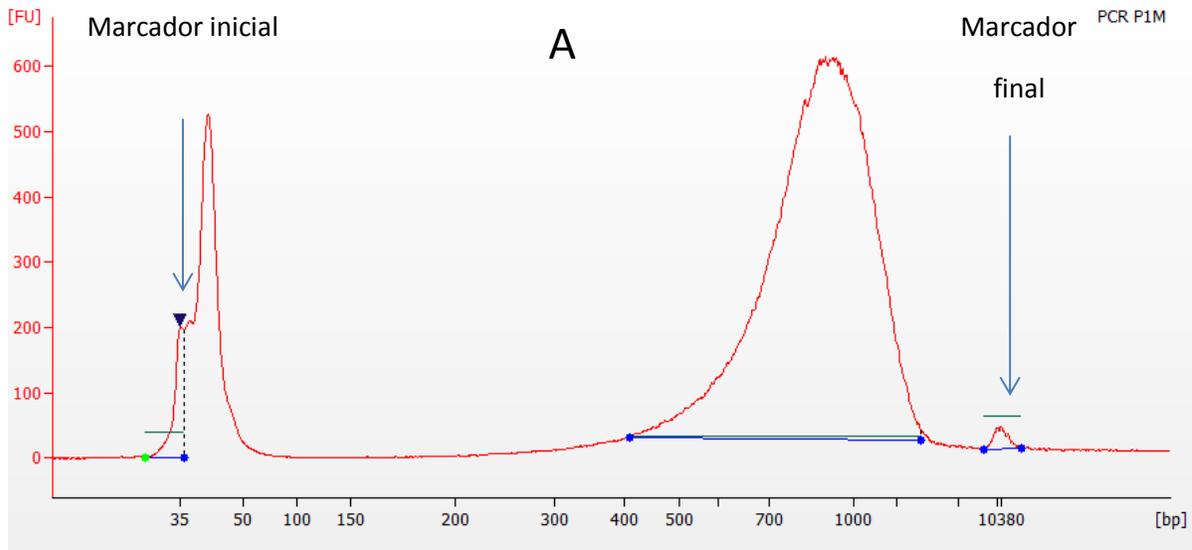
A los productos de RNA y de DNA mezclados se les realizó otra medición con el equipo Bioanalyzer, mientras que para los productos de amplificación para región hipervariable V1-V3 se llevaron a una concentración de 0.2 ng/μL para la adición de los índices y barcode de la plataforma de Illumina.

7.5.- Estimación de la concentración y del tamaño de los productos en equipo Bioanalyzer

Cada uno de los productos obtenidos de la sección 6.3.3.4 para cada uno de los cuatro grupos evaluados se les estimó el tamaño en pares de bases con el equipo Bioanalyzer. En la grafica 1 A se muestra un ejemplo de un resultado adecuado para realizar la estimación de la concentración y el tamaño de los fragmentos, en este caso fueron desde 400 hasta 2000 y una concentración de 12,954.74 pg/μL que correspondió al grupo de adultos con presencia de virus, para la mayoría de los demás grupos se obtuvo la estimación correspondiente y se muestran en la tabla 9.

En el grafica 1 inciso B se muestra el ejemplo del único grupo del cual no se obtuvo un resultado adecuado para estimar la concentración y el tamaño de los fragmentos, este fue el caso del grupo de niños sin identificación de virus, donde se observa que hay una degradación del material genético.

Gráfica 1.- Productos obtenidos del grupo de adultos con virus y niños sin virus



Para el caso del grupo de niños sin virus, el equipo no detecta una concentración ni puede estimar la cantidad de pares de bases ya que como se puede observar en la gráfica 1B en la muestra hay una degradación del material genético. A la derecha de la gráfica se observa la imagen que donde no hay una banda (no hay color negro).

En la tabla 9 se muestran los resultados obtenidos de la estimación de la concentración y del tamaño de los fragmentos en cada uno de los grupos evaluados, estos resultados se muestran en dos rondas de secuenciación, debido a que los grupos en los cuales no se encontraron resultados se volvieron a repetir.

Tabla 9.- Estimación de la concentración y del tamaño de los productos para secuenciación de *novo*.

Grupo	Primera secuenciación		Segunda secuenciación	
	[] ^a	p.b. ^b	[] ^a	p.b. ^b
1 ^e	12 954.7	400 - 2000	477.9	1000 – 2800
2 ^e	2 132.7	350 - 2000	6012.4	700-2040
3 ^e	9 520.0	700 - 2000	NR ^d	NR ^d
4 ^e	SR ^c	SR ^c	521.4	1000-2900

- a) [] : concentración en pg/μL
- b) p.b. : Pares de bases
- c) SR : Sin resultado
- d) NR : No repetido
- e) Del 1 al 4 corresponden a cada uno de los grupos evaluados. 1= grupo de adultos con presencia de virus. 2= grupo de adultos sin presencia de virus. 3= grupo de niños con presencia de virus y 4 grupo de niños sin presencia de virus.

Al analizar las concentraciones obtenidas mediante las estrategias de RT-PCR y PCR, para cada una de las muestras y analizadas con el equipo Nanodrop y Bioanalyzer, tenemos que las concentraciones fueron adecuadas para realizar las diluciones correspondientes. Los productos obtenidos se llevaron a una concentración de 0.2 ng/ μ L y el resultado del tamaño en pares de bases también sirvió para saber que teníamos fragmentos de tamaños adecuados para realizar la fragmentación y preparación de las bibliotecas a secuenciar, con excepción del pool 4 de la primer ronda de secuenciación donde se pudo observar una degradación de la muestra (ver gráfica 1 inciso B) y a pesar de que se secuencio este grupo, los resultados obtenidos no fueron adecuados para su análisis.

En contraste, para el caso del pool 1 del cual se observó una buena concentración y tamaño de los fragmentos (gráfica 1 inciso A) en comparación con los resultados obtenidos en la segunda ronda de secuenciación para el mismo grupo, para el caso del grupo 1 de la primera ronda de secuenciación no se contaba con el complemento de la información de 16S ribosomal y por eso se analizaron los resultados del grupo 1 de la segunda ronda de secuenciación.

7.6.- Resultados de secuenciación en equipo Illumina.

Una vez que se realizaron todos los procesamientos necesarios de cada una de las muestras de cada uno de los grupos evaluados en el presente trabajo, se procedió a construir las bibliotecas de las regiones hipervariables y de *novo* de acuerdo a las secciones 6.4.1 y 6.4.2 respectivamente. En la tabla 10 se muestran los resultados de la cantidad de lecturas obtenidas mediante secuenciación de la primera y segunda secuenciación.

Tabla 10.- Resultados obtenidos de las rondas de secuenciación

1era Secuenciación		2da Secuenciación	
Lecturas leídas	14 millones	Lecturas leídas	20 millones
Clusters que pasaron el filtro de calidad	81.6 %	Clusters que pasaron el filtro de calidad	62.2 %
Lecturas proporcionadas	13 millones	Lecturas proporcionadas	12 millones

Tanto para la primera y segunda secuenciación se obtuvieron dos archivos por cada grupo de pacientes, recordando que se realizaron dos ciclos de 250 (pair end). El primer archivo identificado como R1 (lectura en dirección 5' a 3') y R2 el segundo archivo (lectura 3' a 5').

7.7.- Comparación de análisis bioinformáticos para viroma

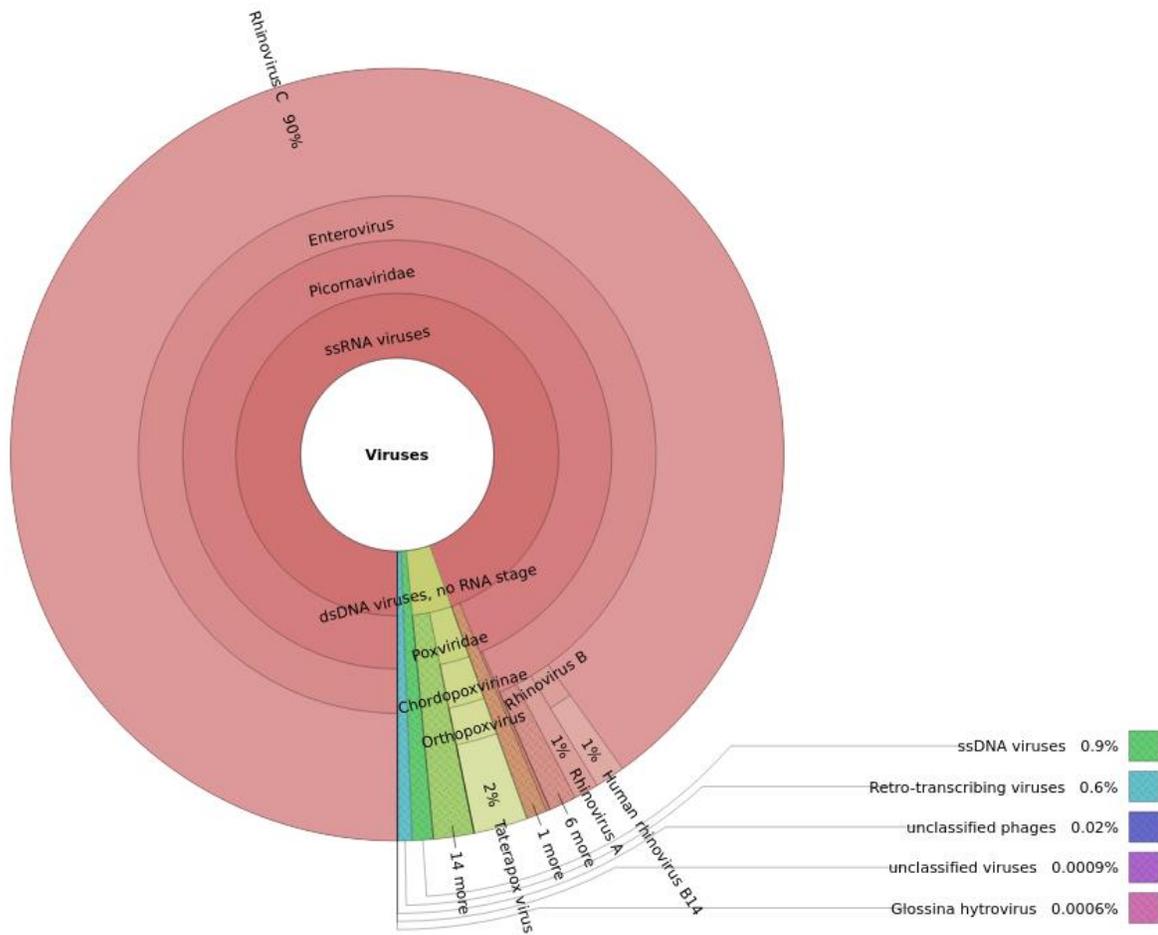
7.7.1.- Análisis con Galaxy y Metavir

El primer análisis hecho en línea mediante la interfaz Galaxy y con los datos obtenidos del grupo de niños con virus presentes, nos proporcionó los siguientes resultados:

De los dos archivos (R1 y R2) obtenidos mediante la secuenciación cada uno tenía un total de 2 216 661 lecturas. Para cada archivo se realizó el mapeado a genoma humano utilizando la herramienta BWA, de las lecturas que no mapearon de cada archivo se parearon, se descartaron las cuales no tuvieron su par y entre los dos archivos se obtuvo finalmente uno con 174 756 lecturas. Éste archivo se utilizó en Metavir para realizar el procesamiento con las herramientas GAAS, tBLASTx y con la ayuda de las curvas de rarefacción se obtuvo la diversidad y riqueza de virus presentes en el grupo de muestras que conformaron al grupo de niños con virus.

Se obtuvo 157 117 lecturas que no corresponden a virus y 17 639 lecturas de virus con las cuales se contruyo la gráfica 2, donde se muestran a los virus presentes y la proporción de éstos.

Grafica 2.- Diversidad viral del grupo de niños con virus.



En la gráfica 2 se puede observar que el virus que se encuentra en mayor proporción es rinovirus C con 90%, 2% de Tetrapox virus, 1% de rinovirus 14, 1% de rinovirus A y menos del 1% corresponden a los virus Torque teno mini virus, influenza A y adenovirus B, el resto corresponde a fagos diversos.

El resultado fue obtenido dos semanas después de haber ingresado las secuencias en el sitio en línea y la gran limitante para utilizar esta herramienta para el procesamiento de los datos es el tamaño de los archivos y la velocidad de transferencia de los datos, dado por el servicio de Internet y la espera para obtener un análisis, esto último depende de la cantidad de usuarios utilizando las plataformas de análisis en la web, así como la cantidad de datos que analiza por cada usuario y el potencial mantenimiento que se le dan a los servidores que prestan el servicio, esto podría aumentar aun más el tiempo de espera.

7.7.2.- Procesos en Biolinux

Primero se obtuvieron las bases de datos de humano y de virus. La referencia de genoma humano (GRCh38), contiene 4 322 secuencias. Para el índice de virus construido sin secuencias espurias son 1 763 688 secuencias. Para realizar las pruebas con los diferentes procesos bioinformáticos se utilizó el archivo R2 filtrado por calidad Q20 con 2 173 502 lecturas del grupo de niños con virus, para comparar los procesos.

7.7.3.- Comparación de los proceso bioinformático A, B y C

Para el proceso A se utilizó la herramienta BLAST para el procesamiento de las lecturas, el procesamiento se llevo a cabo en un lapso de tiempo de 12 horas y 54 minutos. De la cantidad original de 2 173 502 lecturas, el algoritmo utilizó 153 133 para realizar su

alineamiento local utilizando la base de datos de virus como índice. A continuación se procedió a usar la herramienta GAAS y nos dio un total de 1 340 lecturas para virus, de forma manual se corroboró y se obtuvo la proporción de virus como se observa en la gráfica 8, inciso A.

Para el proceso B se utilizó la herramienta bioinformática BWA para realizar el mapeado a genoma humano y partir de las 2 173 502 secuencias filtradas por Q20, se mapearon 2 185 800 lecturas. De las 21 558 lecturas que no mapearon a genoma humano se realizó el alineamiento local con la base de datos de virus y usando la herramienta BLAST, finalmente se utilizó la herramienta GAAS para obtener la diversidad de virus y de forma manual se corroboró la identidad de las lecturas. Este proceso es el que requirió de 1 hora y 53 minutos y se obtuvieron 6 430 secuencias, de las cuales se muestra su proporción en la gráfica 8, inciso B.

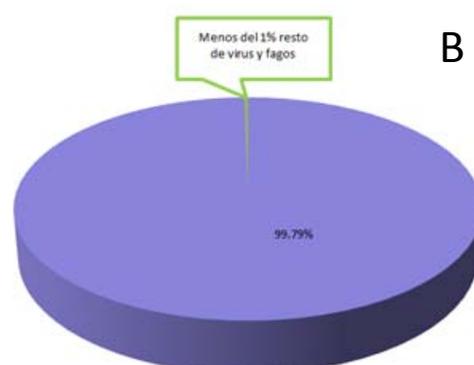
Se necesitó 8 horas y 45 minutos para realizar el proceso C, en éste proceso se utilizó la herramienta SMALT para realizar el mapeado a genoma humano y a partir de las 2 173 502 secuencias filtradas por calidad, mapeo 2 029 737 lecturas a humano. Las 177 621 que no mapearon se les aplicó el alineamiento local con la herramienta BLAST y tomando como referencia el índice de virus, posteriormente se usó el algoritmo GAAS y se realizó su revisión de forma manual de 6 478 lecturas, con estas lecturas se obtuvo las proporciones que se observan en la gráfica 8, inciso C.

Se puede observar en la gráfica 8 inciso A una proporción de 85.16% de Rinovirus C, 14.35% de la presencia de Retrovirus Endógenos y menos del 1% el resto de virus y fagos. En el inciso B un 99.79% de las lecturas corresponden a Rinovirus C y el resto menos del 1% a los demás virus y fagos. Finalmente en el inciso C hay un 98.92% de las lecturas pertenecen a Rinovirus C, el resto menos del 1% corresponden a los demás virus y fagos.

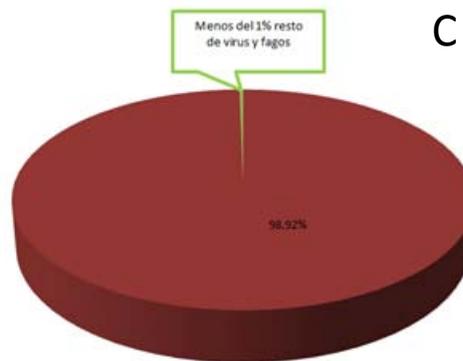
Gráfica 3.- Proporción de virus con proceso bioinformático A, B y C del archivo R2 del grupo de niños con presencia de virus.



Virus	Proporción (%)
Rinovirus C Mex14	85.16
Retrovirus endógenos	14.35
Influenza A H1N1	0.28
Torque Teno minivirus	0.01
Adenovirus B	0.04
Fago de Propionibacterium	0.08
Fago de Streptococcus	0.02
Fago de Staphylococcus	0.05
Fago de Escherichia	0.01



Virus	Proporción (%)
Rinovirus C Mex14	99.79
Influenza A H1N1	0.04
Torque Teno minivirus	0.01
Adenovirus B	0.04
Fago de Propionibacterium	0.06
Fago de Streptococcus	0.01
Fago de Staphylococcus	0.04
Fago de Escherichia	0.01



Virus	Proporción (%)
Rinovirus C Mex14	98.92
Retrovirus endógenos	0.66
Influenza A H1N1	0.06
Torque Teno minivirus	0.02
Adenovirus B	0.04
Fago de Propionibacterium	0.09
Fago de Streptococcus	0.04
Fago de Staphylococcus	0.06
Fago de Enterobacteria	0.01
Fago de Escherichia	0.01

7.7.4.- Elección del procesamiento a utilizar.

Al comparar los resultados obtenidos encontramos con la herramienta BLAST un 85.16% lecturas de rinovirus C contra un 99.79% que arrojo la herramienta BWA y un 98.92% con la herramienta SMALT, también encontramos en 14.35% la presencia de retrovirus endógenos al analizar los resultados de BLAST no se observó la presencia con BWA y sólo 0.66% de SMALT.

Al analizar el tiempo y la cantidad de secuencias utilizadas, tenemos que con BLAST se procesó 153 133 lecturas y requirio 12 horas y 54 minutos, seguido de SMALT con el que se procesó 177 621 lecturas en 8 horas y 45 minutos y con BWA se procesó 21 558 lecturas en un tiempo de 1 hora y 53 minutos. Por lo tanto aunque el mapeo con BWA es el que requirió de menor tiempo, también dejamos de analizar una cantidad importante de lecturas y podríamos dejar a un lado secuencias virales de importancia.

Al ser el mapeador SMALT el más versátil para realizar los procesamientos, al no requerir tantos recursos computacionales y al poder modificar los parámetros de búsqueda como el tamaño de k-meros y el porcentaje de identidad, lo hacen el más sensible de los tres procesos. Muestra de la sensibilidad de esta herranmienta es que pudo procesar lecturas del fago de Enterobacteria, mientras que los otros dos procesos no se observan estas lecturas.

Por lo tanto escogemos al proceso bioinformático C para realizar el procesamiento del total de los archivos, recordando que por cada pool se tiene un par de archivos asignados como R1 y R2, de estos se evalúan las lecturas de cada archivo de forma manual y debido a que son resultados complementarios estos se muestran en una sola gráfica y como se analizaron cuatro grupos son cuatro graficas, una para cada grupo. El total de lecturas obtenidas para el total de las muestras se muestra en la tabla 11.

Adicionalmente se muestran aquellos datos obtenidos mediante el análisis de la región hipervariable V1-V3, hecho con la aplicación Metagenomics en Base Space. Este análisis se realizó en los cuatro grupos estudiados y se muestran a la par con los resultados de virus. En la tabla 12 se muestra al total de lecturas obtenidas, para cada uno de los grupos a los que se les realizó este análisis.

Tabla 11.- Lecturas obtenidas mediante el proceso de análisis bioinformático C.

Lecturas								
Grupo	Al inicio	Filtradas Q20	Mapeadas a humano	% ^a	No mapeadas a humano	% ^a	Virus y fagos	% ^a
Adultos con virus	415 116	R1 = 395 048	R1 = 234 524	56	R1 = 160 524	38	R1 = 44	0.01
		R2 = 376 460	R2 = 226 373	54	R2 = 150 087	36	R2 = 31	0.007
Adultos sin virus	2 264 145	R1 = 2 222 580	R1 = 2 013 922	88	R1 = 208 658	9	R1 = 1 032	0.04
		R2 = 2 201 779	R2 = 1 995 519	88	R2 = 206 260	9	R2 = 850	0.03
Niños con virus	2 216 661	R1 = 2 207 358	R1 = 2 047 116	90	R1 = 160 242	7	R1 = 6 527	0.29
		R2 = 2 173 492	R2 = 1 995 871	90	R2 = 177 621	8	R2 = 6 478	0.29
Niños sin virus	895 826	R1 = 850 096	R1 = 567 821	63	R1 = 282 275	31	R1 = 121	0.01
		R2 = 820 493	R2 = 556 436	62	R2 = 264 057	29	R2 = 392	0.04

a) % = Proporción en porcentaje

Tabla 12.- Lecturas obtenidas mediante el proceso de análisis de la aplicación 16S Metagenomics de Base Space.

Lecturas						
Grupo	Al inicio	Filtradas Q20	No clasificadas	% ^a	Calsificadas por especie	% ^a
Adultos con virus	36 675	28 710	1 957	5	26 753	72
Adultos sin virus	41 985	33 433	1 845	4	31 588	75
Niños con virus	1 606 643	1490 590	74 674	5	1 415 916	88
Niños sin virus	103 466	77 763	3 360	3	72 659	71

a) % = Proporción en porcentaje

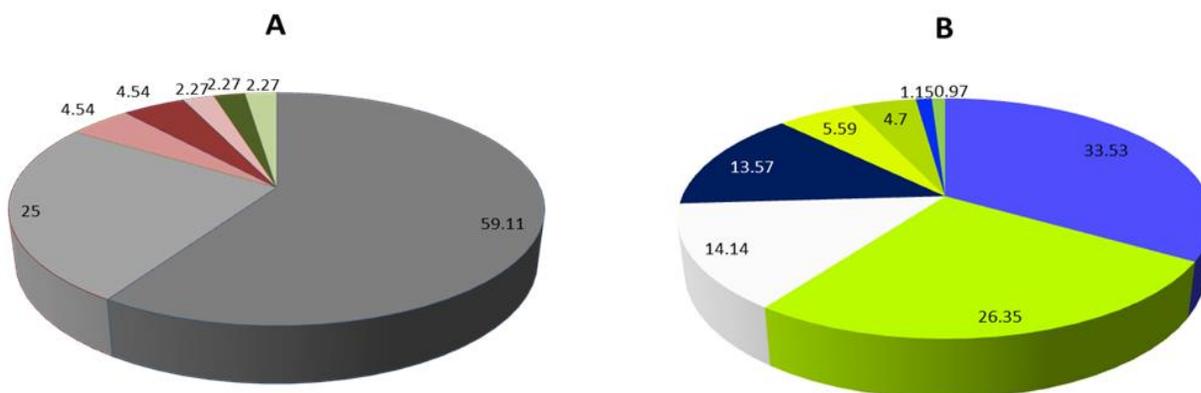
7.8.- Viroma y Bacterioma del pool 1 o adultos con agente etiológico viral.

Una vez que se escogió el procesamiento bioinformático C para realizar los siguientes análisis comenzamos con el pool 1, tenemos que de un archivo inicial con 415 116 lecturas y recordando que se tenían dos archivos, el primero R1 con sentido de lectura de 5' a 3' y R2 con un sentido de lectura de 3' a 5'. Cada archivo fue filtrado por calidad Q20 y después del mapeo a genoma humano con mapeador SMALT quedaron 160 524 lecturas para el archivo R1 y para R2 150 087 lecturas. Posteriormente se realizó el procesamiento con la herramienta BLAST y con el índice de virus, posteriormente se aplicó el algoritmo de GAAS y las lecturas obtenidas se corroboraron de forma manual. En el archivo R1 se obtuvieron 44 secuencias que pertenecen a virus, mientras que con R2 se obtuvieron 31 secuencias, en la gráfica 9 inciso A se muestra la proporción de virus obtenida.

Al analizar la región hipervariable V1-V3 tenemos que para el grupo de adultos con virus tenemos de un archivo original de 36 675 lecturas, 28 710 pasaron el filtro de calidad, después del análisis con las herramienta 16S Metagenomics de Illumina tenemos un total de

26 753 secuencias reportadas a nivel de especie, la clasificación de estas lecturas se muestran en la gráfica 4 inciso B.

Gráfica 4.- Viroma y Bacterioma del grupo de adultos con virus



Virus	Proporción (%)
Retrovirus endógenos	59.11
Herpesvirus	25
Adenovirus B	4.54
Influenza A H1N1	4.54
Coronavirus OC43	2.27
Fago de <i>Staphylococcus</i>	2.27
Fago de <i>Enterobacteria</i>	2.27

Bacteria	Proporción (%)
<i>Prevotella</i>	33.53
<i>Staphylococcus</i>	26.35
Sin clasificar	14.14
<i>Escherichia</i>	13.57
<i>Streptococcus</i>	5.59
<i>Corynebacterium</i>	4.7
<i>Veionella</i>	1.15
<i>Megasphaera</i>	0.97

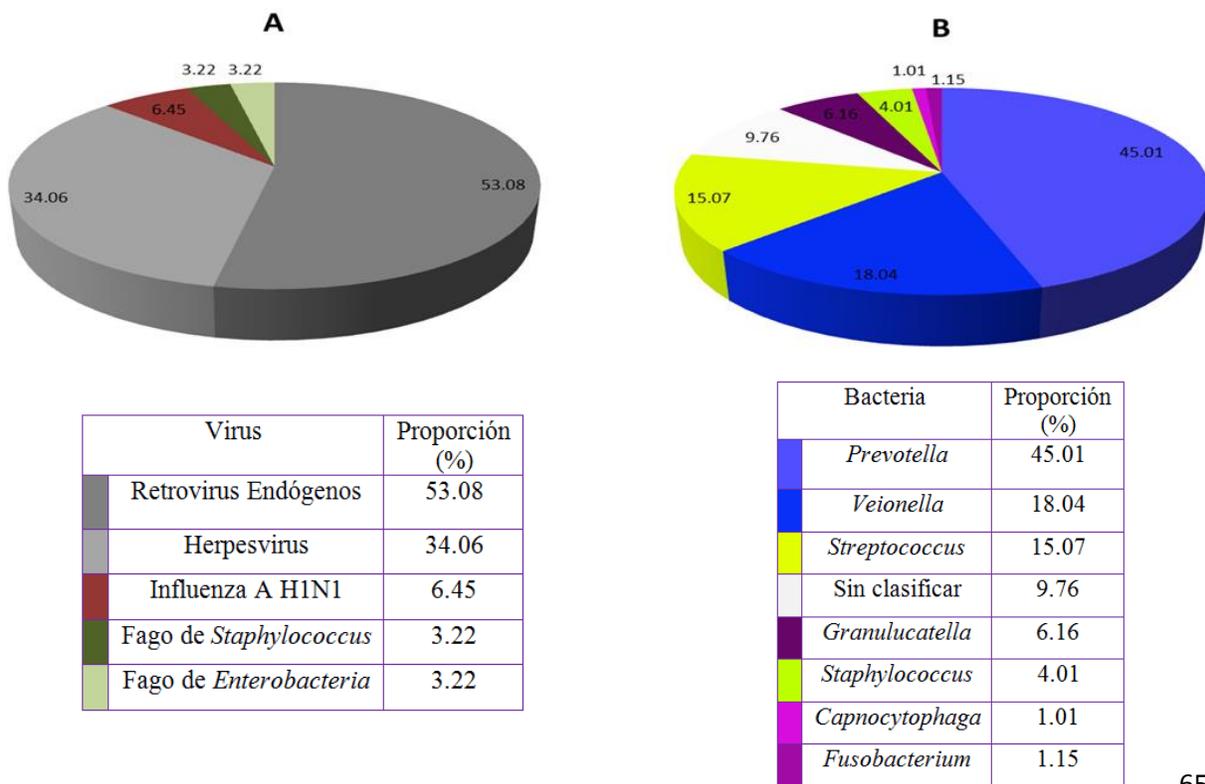
En la gráfica 4 inciso A podemos observar la presencia de 59.11% de retrovirus endógenos, 25% de Herpesvirus, seguido de Adenovirus B con 4.54%, con el mismo porcentaje Influenza A H1N1, coronavirus OC43 y al revisar los datos obtenidos mediante qPCR se puede observar que se había tipificado como coronavirus 229E, finalmente se tiene al fago de *Staphylococcus* en un 2.27% y en la misma proporción fago de *Enterobacteria*. En la gráfica 4 inciso B observamos la proporción de bacterias *Prevotella* 33.53%, *Staphylococcus* 26.35%, 14.14% no se pudo clasificar, *Escherichia* con 13.57%, *Streptococcus* 5.59%, *Corynebacterium* 4.7%, *Veionella* 1.15% y *Megasphaera* 0.97%.

7.9.- Viroma y Bacterioma del pool 2 o adultos sin agente etiológico viral.

En pool 2 tenía una cantidad de 2 264 145 lecturas iniciales, se filtraron por calidad y se mapearon con el índice de genoma humano y se obtuvo 208 658 y 206 260 lecturas para R1 y R2 respectivamente. Se realizó el alineamiento local con el índice de virus y BLAST, finalmente se usó el algoritmo de GAAS y la revisión manual de las secuencias procesadas. Para R1 se tiene que hay un total de 1032 secuencias que pertenecen a virus, mientras que para R2 son 850 lecturas, en la gráfica 5 inciso A se muestra la proporción de virus para este grupo de pacientes.

Para la fracción bacteriana se tenía un archivo original con 41 985 lecturas, 33 433 pasaron el filtro Q20 y después del análisis con la herramienta 16S Metagenomics de illumina tenemos un total de 31 588 secuencias reportadas a nivel de género, la clasificación de estas lecturas se muestran en la gráfica 5 inciso B.

Gráfica 5.- Viroma y Bacterioma del grupo de adultos sin virus



Para el caso del grupo de adultos en los que no se encontró la presencia de virus mediante la prueba de qPCR se observa que tenemos en el inciso A de la gráfica 5, un 44.14% de Herpesvirus, 33.33% del fago de *Staphylococcus*, 14.41% de Retrovirus Endógenos, 3.6% del fago de *Shigella* y en igual proporción el fago de *Enterobacteria*, finalmente con un 0.9% el fago *Evironmental*.

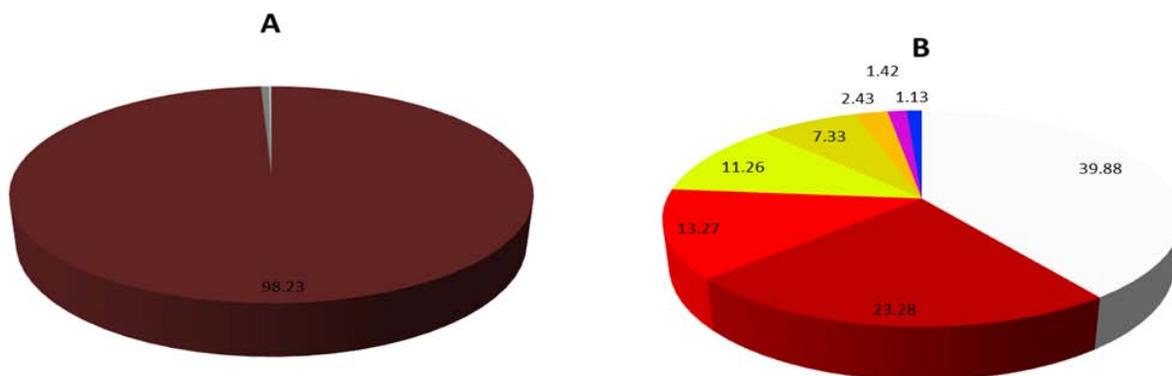
Para la fracción bacteriana tenemos en el gráfica 5 inciso B de la gráfica 10 un 45.01% de *Prevotella*, 18.04% de *Veionella*, 15.07% de *Streptococcus*, un 9.76% de lecturas sin clasificar, 6.16% de *Granullicatella*, 4.01% de *Staphylococcus*, 1.01% de *Capnocytophaga* y 0.94% de *Fusobacterium*.

7.10.- Viroma y Bacterioma del pool 3 o niños con agente etiológico viral.

Al realizar el proceso bioinformático C en el pool 3 tenemos que de un archivo original con 2 216 661 lecturas, al realizar el filtrado por calidad y el mapeo a genoma humano, tenemos que 160 242 secuencias no mapearon a genoma humano para el archivo R1, mientras que para el archivo R2 tenemos 177 621 secuencias que no mapearon al índice de humano. A partir de estos archivos se hace el alineamiento local con herramienta BLAST y utilizando el índice de virus, posteriormente se usó el algoritmo GAAS y de forma manual se verificó la identidad de las lecturas. Resultó en 6 527 secuencias que pertenecen a virus para el archivo R1 y para R2 6 478 secuencias que pertenecen a virus, en la gráfica 6 inciso A se muestra la proporción de virus.

Para la parte bacteriana se tenía un archivo original con 1 606 643 lecturas que se filtraron por calidad y quedaron 1 490 590, después del análisis con las herramientas de Illumina tenemos un total de 1 415 916 secuencias reportadas a nivel de género, la clasificación de estas lecturas se muestran en la gráfica 6 inciso B.

Gráfica 6.- Viroma y Bacterioma del grupo de niños con virus



Virus	Proporción (%)
Rinovirus C	98.23
Retrovirus Endógenos	0.47
Influenza A H1N1	0.03
Torque Teno Mínavirus	0.03
Adenovirus B	0.04
Fago de <i>Propionibacterium</i>	0.10
Fago de <i>Streptococcus</i>	0.03
Fago de <i>Staphylococcus</i>	0.04
Fago de <i>Enterobacteria</i>	0.01
Fago de <i>Escherichia</i>	0.01

Bacteria	Proporción (%)
Sin clasificar	39.88
<i>Serratia</i>	23.28
<i>Achrombacter</i>	13.27
<i>Streptococcus</i>	11.26
<i>Neisseria</i>	7.33
<i>Burkholderia</i>	2.43
<i>Porphyromonas</i>	1.42
<i>Veionella</i>	1.13

Los resultados que se observan de la gráfica 6 inciso A muestran que Rinovirus C se encuentra en mayor proporción un 99.23%, mientras que el resto, menos del 1% corresponde a un 0.47% de retrovirus endógenos, 0.10 del fago de *Propionibacterium*, 0.04% de Adenovirus B y en la misma proporción el fago de *Staphylococcus*, seguido de 0.03% de Influenza A H1N1, en la misma proporción Torque teno minivirus, así como el fago de *Streptococcus*, para el caso del fago de *Enterobacteria* lo encontramos en una proporción del 0.01% al igual que al fago de *Escherichia*.

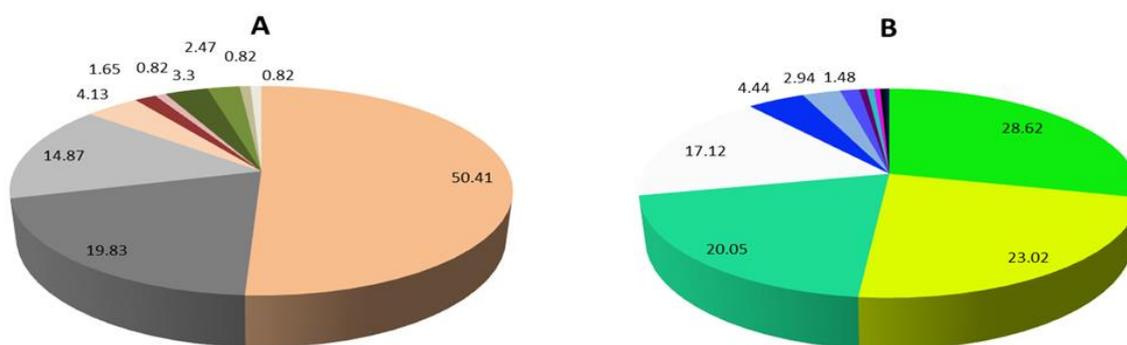
En la gráfica 6 inciso B se puede observar que un 39.88% de las lecturas no se pudo clasificar, seguido de *Serratia* con 23.28%, *Achrombacter* 13.27%, *Streptococcus* 11.26%, *Neisseria* 7.33%, *Burkholderia* 2.43%, *Porphyromonas* 1.42% y *Veionella* 1.13%.

7.11.- Viroma y Bacterioma del pool 4 o niños sin agente etiológico viral.

Para la parte del viroma del pool 4 se inició con 895 826 lecturas, se filtró por calidad y se mapeo a genoma humano con el mapeador SMALT, resulto en 282 275 secuencias que no mapearon a humano para R1 y 264 057 secuencias para R2. Con estos archivos se realizó el alineamiento local utilizando el índice de virus y herramienta BLAST. Posteriormente se usó el algoritmo GAAS, para finalmente corroborar de forma manual la identidad de las lecturas. Para R1 se obtuvo 121 secuencias que pertenecían a virus, mientras que para R2 se tiene 392, la proporción de virus se puede observar en la gráfica 7 inciso A.

En la parte bacteriana tenemos que de un archivo original con 103 466 lecturas, 77 763 pasaron el filtro de calidad y después del análisis con las herramientas de Illumina tenemos un total de 74 403 secuencias reportadas a nivel de especie, la clasificación de estas lecturas se muestran en la gráfica 7 inciso B.

Gráfica 7.- Viroma y Bacterioma del grupo de niños sin virus



Virus	Proporción (%)
Parvovirus B19	50.41
Retrovirus Endógenos	19.83
Torque Teno Minivirus	14.87
Enterovirus D68	4.13
Influenza A H1N1	1.65
Parainfluenza 3	0.82
Fago de <i>Staphylococcus</i>	3.3
Fago de <i>Streptococcus</i>	2.47
Fago de <i>Shigella</i>	0.82
Fago de <i>Lactococcus</i>	0.82

Bacteria	Proporción (%)
<i>Moraxella</i>	28.62
<i>Streptococcus</i>	23.02
<i>Haemophilus</i>	20.05
Sin clasificar	17.12
<i>Veionella</i>	4.44
<i>Alloicoccus</i>	2.94
<i>Prevotella</i>	1.48
<i>Escherichia</i>	0.59
<i>Granullicatella</i>	0.57
<i>Mycoplasma</i>	0.47
<i>Porphyromonas</i>	0.43
<i>Alkalibacterium</i>	0.27

Para el grupo de niños sin virus tenemos que en la gráfica 7 inciso A un 66.07% de Retrovirus Endógenos, un 19.64% de Parvovirus B19, un 8.16% de Torque Teno Minivirus, le sigue con 2.04% Enterovirus D68, 1.27% de Influenza A H1N1 y un 0.25% de Parainfluenza 3, para el caso de los fagos tenemos 1.53% de la presencia del fago de *Staphylococcus*, 0.76% del fago de *Streptococcus* y finalmente un 0.25% del fago de *Lactococcus*.

En la grafica 7 inciso B se muestran los resultados obtenidos para la región hipervariable V1-V3 del 16S ribosomal para el grupo de niños sin presencia de virus. En 28.62% *Moraxella*, 23.02% *Streptococcus*, 20.05% *Haemophilus*, un 17.12% sin clasificar, 4.44% *Veionella*, *Alloiococcus* 2.94%, *Prevotella* 1.48% y menos del 1% *Escherichia*, *Granullicatella*, *Mycolasma*, *Porphyromonas* y *Alkalibacterium*.

7.12.- Ensamblado de novo

De acuerdo a los resultados anteriores se realizó el ensamblado de rinovirus C, ya que es el virus con el cual se tienen suficientes secuencias para tal propósito. De la longitud de k-meros probados mediante velvet y velveth, con el k-mero de 189 fue con el que se obtuvo el mejor resultado. Se obtuvo 3 638 secuencias para construir el genoma, con una cobertura media de 28.84 en promedio y una longitud consenso de 6 644 nucleótidos. Esto se visualizó con el programa tablet. A partir del consenso obtenido se reportó en el NCBI la secuencia de rinovirus C Mex 14 con número de acceso KM486097.

VIII. DISCUSION

8.1.- Toma de muestra y prevalencia de virus respiratorios.

La toma de muestras se realizó de forma constante durante octubre de 2013 hasta junio de 2014 y sólo en pacientes que requirieron de hospitalización debido a IRAs, esto con el propósito de tener muestras representativas de la enfermedad y de la temporada invernal evaluada. Cada una de las muestras fue dividida en tres alícuotas, particularmente las que se tomaron para secuenciación fueron procesadas en las horas posteriores a la toma de muestra y se obtuvieron 36 muestras de las 40 propuestas al inicio del estudio.

El procesamiento en fresco se realizó con el fin de obtener la mayor sensibilidad de detección durante la secuenciación, es decir, su procesamiento casi de forma inmediata después de la toma de muestra, debió de ayudar a eliminar agentes causales de la degradación de los ácidos nucleicos y por lo tanto aumentar la posibilidad de detección de los agentes etiológicos causantes de las IRAs. Pero al analizar los resultados de la primera secuenciación solamente en el grupo de niños que se les detectó virus por qPCR de Biomark, se obtuvieron resultados adecuados para su análisis. Por lo tanto se decidió repetir a los mismos grupos con muestras disponibles en el banco de muestras congeladas a -80°C , esto fue de suma importancia, debido a que al analizar los resultados estos sufren cambios importantes.

Al comparar los grupos de niños con presencia de virus, las lecturas obtenidas son consistentes casi en su totalidad con lo detectado por qPCR, pero no resulta igual para el grupo de adultos con presencia de virus, donde los resultados obtenidos por secuenciación resultan distintos. Sugerimos que se debe principalmente a que las muestras congeladas sufren un proceso de degradación, mientras que el procesamiento en fresco no, en la sección de comparación entre grupos de pacientes se abordará con mayor detalle las diferencias obtenidas.

Otra característica de nuestro estudio fue realizar grupos de pacientes, se dividió entre adultos y niños y estos a su vez fueron divididos entre aquellos que se les detectó virus y entre los que no se les detectó virus utilizando qPCR con Biomark. Los cuatro grupos se formaron con varios fines:

El primero fue clasificar a los pacientes para poder evaluar las diferencias entre los grupos, ya que en la mayoría de los trabajos descritos en la tabla 3 sólo evalúan adultos ó niños sanos y/o con comorbilidades asociadas, la clasificación también tuvo otro propósito, que se pudieran asociar a los agentes etiológicos presentes con la severidad de la enfermedad. Aunque en el presente trabajo quedo pendiente recabar la información de las historias clínicas de los pacientes y poder hacer asociaciones.

En segundo lugar tenemos que se propuso realizar grupos para tratar de sumar los ácidos nucleicos de los posibles agentes etiológicos presentes y con esto obtener mayor cantidad de lecturas.

Tercero en base a lo propuesto por organizaciones de vigilancia epidemiológica a nivel mundial, se aprovecho la calificación de muestras propuesta, para el monitoreo de virus respiratorios con en un cierto lapso de tiempo y particularmente con las muestras de pacientes que resultaron negativas con las técnicas de diagnóstico por qPCR (Zumla et al. 2014).

Cuarto, al agrupar cierta cantidad de muestras se reduce la cantidad de reactivos, consumibles y en consecuencia se reducen costos, con la finalidad de optimizar recursos.

Independientemente de sólo haber secuenciado 116 muestras, en los 655 pacientes evaluados en la temporada invernal 2013-2014 se les realizó la detección de virus mediante qPCR con Biomark y de estos resultados tenemos lo siguiente:

En la parte introductoria ya se había mencionado que aproximadamente 60% de las IRAs son causadas por virus. Al observar los resultados obtenidos en nuestra cohorte de estudio

tenemos para pacientes menores de 15 tenemos una prevalencia de virus del 53%, mientras que en mayores a los 15 años tenemos casi un 30%. Cabe resaltar que el estudio se realizó desde octubre de 2013 hasta junio de 2014, debido a la toma de muestra y a los inconvenientes obtenidos en la primera secuenciación. A pesar que para el periodo 2013 a 2014 hubo una alta incidencia de influenza en el centro de México, también se sabe que la prevalencia de los virus es más alta entre los meses de octubre a febrero y el estudio se realizó con muestras tomadas hasta junio de 2014, siendo a partir de marzo donde la actividad de virus declinó y esto contribuyó a la baja cantidad de virus detectados por qPCR en Biomark (Dávila-Torres et al. 2015; Essa et al. 2015).

Los datos obtenidos son consistentes con otros estudios en México donde el virus con mayor prevalencia en población adulta es rinovirus, seguido de influenza y se invierten en temporada alta de influenza. Para la población de niños menores de 6 años se tiene una mayor prevalencia de influenza, seguido de virus sincicial respiratorio y rinovirus como los tres principales. En contraste en nuestra población menor de 5 años tenemos primero enterovirus/rinovirus, seguido de virus sincicial respiratorio e influenza como los tres principales agentes etiológicos presentes en la temporada evaluada.

Estos datos aportan a observaciones recientes, que la frecuencia de rinovirus se ha incrementado en los últimos años en población mexicana tanto en niños y adultos, siendo una importante causa de morbilidad en todas las edades (Galindo-Fraga et al. 2013; Landa-Cardena et al. 2012; Dávila-Torres et al. 2015)

8.2.- Preparación de muestras para estrategia de secuenciación shot gun.

Una vez obtenida la alícuota para secuenciación se procedió a realizar una serie de estrategias modificadas de RT-PCR y PCR como se había descrito por Taboada et al. y Prachayangprecha et al. en 2014. Con el uso de primers hexámeros aleatorios y hexámeros aleatorios con secuencia de bacteriófago M13 para amplificar a las moléculas de los agentes etiológicos presentes, de una forma azarosa para evitar sesgos, debido a que la secuencia de este bacteriófago no se encuentra presente en el hospedero, y no se ha

reportado su asociación con IRA. Al contrario si se tomara como base genomas de referencia y a partir de estos se amplificarán las moléculas presentes. La incorporación de las secuencias de bacteriófagos M13 mediante estuche First Strand cDNA Synthesis y con fragmento de klenow se realizó para adicionar los oligonucleótidos a los ácidos nucleicos presentes en las muestras y posteriormente realizar una ronda de amplificación. La obtención de estas moléculas a partir de la amplificación con los primers específicos hacia los oligonucleótidos de bacteriófago M13 aumento la posibilidad de tener mayor cantidad de lecturas de los agentes etiológicos presentes en las muestras.

Otro de los aspectos a considerar en nuestro estudio es que al utilizar filtros de 0.20 μm se pierden a los virus gigantes, los cuales podrían constituir un componente importante del microbioma en vías respiratorias, como se ha descrito en el microbioma del estómago (Lagier et al. 2012). Actualmente se proponen alternativas para modificar la poca cantidad de lecturas obtenidas y evitar el uso de filtros que pudieran afectar el estudio de virus que no puedan pasar a través de estos. El uso de sondas de captura de alrededor de 100-mer, construidas a partir de genomas de referencia de los virus de interés, permite capturar a los agentes etiológicos y remover el resto de la muestra. El inconveniente de utilizar genomas de referencia para capturar el material genético de los microorganismos, es que hay un sesgo hacia los cuales se diseño el experimento, dejando a un lado aquellos agentes emergentes. Pero el objetivo es positivo al tratar de tener una mayor cantidad de lecturas de virus, pues se demostró en este estudio y en otros realizados con técnicas similares, que la cantidad de lecturas del hospedero corresponden casi en su totalidad a los datos obtenidos (Grover, Salmon, y Wendel 2012).

Aunque hasta el momento estas tecnologías no aseguran enriquecer la muestras sin un sesgo implícito, el aumentar la cantidad de lecturas de los agentes etiológicos permitiría una mejor descripción de estos, así como mejores coberturas de su genoma. Mas adelante se abordará con mayor detalle el hecho de que al mezclar las muestras en grupos, se podría tener un efecto de dilución para algunos de los agentes etiológicos presentes lo cual

disminuye la capacidad de la técnica para poder analizar a los microorganismos presentes (Wylie et al. 2012; Grover, Salmon, y Wendel 2012).

8.3.- Comparación de los métodos bioinformáticos

Para realizar las comparaciones de los métodos bioinformáticos tenemos primero que mencionar que los resultados obtenidos con los procesamientos con Galaxy y Metavir son distintos de los otros métodos de análisis, debido a la naturaleza de los algoritmos y a las bases de datos que se utilizan. Pero la mayor diferencia la tenemos por el mapeador utilizado en Galaxy, tBLASTx es una herramienta que se basa en la búsqueda de semejanzas por proteínas, mientras que los procesos realizados en plataforma Biolinux se basan en el algoritmo BLASTn, esta herramienta se basa en la búsqueda de semejanzas mediante el uso de nucleótidos.

Aunque 17 639 lecturas de virus proporcionadas por Metavir son mas del doble que las 6 527 lecturas de virus obtenidas con el proceso bioinformático C, no podrían ser comparables ya que los resultados de Metavir resultan de la comparación de los motivos de proteínas. Por lo tanto con esta herramienta podría haber semejanzas que no necesariamente correspondan a los virus, mientras que el análisis por comparación de nucleótidos resulta más adecuado para proveer de una información más confiable, al tener la seguridad de que estas lecturas si corresponden a virus.

Posteriormente se utilizaron tres procesos similares en la plataforma de Biolinux y se realizaron con el mismo fin, analizar y comparar los nucleótidos de las lecturas obtenidas por secuenciación y utilizando bases de datos de humano y virus como referencia para su mapeo. En los tres procesos se cambio únicamente a los mapeadores para realizar el análisis, por lo tanto, lo único diferente fue el algoritmo de análisis utilizado.

El proceso escogido fue el proceso bioinformático C, ya que fue el proceso que mostró mayor sensibilidad al detectar al fago de *Enterobacteria*, poca cantidad de Retrovirus Endógenos y a pesar de no ser el procesamiento mas rápido, tampoco fue el que requirió de

mayor tiempo. Adicionalmente esta herramienta es la que menor requerimiento computacional necesita, lo cual es de gran utilidad para quienes no cuentan con una gran infraestructura de cómputo.

A pesar de que ya se encuentran disponibles mayor cantidad de herramientas para el análisis de lecturas de secuenciación, el objetivo del presente estudio no fue realizar una comparativa entre herramientas bioinformáticas, pero sí se decidió tener al menos tres opciones para escoger la que mejores resultados de análisis proveyera. La razón de utilizar los parámetros más excluyentes en la herramienta GAAS fue con el propósito de diferenciar bien entre las lecturas obtenidas y evitar los asignamientos erróneos de lectura y asegurar tener los resultados lo más confiables posibles. Finalmente se realizó la verificación manual de cada una de las lecturas para poder corroborar los resultados obtenidos.

En el caso del análisis hecho con la herramienta 16S Metagenomics de Base Space no se tiene punto de comparación, sólo se muestra en el presente trabajo un acercamiento a las bacterias presentes en las muestras de los grupos de pacientes analizados. Pero proponemos realizar un análisis más profundo con los datos obtenidos con la secuenciación de *de novo* y bases de datos de bacterias, esto permitiría comparar los resultados en este análisis y proponer cual tipo de procesamiento y análisis es el más adecuado para poder describir a las bacterias presentes en cada muestra.

También proponemos realizar otros análisis para la parte de bacterias y adicionar bases de datos de parásitos y hongos, lo cual nos proporcionaría una información más completa. Finalmente de las lecturas no identificadas después de todos estos análisis se les analice con otras herramientas más adecuadas para la búsqueda de *de novo* de virus como MetaGene, para la búsqueda por proteínas virales como las cápsides (Cobian Guemes y Arias-Ortiz 2013).

8.4.- Comparación entre qPCR en Biomark y NGS

8.4.1.- Adultos y niños con presencia de virus con Biomark

Al comparar los grupos de adultos con virus y niños con virus, clasificados así por el resultado obtenido por qPCR de Biomark, tenemos que no concuerdan los resultados obtenidos con NGS. En el grupo de adultos con virus esperábamos encontrar mayor cantidad de lecturas de influenza A, pero sólo obtuvimos 2 lecturas de 22 pacientes que previamente se les había tipificado influenza A mediante qPCR.

Encontramos 1 lectura de coronavirus OC43, pero por qPCR se había detectado en dos pacientes como coronavirus 229E, esto podría deberse a que en esta última técnica se tienen 48 ensayos distintos de las 6 familias principales de virus respiratorios y pudo haber un error en la interpretación de los datos ó durante el procesamiento de las muestras y ensayos, por lo tanto se tipificó erróneamente, por que al evaluar la secuencia obtenida tenemos la seguridad de que la presencia de coronavirus se debe al subtipo OC43. Finalmente tenemos que por qPCR también se tenía a un paciente con Enterovirus/Rinovirus, pero al observar los resultados de NGS no se tiene ninguna secuencia.

En el caso del grupo de niños con virus los datos son más congruentes entre las dos técnicas utilizadas, mediante qPCR se habían detectado 6 pacientes con la presencia de Enterovirus/Rinovirus y por NGS se obtuvieron 6475 secuencias de rinovirus C. Con qPCR se detectó Influenza A en 1 paciente y se obtuvieron 4 lecturas por NGS, a otro paciente se le detectó la presencia de Adenovirus y se obtuvieron 3 lecturas, pero en un paciente que se le detectó metapneumovirus por secuenciación no hubieron secuencias.

Podemos resaltar dos particularidades con estos datos, primero al comparar los grupos, en el de niños hay mayor congruencia entre los resultados de ambas técnicas y menor congruencia con los resultados obtenidos para el grupo de adultos. Hay que resaltar que el grupo de niños se procesó en la primera secuenciación a partir de muestras procesadas en fresco, mientras que el resto de grupos incluyendo el de adultos con presencia de virus se

realizaron en la segunda secuenciación, a partir de muestras congeladas, los resultados anteriores sugieren que el procesamiento de las muestras en las horas posteriores a su recolección aumenta la posibilidad de obtener mayor cantidad de lecturas.

También hay que resaltar que en ambos grupos puede haber un efecto de dilución para aquellos virus que se encuentran en concentraciones bajas, por ejemplo, en el grupo de niños había la presencia de metapneumovirus en un paciente, pero al realizar la secuenciación no se obtuvieron secuencias de este virus. La cantidad de individuos que conformaron a este grupo fueron 9 y se sabe que las técnicas de diagnóstico molecular son más sensibles que las técnicas de secuenciación, en consecuencia al mezclar las muestras para su secuenciación no se obtienen secuencias de metapneumovirus. Proponemos que esto se puede deber al volumen final alcanzado, en el cual la concentración de los ácidos nucleicos disminuye junto con la posibilidad de secuenciarlos.

8.4.1.- Adultos y niños sin presencia de virus con Biomark

Ahora revisaremos los grupos que se habían clasificado como negativos a virus en base a los resultados por qPCR. En el grupo de adultos tenemos la presencia de 23 lecturas de influenza, esto podría tener 2 posibles explicaciones, la primera es que los virus detectados por secuenciación tienen cambios importantes en su secuencia, de tal forma que los oligonucleótidos diseñados para su detección en qPCR no tengan la capacidad de detectarlos, mientras que la técnica de secuenciación al ser un método que no es blanco específico, esta técnica sí puede detectar estas secuencias. La segunda explicación posible es que estas lecturas sean el producto de una contaminación, resultado de toda la manipulación de las muestras al realizar todas las estrategias de extracción, de RT-PCR y PCR previo a la preparación de las bibliotecas.

Una solución podría ser diseñar oligonucleótidos específicos en base a las secuencias obtenidas por ejemplo de influenza y volver a probar con la técnica de qPCR en todas las muestras que conforman al grupo y saber si hay la presencia de estos virus, como ya se había mencionado esta técnica es más sensible. Otra solución es incluir un grupo control sin

IRA y así comparar los resultados obtenidos en este nuevo grupo y saber si los resultados anteriores fueron resultado de una contaminación.

En el caso del grupo de niños sin virus, tenemos que al igual que el grupo anterior a pesar de haber sido clasificado sin la presencia de virus mediante qPCR, también hay la presencia de Influenza A y Parainfluenza 3. Las secuencias de estos virus también podrían ser resultado de una contaminación ó debido a que la técnica de qPCR a pesar de tener oligonucleótidos específicos para la detección de estos virus, estos podrían tener cambios importantes en su secuencia que no permitan su detección por esta técnica. Al igual que en el grupo anterior proponemos generar oligonucleótidos específicos y probar en cada uno de los pacientes que conforman al grupo para poder concluir.

Adicionalmente en este grupo encontramos otros virus, 77 lecturas de Parvovirus B19 y 8 lecturas de Enterovirus D68. Parvovirus B19 está asociado a varios síndromes clínicos, incluyendo fiebre, secreción nasal y dolor de cabeza, erupción petequial purpúrea, dolor o inflamación de las articulaciones, artritis y manifestaciones neurológicas. Se sabe que su propagación es a través de las secreciones respiratorias y no existe tratamiento para la también conocida como la quinta enfermedad (debido a que alguna vez fue la quinta erupción cutánea más comúnmente diagnosticada por los médicos). Recientemente en Tokio Japón, se reportó un brote importante de más de 400 casos activos de erupción en niños (Tuccio et al. 2014; Kidd 2015; Barah et al. 2014).

Enterovirus D68 es un virus respiratorio reemergente, identificado desde 1962 en niños con neumonía. Además de producir enfermedad respiratoria en niños, puede causar enfermedad sistémica y enfermedad neurológica aguda. Este fue el caso de agosto de 2014 en los Estados Unidos de Norte América, donde comenzaron a llegar niños a los hospitales con síntomas severos de vías respiratorias bajas, inclusive en una clínica de Cleveland cerca del 45% de los pacientes requirieron ingresar a la unidad de cuidados intensivos (Foster et al. 2015). Al hacer estudios retrospectivos en cohortes de Filipinas y en California, E.U.A. se descubre que desde 2012 hay una asociación entre enterovirus D68 y síntomas neurológicos

o mielitis flácida aguda, así como enfermedad de vías respiratorias altas (Furuse et al. 2015; Greninger et al. 2015)

A partir de la información obtenida de las lecturas de secuenciación proponemos generar oligonucleótidos específicos para cada microorganismo, realizar pruebas de qPCR y conocer si la presencia de algunos virus, fagos ó bacterias fueron resultado de una contaminación ó si estos se encontraban presentes y en que pacientes, ya que hasta el momento no podemos saber si sólo un individuo ó la suma de todos contribuyeron a la obtención de las lecturas. Con esto también conoceríamos si la técnica de qPCR no fue capaz de detectarlos..

Hay otros virus que hasta el momento no se han mencionado y son Retrovirus Endógenos, Herpesvirus y Torque Teno Minivirus. Se sabe que Retrovirus Endógenos conforman el 8% del genoma humano, mientras que el caso de Herpesvirus tenemos que la mayoría de la población adquiere a este virus y que en ocasiones no ocasiona sintomatología alguna, pero que a lo largo de la vida se van teniendo picos de viremia correlacionados con supresión del sistema inmunológico y que para nuestro caso no esta relacionado como causa directa de IRA, mas bien como consecuencia del desequilibrio de salud de los individuos evaluados (Wenner 2008).

También tenemos la presencia de Torque Teno Minivirus el cual ya se había descrito en trabajos previos, donde no hay evidencia de que sea causa directa de IRAs, pero que si se ha detectado en adultos sanos y en casos severos de pacientes con EPOC (Feyzioglu et al. 2014; Burián et al. 2011; Hino y Miyata 2007). Inclusive se le usa como biomarcador en inmunosupresión (Béland et al. 2014).

En todos nuestros grupos de estudio encontramos la presencia de Retrovirus Endogenos e interesantemente Herpesvirus sólo lo encontramos presentes en los dos grupos de adultos y Torque Teno Minivirus sólo lo encontramos en los dos grupos de niños. Con los datos disponibles en los estudios revisados y en el presente estudio, aunque indican que estos

virus no necesariamente son patógenos, la contribución de éstos a la enfermedad aun queda por estudiarse.

8.5.- Comparación de Bacterias y fagos en todos los grupos de pacientes

Se estudio a la molecula 16S ribosomal para describir el bacterioma lo que es muy común, inclusive en la tabla 3 se muestra que en la mayoría de los trabajos enfocan su estudio a esta molécula, gracias a que ésta molécula no está presente en el hospedero y la información que nos aporta es suficiente para conocer la diversidad y composición de las bacterias dentro de las muestras analizadas (“- BaseSpace Application -” 2015; Illumina 2015).

La región hipervariable V1-V3 fue la región que se decidió estudiar, debido al soporte que se tiene por parte de Illumina para el análisis con la aplicación en BaseSpace. De acuerdo al reporte generado por grupo de muestras, se decidió mostrar los resultados a nivel de género (“- BaseSpace Application -” 2015; Illumina 2015)

Por otra parte el viroma del humano incluye virus que infectan a células eucariontes, pero también hay presencia de fagos o bacteriófagos que son virus que infectan exclusivamente bacterias, estos fagos se encuentran asociados a humanos sanos (Pride et al. 2012; Twomey et al. 2013). En general en nuestros resultados se encontraron poca cantidad de ellos y a continuación se comparan con los resultados obtenidos.

8.5.1.- Adultos con presencia de virus

Todas las bacterias encontradas en adultos sanos forman parte del microbioma de la cavidad nasal, mucosas, piel, boca, vagina y estómago. Son bacterias oportunistas que en un desequilibrio inmunológico producen enfermedad y se les pueden encontrar relacionadas con IRA, neumonía, difteria, lesiones de piel y vaginitis (Ramakrishnan et al. 2013; Martín et al. 2014; Adler, Mahony, y Friedman 2013; Tanaka et al. 2008; Datcu 2014).

Para el caso de los fagos encontrados en este grupo sólo tenemos 1 lectura de *Staphylococcus* y 1 de *Enterobacteria*, aunque si correlacionan los fagos encontrados con la presencia de las bacterias en este grupo de pacientes no podríamos saber si realmente se correlacionan con el ó los pacientes que tenían la presencia de bacteria y su respectivo fago, por lo que aquí también proponemos hacer oligonucleótidos específicos y evaluar a los individuos que conformaban este pool y así poder concluir si hay correlacion. Por otra parte destaca la poca cantidad de lecturas lo cual se pueda deber al efecto de dilución que ya se había mencionado anteriormente.

8.5.2.- Adultos sin presencia de virus

En el grupo de adultos sin presencia de virus, tenemos la prevalencia de *Prevotella* y *Veilonella*, se sugiere que un desequilibrio inmunológico favorezca la colonización de estas bacterias que se les encuentra como residentes de la cavidad oral, respiratoria y vaginal, de personas adultas sanas y también se les ha aislado en procesos infecciosos del tracto respiratorio (Tanaka et al. 2008; Martín et al. 2014; Ramakrishnan et al. 2013; Venkataraman et al. 2015).

Streptococcus y *Staphylococcus* también forman parte del microbioma de piel, mucosas, cavidad oral y respiratoria, siendo estas bacterias mas correlacionadas como parte del proceso infeccioso en IRA (S. NCBI 2015; Frank et al. 2010; Ramakrishnan et al. 2013). Las bacterias del género *Granulicatella* y *Fusobacterium* se les ha encontrado en co-agregación, formando biofilms en placa dental (Karched, Bhardwaj, y Asikainen 2015), a *Fusobacterium* se le ha encontrado en sujetos con periodontitis y al género *Capnocytophaga* en individuos con gingivitis (Park et al. 2015).

En este grupo encontramos la presencia de fagos de *Staphylococcus* y *Enterobacteria*, aunque si hay la presencia de bacterias que puedan infectar, ya se había mencionado que podrían realizarse otros análisis para poder correlacionar de manera mas adecuada la presencia de ambos microorganismos. Aunado a lo anterior se tiene la presencia de fagos

de *Propionibacterium*, *Shigella* y *Environmental*, los cuales no correlacionan su presencia con bacterias encontradas en este grupo (Oh et al. 2014).

Revisando el caso de *Propionibacterium* se sabe que si se encuentra presente en narinas anteriores, pero estos fagos también están relacionados con otras partes anatómicas, por ejemplo, piel. Lo anterior pasa de igual manera para *Shigella* y *Environmental*, la presencia de estos fagos podrían ser resultado de una contaminación de la piel, esto debido a la toma de muestra, ya que como se describió en la metodología, durante la toma de muestra se pasa por la cavidad nasal y se va arrastrando células que potencialmente están infectadas por virus y en ese momento arrastrar células del epitelio (Oh et al. 2014).

8.5.3.- Niños con presencia de virus

En el grupo de niños con virus, hay la presencia de *Serratia* que es el género de bacterias con mayor prevalencia, seguido de *Achrombacter*, ambos se les consideran como los más importantes patógenos nosocomiales oportunistas y particularmente a *Achrombacter* se le encuentra incrementado en grupo de pacientes con infección de influenza (Kraus-Haas, Mielke, y Simon 2015; Méndez et al. 2015; Lynch 2014).

Streptococcus y *Neisseria* se les encuentra como parte del microbioma de la cavidad nasal y vaginal, pero también se les encuentra asociadas a IRA (Datcu 2014; Martín et al. 2014; Ramakrishnan et al. 2013). A *Burkholderia* se le asocia con infecciones que afectan al pulmón principalmente en Asia y Australia de donde es endémica y en pacientes con fibrosis quística (Smet et al. 2015; Voronina et al. 2015). *Pophyromonas* forma parte del tracto respiratorio y se les encuentra relacionado con gingivitis en humanos (Li et al. 2015; Venkataraman et al. 2015). Finalmente *Veilonella* ya se había mencionado que se encuentra en cavidad respiratoria y oral, siendo un género de bacterias oportunistas (Ramakrishnan et al. 2013; Martín et al. 2014).

Los fagos encontrados en este grupo de pacientes son *Staphylococcus*, *Streptococcus*, *Enterobacteria*, *Escherichia* y *Propionibacterium*, al revisar las bacterias presentes en este

grupo observamos que no correlaciona ningún fago con las bacterias presentes, lo cual se podría deber a lo mencionado anteriormente, que durante la toma de muestra se pudo haber arrastrado a estos fagos de piel y no reflejan a la comunidad de fagos de la cavidad nasal (Oh et al. 2014).

8.5.4.- Niños sin presencia de virus

Todos los géneros de bacterias presentes en el grupo de niños sin virus se encuentran como parte del microbioma en humano y potencialmente están implicados en enfermedades, excepto *Alkalibacterium* el cual se ha aislado de productos marinos y no se le encuentra correlación con enfermedad respiratoria (Ramakrishnan et al. 2013; Martín et al. 2014; Ishikawa et al. 2009).

Moraxella es el género de bacteria con mayor prevalencia en este grupo y se le encuentra en cavidad respiratoria, pudiendo ser responsable de infecciones en tracto respiratorio, sinusitis y otitis media (Ramakrishnan et al. 2013; PHAC 2015). *Haemophilus* también se le ha encontrado en cavidad nasal en pacientes sanos, pero también se le conoce como un patógeno oportunista que causa IRA (Dajani, Asmar, y Thirumoorthi 1979).

Las bacterias del género *Alloicoccus* se le atribuye junto con *Haemophilus* y *Moraxella*, como causa de otitis media (Farajzadah Sheikh et al. 2015). Al género *Mycoplasma* se le encuentra en mucosas, en tracto respiratorio y urogenital, son bacterias causantes de neumonía en humanos (M. NCBI 2015). *Streptococcus*, *Prevotella*, *Veilonella*, *Escherichia*, *Granulicatella* y *Porphyromonas*, ya se habían descrito que forman parte del microbioma de piel, estomago, cavidad oral y respiratoria, pueden causar infecciones oportunistas como resultado de un desequilibrio inmunológico (Ramakrishnan et al. 2013; Venkataraman et al. 2015; Martín et al. 2014; Karched, Bhardwaj, y Asikainen 2015; Li et al. 2015).

Los fagos presentes en este grupo son *Staphylococcus*, *Streptococcus*, *Propionibacterium* y *Lactococcus*, nuevamente se tendrían que realizar oligonucleótidos específicos para discernir que pacientes contribuyen con estas lecturas y pensar que otras estrategias de toma de muestra podrían evitar posibles contaminaciones de piel, ya que como se pudo ver en todos los grupos de pacientes las bacterias y los fagos presentes en cada grupo no correlacionan.

En general podemos mencionar que al comparar a las bacterias presentes de acuerdo a los resultados obtenidos con 16S, tenemos interesantemente que en niños no hay la presencia de *Staphylococcus*, mientras que en adultos sí. En adultos y niños sin presencia de virus tenemos mayor presencia de bacterias residentes de la cavidad bucal, mientras que en adultos y niños con virus, encontramos más bacterias patógenas del tracto respiratorio o aquellas que se adquieren de forma nosocomial.

Aunado a la controversia que existe alrededor de algunos virus y fagos presentes durante una IRA, acerca del papel que juegan, como pueden modificar la microbiota y que impacto tienen en el nacimiento, desarrollo, nutrición y protección ante patógenos. El estilo de vida en las sociedades industrializadas reduce la exposición de las personas y los microbios, lo que ha modificado la función inmune por evitar o cambiar el establecimiento de las comunidades que conforman la microbiota de los individuos y que difieren substancialmente de las de nuestros antepasados (Farsani et al. 2015; Ursell et al. 2013). Junto con el problema en el uso de antibióticos, ya que cada vez es más común las infecciones con bacterias multidrogaresistentes, por lo que recientemente se han comenzado a retomar tratamientos alternativos con fagos para potenciar el efecto de los antibióticos y reducir la presencia de bacterias resistentes a los tratamientos.

8.6.- Comparación con otros estudios de la microbiota en IRA.

En la tabla 3 se muestran trabajos que estudian a la microbiota, principalmente en fibrosis quística y EPOC, utilizan 16S para el análisis de las bacterias, siendo su procesamiento de muestras diferente al nuestro y se realizan en otras partes del mundo. Por lo cual nos centraremos básicamente en el trabajo de Cobian et al y Taboada et al. Ya que los estudios se realizaron con pacientes del D.F. y Veracruz, México respectivamente y el procesamiento de muestras fue similar.

Para el trabajo de Cobian y colabs tenemos que la secuenciación se dividió en dos, una en secuenciación de *novo* y la otra en secuenciación de amplicones de influenza, al revisar los resultados de las lecturas de influenza obtenidas mediante la estrategia de *novo* los resultados son similares a los nuestros, excepto para los amplicones de Influenza, lo cual puede sugerir que los tratamientos para la obtención de los productos de RT-PCR y PCR con hexámeros M13 podrían ser muy agresivos para las moléculas de RNA, en consecuencia se reduce el rendimiento y se pierde información valiosa. De acuerdo a lo anterior en el equipo de Cobian decidieron hacer amplicones directamente de los ácidos nucleicos de influenza con oligonucleótidos específicos y con lo cual obtuvieron inclusive genoma completo.

En el trabajo desarrollado por Taboada y colabs, primero destaca que algunas muestras fueron procesadas de forma individual y otras en grupo, al revisar los datos que obtienen son muy diversos y no concluyen si es mejor la preparación de muestras en grupos o de forma individual. Lo que sí es ventajoso al realizar la preparación de muestras de forma individual es que se puede saber con exactitud que microbiota tiene el paciente evaluado y se evita el riesgo de sufrir dilución de las muestras y perder información valiosa.

Particularmente Cobian y colabs evaluaron la temporada de mayo a septiembre de 2009, mientras que Taboada y colabs evaluaron las temporadas de marzo de 2010 a abril de 2011 y de septiembre de 2011 a abril de 2012, eran muestras congeladas a $-70\text{ }^{\circ}\text{C}$, en contraste nuestro trabajo se realizó de octubre de 2013 a junio de 2014 y las muestras se procesaron

en fresco. Ellos no encontraron nuevos virus respiratorios, ni virus emergentes, en comparación nosotros encontramos enterovirus D68 y Parvovirus B19 que aunque no son nuevos virus, si son reemergentes. Otro virus que se encontró fue una nueva cepa rinovirus C y aunque en su trabajo también tienen presencia de rinovirus C no describen algo adicional.

Al comparar los trabajos previos por Lysholm y colabs, así como Taboada y colabs, con preparación de muestras similar a la nuestra tuvieron rendimientos del 10-20% de lecturas que no mapearon a humano, de donde obtuvieron lecturas de virus, fagos y bacterias. En comparación con nuestros datos obtuvimos desde 10-54% de lecturas que no mapearon a humano, el resto pertenecen en su mayoría a lecturas del hospedero. Los porcentajes obtenidos en nuestro estudio son prometedores y aunque en el grupo preparado en fresco fue donde menor cantidad de lecturas no mapeadas a humano se obtuvieron (10%), los resultados en comparación con los otros grupos fueron mejores, por lo cual concluimos que la preparación en fresco aumenta la cantidad de lecturas disponibles para secuenciar.

8.7.- Importancia del estudio de la microbiota en humanos

Se estima que entre 36 y 562 virus patógenos quedan por descubrirse y hay enfermedades en las cuales su etiología es desconocida, se sospecha que un virus pueda ser la causa, pero hacen falta más estudios para conocer el funcionamiento de la microbiota y como la interacción con el hospedero pueda conferir cierta protección ante algunos microbios patógenos o las enfermedades (Bibby 2013; Wylie, Weinstock, y Storch 2012). Y de acuerdo a los resultados obtenidos en nuestro estudio es importante estudiar aquellas lecturas donde no se les pudo clasificar como bacterias, virus o fagos, adicionar bases de datos de parásitos, hongos y aquellas que todavía queden sin clasificar usar otras estrategias de análisis, con MetaGene por ejemplo.

Hace falta realizar análisis que puedan contribuir más datos de importancia, pero también es importante la difusión de estos estudios y como pueden ayudar a modificar el esquema de tratamiento actual en base a antibióticos. Debido a la problemática con antibióticos,

aunque indudablemente han salvado billones de vidas, últimamente la aparición de cepas bacterianas resistentes a los fármacos es más frecuente cada vez, además que el impacto que tienen estos tratamientos en la microbiota de las personas apenas se comienza a estudiar y donde las plataformas de secuenciación masiva juegan un papel crucial, al proporcionarnos información de la composición de los virus, bacterias, hongos y parásitos que se encuentran presentes durante una enfermedad o en un estado sano (Kiley y Caler 2014; Ursell et al. 2013).

Recientemente hay avances de cómo podría beneficiar la microbiota a la salud de los humanos, estos estudios se han realizado al evaluar a recién nacidos que participaron en el proyecto del microbioma humano. El estudio lo realizan evaluando las heces de los niños que representa la comunidad microbiana del intestino, primero observan que hay diferencias entre los individuos que nacen por parto natural y aquellos que nacen por cesarea, aquellos niños que nacen por parto tienen un recubrimiento de bacterias del tracto vaginal proveniente de la madre, mientras que los que nacen por cesarea no.

Evalúan aquellos niños nacidos por parto normal hasta los 838 días después de nacer y observan que conforme van creciendo, aquel recubrimiento de bacterias provenientes de la madre se va modificando y se parece cada vez más al bacterioma de adultos sanos. Interesantemente también evalúan a los niños que por alguna infección requieren un tratamiento con antibióticos y como repercute este tratamiento en la comunidad bacteriana, observan como su bacterioma revierte súbitamente del bacterioma de adulto sano, al bacterioma vaginal con el que comenzaron al nacer, terminado el tratamiento se recupera al punto antes de recibir el tratamiento (Ursell et al. 2013).

Este trabajo lo mencionamos por que nos parece importante resaltar lo que hasta el momento se ha desarrollado con el estudio del microbioma humano y como interviene en la salud de los individuos o como puede revertir una enfermedad, por lo que proponemos se realicen trabajos a partir de la información aportada con este estudio, encaminada a realizar estudios como lo realizados por Ursell y colabs.

Finalmente cabe mencionar que las plataformas de secuenciación masiva, proveen de una importante información de la composición de la microbiota, pero el aislamiento de los virus para estudiar sus propiedades fenotípicas, tropismo y como interaccionan con los otros componentes de la microbiota, así como con el sistema inmunológico del hospedero son necesarios para que se pueda integrar toda la información multidisciplinariamente. El propósito sería realizar la vigilancia epidemiológica, el descubrimiento de virus, así como su diversidad, evolución y patogénesis, esto nos podría ayudar a entender la dinámica de la infección y poder mitigar o evitar los efectos de las infecciones en un futuro (Farsani et al. 2015; Smits y Osterhaus 2013).

IX. CONCLUSIONES

Los resultados obtenidos por qPCR de Biomark aportan datos que refuerzan observaciones recientes, que la frecuencia de Rinovirus se ha incrementado en los últimos años en población mexicana, siendo causa importante de morbilidad en todas las edades.

A partir de muestras de hisopado nasofaríngeo y realizando las estrategias de extracción de ácidos nucleicos, RT-PCR y PCR, se obtuvieron buenas concentraciones de ácidos nucleicos, permitiendo la preparación de bibliotecas y su posterior secuenciación. Con las lecturas obtenidas se pudo detectar virus reemergentes y aquellos que mediante técnicas de diagnóstico molecular no se pueden detectar, como es el caso de Parvovirus B19, Enterovirus D68 y a la variante de rinovirus C, con las lecturas de éste último virus se realizó el ensamble de su genoma para su posterior caracterización.

Con las estrategias de extracción y de amplificación del material genético se obtuvieron rendimientos desde 10% hasta el 54% de lecturas que no correspondían a humano, comparando nuestros resultados en trabajos como los de Lysholm y Taboada, donde obtuvieron desde un 10% hasta un 20% de lecturas no humanas, tenemos que obtuvimos mayor rendimiento. Interesantemente en el grupo de muestras procesadas en las horas posteriores a su recolección tenemos que fue en el grupo donde mayor cantidad de lecturas obtuvimos, por lo que recomendamos el procesamiento de las muestras en las horas posteriores a su recolección.

El procesamiento de muestras en grupos de pacientes ayudo a reducir la cantidad de reacciones, consumibles y en consecuencia reducir costos, además de permitir la vigilancia epidemiológica de los sectores de la población evaluados, pudiendo detectar a los virus reemergentes, enterovirus D68 y Parvovirus B19, los cuales causaron brotes importantes en E.U.A y Japón respectivamente. Pero también se pudo observar que algunos virus que se encuentran en baja concentración, no pudieron ser secuenciados, en contraste las técnicas de qPCR de Biomark si pudieron detectarlos, lo cual ya se conocía, que los métodos de diagnóstico molecular son mas sensibles.

El análisis bioinformático elegido (con mapeador SMALT) resultó tener la mayor sensibilidad de detección. Además de que por sus características, no necesita muchos recursos computacionales y por lo tanto no es necesaria una infraestructura computacional costosa.

La presencia de algunos virus no necesariamente indica que son patógenos como se describió para herpesvirus y torque teno minivirus, pero este estudio ayuda a describir la presencia de estos, quedando pendiente el estudiar la contribución de los virus en la salud o en la enfermedad.

Al comparar los resultados de las bacterias presentes en los cuatro grupos analizados, tenemos que en los grupos de adultos había la presencia de Staphylococcus, pero en niños no. En los grupos de adultos y niños sin presencia de virus, había mayor cantidad de bacterias residentes de la cavidad bucal, en contraste con los grupos de adultos y niños con presencia de virus, encontramos mayor cantidad de bacterias patógenas del tracto respiratorio o aquellas que se adquieren de forma nosocomial.

En los grupos tipificados como negativos con la técnica de qPCR de Biomark, al comparar los resultados obtenidos por secuenciación tenemos que en todos los grupos de pacientes en realidad si hay presencia de virus, como Enterovirus D68 y Parvovirus B19. Aunque quedaría pendiente corroborar algunos datos como los de influenza, para lo cual sería interesante adicionar un grupo control y generar oligonucleótidos específicos, esto permitiría saber si hay contaminación por los procesos de extracción y amplificación.

X. PERSPECTIVAS

Evaluar las lecturas obtenidas por NGS de *novo* no utilizadas en este estudio, con bases de datos de bacterias, parásitos y hongos, para conocer si hay la presencia de estos microbios y discriminar cuantas secuencias no se clasifican en humano y microorganismos, con estas lecturas aplicar otras estrategias para la búsqueda de nuevos virus (Edwards y Rohwer 2005).

La identificación de nuevos virus se puede realizar mediante el estudio de sus proteínas, al evaluar similitud en sus secuencias con virus conocidos y el estudio de marcos de lecturas abiertos para la búsqueda de motivos que nos puedan dar indicios de microorganismos presentes (Delwart 2013). También se pueden utilizar modelos ocultos de Markov, sus algoritmos se basan en buscar semejanzas con proteínas, polimerasas o capsides virales (Cobian Guemes y Arias-Ortiz 2013).

Sugerimos revisar la estrategia de filtrado con membranas de 0.22 μm , debido a que al utilizarlas se pierden a los virus gigantes, los cuales podrían constituir un componente importante de la microbiota de la vías respiratorias. También sugerimos revisar los procedimientos de preparación de muestras de Lysholm y colabs, debido a que ellos reportan proporciones de lecturas similares a las obtenidas en este trabajo, pero a partir de muestras congeladas.

Proponemos revisar otras estrategias de secuenciación masiva, a partir de muestras individuales, utilizando grupos de muestras, pero etiquetadas de forma individual y por otra parte probar las técnicas de captura y comprar cual de las estrategias es la más adecuada para el diagnóstico y la vigilancia epidemiológica.

Evaluar a cada uno de los individuos que conformaban los grupos estudiados con oligonucleótidos específicos, generados a partir de las lecturas obtenidas en este estudio. Lo anterior permitiría conocer con técnicas más sensibles, si algunas de las lecturas corresponden verdaderamente a los virus encontrados ó son producto de una

contaminación. Aunado a lo anterior proponemos adicionar un grupo control sin IRA, para conocer si algunos microorganismos se encuentran presentes en individuos sanos y poder hacer asociaciones con la presencia ó ausencia de los microorganismos y la enfermedad.

Correlacionar la historia clínica de los pacientes y la severidad de la enfermedad, con la presencia de los virus, bacterias, parásitos y hongos. También proponemos estudios complementarios a la secuenciación, donde se evalúen las propiedades fenotípicas, tropismo y como interaccionan los microorganismos con los otros componentes de la microbiota, así como el sistema inmunológico del hospedero para poder integrar los resultados y conocer los efectos moduladores de los microbios y determinar la dinámica de las enfermedades para poder mitigarlas o evitarlas.

XI. LITERATURA CITADA

- “2100 Agilent Bioanalyzer”. 2015. Consultado junio 28.
<http://www.chem.agilent.com/Library/applications/5991-3322EN.pdf>.
- Adler, N. R., A. Mahony, y N. D. Friedman. 2013. “Diphtheria: Forgotten, but Not Gone”. *Internal Medicine Journal* 43 (2): 206–10. doi:10.1111/imj.12049.
- “Agencourt AMPure XP.book - techdocs”. 2015. Consultado junio 28.
<https://www.beckmancoulter.com/wsrportal/techdocs?docname=B37419AA>.
- Agilent Technologies. 2015. “2100 Bioanalyzer Instruments”. Consultado junio 28.
http://www.genomics.agilent.com/en/product.jsp?cid=AG-PT-106&_requestid=266380.
- Allen, E Kaitlynn, Alex F Koeppel, J Owen Hendley, Stephen D Turner, Birgit Winther, y Michèle M Sale. 2014. “Characterization of the nasopharyngeal microbiota in health and during rhinovirus challenge”. *Microbiome* 2 (junio): 22. doi:10.1186/2049-2618-2-22.
- Angly, Florent E., Dana Willner, Alejandra Prieto-Davó, Robert A. Edwards, Robert Schmieder, Rebecca Vega-Thurber, Dionysios A. Antonopoulos, et al. 2009. “The GAAS Metagenomic Tool and Its Estimations of Viral and Microbial Average Genome Size in Four Major Biomes”. *PLoS Computational Biology* 5 (12): e1000593.
doi:10.1371/journal.pcbi.1000593.
- “Applied Biosystems”. 2015. Consultado junio 27.
https://tools.lifetechnologies.com/content/sfs/manuals/sp_9520.pdf.
- Arias, Carlos F., Marina Escalera-Zamudio, María de Los Dolores Soto-Del Río, Ana Georgina Cobián-Güemes, Pavel Isa, y Susana López. 2009. “Molecular Anatomy of 2009 Influenza Virus A (H1N1)”. *Archives of Medical Research* 40 (8): 643–54.
doi:10.1016/j.arcmed.2009.10.007.
- Barah, Faraj, Sigrid Whiteside, Sonia Batista, y Julie Morris. 2014. “Neurological Aspects of Human Parvovirus B19 Infection: A Systematic Review”. *Reviews in Medical Virology* 24 (3): 154–68. doi:10.1002/rmv.1782.
- “- BaseSpace Application -”. 2015. Consultado julio 7.
<https://basespace.illumina.com/apps/593593/16S-Metagenomics?preferredversion>.
- Bassis, Christine M., John R. Erb-Downward, Robert P. Dickson, Christine M. Freeman, Thomas M. Schmidt, Vincent B. Young, James M. Beck, Jeffrey L. Curtis, y Gary B. Huffnagle. 2015. “Analysis of the Upper Respiratory Tract Microbiotas as the Source of the Lung and Gastric Microbiotas in Healthy Individuals”. *mBio* 6 (2). doi:10.1128/mBio.00037-15.
- Beck, Eric T., y Kelly J. Henrickson. 2010. “Molecular Diagnosis of Respiratory Viruses”. *Future Microbiology* 5 (6): 901–16. doi:10.2217/fmb.10.48.
- Beck, James M., Vincent B. Young, y Gary B. Huffnagle. 2012. “The Microbiome of the Lung”. *Translational research : the journal of laboratory and clinical medicine* 160 (4): 258–66.
doi:10.1016/j.trsl.2012.02.005.
- Béland, Kathie, Michael Dore-Nguyen, Marie-Josée Gagné, Nathalie Patey, Julie Brassard, Fernando Alvarez, y Ugur Halac. 2014. “Torque Teno Virus Load as a Biomarker of Immunosuppression? New Hopes and Insights”. *The Journal of Infectious Diseases* 210 (4): 668–70. doi:10.1093/infdis/jiu210.
- Bender, L., M. Ott, R. Marre, y J. Hacker. 1990. “Genome Analysis of Legionella Ssp. by Orthogonal Field Alternation Gel Electrophoresis (OFAGE)”. *FEMS Microbiology Letters* 60 (3): 253–57.

- Berry, Michael, Junaid Gamielien, y Burtram C. Fielding. 2015. "Identification of New Respiratory Viruses in the New Millennium". *Viruses* 7 (3): 996–1019. doi:10.3390/v7030996.
- Bialasiewicz, Seweryn, Jodie McVernon, Terry Nolan, Stephen B Lambert, Guoyan Zhao, David Wang, Michael D Nissen, y Theo P Sloots. 2014. "Detection of a divergent Parainfluenza 4 virus in an adult patient with influenza like illness using next-generation sequencing". *BMC Infectious Diseases* 14 (mayo): 275. doi:10.1186/1471-2334-14-275.
- Bibby, Kyle. 2013. "Metagenomic Identification of Viral Pathogens". *Trends in Biotechnology* 31 (5): 275–79. doi:10.1016/j.tibtech.2013.01.016.
- BioFire Diagnostics. 2015. "The FilmArray System". Consultado junio 24. <http://filmarray.com/the-filmarray-system/>.
- Bogaert, Debby, Bart Keijser, Susan Huse, John Rossen, Reinier Veenhoven, Elske van Gils, Jacob Bruin, Roy Montijn, Marc Bonten, y Elisabeth Sanders. 2011. "Variability and Diversity of Nasopharyngeal Microbiota in Children: A Metagenomic Analysis". *PLoS ONE* 6 (2). doi:10.1371/journal.pone.0017035.
- Burián, Zsófia, Hajnalka Szabó, Gyöngyi Székely, Kálmán Gyurkovits, Péter Pankovics, Tibor Farkas, y Gábor Reuter. 2011. "Detection and Follow-up of Torque Teno Midi Virus ('small Anelloviruses') in Nasopharyngeal Aspirates and Three Other Human Body Fluids in Children". *Archives of Virology* 156 (9): 1537–41. doi:10.1007/s00705-011-1021-0.
- "Burrows-Wheeler Aligner". 2015. Consultado julio 7. <http://bio-bwa.sourceforge.net/>.
- Carmody, Lisa A., Jiangchao Zhao, Patrick D. Schloss, Joseph F. Petrosino, Susan Murray, Vincent B. Young, Jun Z. Li, y John J. LiPuma. 2013. "Changes in Cystic Fibrosis Airway Microbiota at Pulmonary Exacerbation". *Annals of the American Thoracic Society* 10 (3): 179–87. doi:10.1513/AnnalsATS.201211-107OC.
- CDC. 2015. "Pneumococcal Disease | Clinical | Streptococcus pneumoniae | CDC". Consultado junio 20. <http://www.cdc.gov/pneumococcal/clinicians/streptococcus-pneumoniae.html>.
- Chaban, Bonnie, Arianne Albert, Matthew G. Links, Jennifer Gardy, Patrick Tang, y Janet E. Hill. 2013. "Characterization of the Upper Respiratory Tract Microbiomes of Patients with Pandemic H1N1 Influenza". *PLoS ONE* 8 (7). doi:10.1371/journal.pone.0069559.
- Charlson, Emily S., Kyle Bittinger, Jun Chen, Joshua M. Diamond, Hongzhe Li, Ronald G. Collman, y Frederic D. Bushman. 2012. "Assessing Bacterial Populations in the Lung by Replicate Analysis of Samples from the Upper and Lower Respiratory Tracts". *PLoS ONE* 7 (9). doi:10.1371/journal.pone.0042786.
- Charlson, Emily S., Kyle Bittinger, Andrew R. Haas, Ayannah S. Fitzgerald, Ian Frank, Anjana Yadav, Frederic D. Bushman, y Ronald G. Collman. 2011. "Topographical Continuity of Bacterial Populations in the Healthy Human Respiratory Tract". *American Journal of Respiratory and Critical Care Medicine* 184 (8): 957–63. doi:10.1164/rccm.201104-0655OC.
- Charlson, Emily S., Jun Chen, Rebecca Custers-Allen, Kyle Bittinger, Hongzhe Li, Rohini Sinha, Jennifer Hwang, Frederic D. Bushman, y Ronald G. Collman. 2010. "Disordered Microbial Communities in the Upper Respiratory Tract of Cigarette Smokers". *PLoS ONE* 5 (12). doi:10.1371/journal.pone.0015216.
- Chiu, Charles Y. 2013. "Viral pathogen discovery". *Current Opinion in Microbiology*, Host–microbe interactions: fungi/parasites/viruses, 16 (4): 468–78. doi:10.1016/j.mib.2013.05.001.
- Cobian Guemes, Ana Georgina, y Carlos Federico Arias-Ortiz. 2013. "Metagenómica viral y bacteriana de enfermedades respiratorias graves". Cuernavaca, Morelos: Universidad Nacional Autónoma de México. <http://oreon.dgbiblio.unam.mx/F/S7Q3D8ILEL7JD1N1YUIX223LBM46BJ4SL3USDIKCHIJE341KMU-22157?func=find->

- b&request=metagenomica+viral+y+bacteriana+de+enfermedades+respiratorias+graves&find_code=WRD&adjacent=N&local_base=TES01&x=0&y=0&filter_code_2=WYR&filter_request_2=&filter_code_3=WYR&filter_request_3=.
- Coburn, Bryan, Pauline W. Wang, Julio Diaz Caballero, Shawn T. Clark, Vijaya Brahma, Sylva Donaldson, Yu Zhang, et al. 2015. "Lung microbiota across age and disease stage in cystic fibrosis". *Scientific Reports* 5 (mayo). doi:10.1038/srep10241.
- "Copan". 2015. Consultado junio 27.
<http://www.copaninnovation.com/studies/download.php?id=510>.
- Cosio Villegas, Ismael, Alejandro Celis Salazar, Miguel Cosio Pascal, y Manuel Alfredo Urbina Alvarenga. 2006. *Aparato Respiratorio*. 17a ed. Méndez Editores.
- Cox, Michael J., William O. C. M. Cookson, y Miriam F. Moffatt. 2013. "Sequencing the Human Microbiome in Health and Disease". *Human Molecular Genetics* 22 (R1): R88–94. doi:10.1093/hmg/ddt398.
- Cremers, Amelieke JH, Aldert L Zomer, Jenna F Gritzfeld, Gerben Ferwerda, Sacha AFT van Hijum, Daniela M Ferreira, Joshua R Shak, et al. 2014. "The adult nasopharyngeal microbiome as a determinant of pneumococcal acquisition". *Microbiome* 2 (diciembre). doi:10.1186/2049-2618-2-44.
- Dajani, A. S., B. I. Asmar, y M. C. Thirumoorthi. 1979. "Systemic Haemophilus Influenzae Disease: An Overview". *The Journal of Pediatrics* 94 (3): 355–64.
- Datcu, Raluca. 2014. "Characterization of the Vaginal Microflora in Health and Disease". *Danish Medical Journal* 61 (4): B4830.
- Dávila-Torres, Javier, Gerardo Chowell, Víctor H. Borja-Aburto, Cécile Viboud, Concepción Grajales-Muñiz, y Mark A. Miller. 2015. "Intense Seasonal A/H1N1 Influenza in Mexico, Winter 2013-2014". *Archives of Medical Research* 46 (1): 63–70. doi:10.1016/j.arcmed.2014.11.005.
- Delwart, Eric. 2013. "A Roadmap to the Human Virome". *PLoS Pathogens* 9 (2): e1003146. doi:10.1371/journal.ppat.1003146.
- Dewall, Michael Thomas, y Davis W. Cheng. 2011. "The Minimal Genome: A Metabolic and Environmental Comparison". *Briefings in Functional Genomics* 10 (5): 312–15. doi:10.1093/bfpg/elr030.
- Donatin, E., y M. Drancourt. 2012. "DNA Microarrays for the Diagnosis of Infectious Diseases". *Médecine Et Maladies Infectieuses* 42 (10): 453–59. doi:10.1016/j.medmal.2012.07.017.
- Edwards, Robert A., y Forest Rohwer. 2005. "Viral Metagenomics". *Nature Reviews. Microbiology* 3 (6): 504–10. doi:10.1038/nrmicro1163.
- Eiros, José M., Raúl Ortiz de Lejarazu, Alberto Tenorio, Inmaculada Casas, Francisco Pozo, Guillermo Ruiz, y Pilar Pérez-Breña. 2009. "[Microbiological diagnosis of viral respiratory infections]". *Enfermedades Infecciosas Y Microbiología Clínica* 27 (3): 168–77. doi:10.1016/j.eimc.2008.03.004.
- Epicentre. 2015. "Exo-Minus Klenow DNA Polymerase (D355A- E357A)". Consultado junio 28.
[http://www.epibio.com/enzymes/dna-and-rna-polymerases/dna-polymerases-mesophilic/exo-minus-klenow-dna-polymerase-\(d355a-e357a\)](http://www.epibio.com/enzymes/dna-and-rna-polymerases/dna-polymerases-mesophilic/exo-minus-klenow-dna-polymerase-(d355a-e357a)).
- Epstein, S. S. 2013. "The Phenomenon of Microbial Uncultivability". *Current Opinion in Microbiology* 16 (5): 636–42. doi:10.1016/j.mib.2013.08.003.
- Erb-Downward, John R., Deborah L. Thompson, Meilan K. Han, Christine M. Freeman, Lisa McCloskey, Lindsay A. Schmidt, Vincent B. Young, et al. 2011. "Analysis of the Lung Microbiome in the 'Healthy' Smoker and in COPD". *PLoS ONE* 6 (2): e16384. doi:10.1371/journal.pone.0016384.

- Essa, Sahar, Abdullah Owayed, Haya Altawalrah, Mousa Khadadah, Nasser Behbehani, y Widad Al-Nakib. 2015. "The Prevalence of Human Bocavirus, Human Coronavirus-NL63, Human Metapneumovirus, Human Polyomavirus KI and WU in Respiratory Tract Infections in Kuwait". *Medical Principles and Practice: International Journal of the Kuwait University, Health Science Centre* 24: 382–87. doi:10.1159/000381422.
- Farajzadah Sheikh, Ahmad, Nader Saki, Mitra Roointan, Reza Ranjbar, Mohammad Jaafar Yadyad, Abbas Kaydani, Sajad Aslani, Mansoor Babaei, y Hamed Goodarzi. 2015. "Identification of Alloiococcus Otitidis, Streptococcus Pneumoniae, Moraxella Catarrhalis and Haemophilus Influenzae in Children With Otitis Media With Effusion". *Jundishapur Journal of Microbiology* 8 (3): e17985. doi:10.5812/jjm.17985.
- Farsani, Seyed Mohammad Jazaeri, Martin Deijis, Ronald Dijkman, Richard Molenkamp, Rien E. Jeeninga, Margareta Ieven, Herman Goossens, y Lia van der Hoek. 2015. "Culturing of Respiratory Viruses in Well-Differentiated Pseudostratified Human Airway Epithelium as a Tool to Detect Unknown Viruses". *Influenza and Other Respiratory Viruses* 9 (1): 51–57. doi:10.1111/irv.12297.
- "FASTX-Toolkit". 2015. Consultado julio 7. http://hannonlab.cshl.edu/fastx_toolkit/.
- Feyzioglu, Bahadir, Turgut Teke, Mehmet Ozdemir, Adnan Karaibrahimoğlu, Metin Doğan, y Mehmet Yavşan. 2014. "The Presence of Torque Teno Virus in Chronic Obstructive Pulmonary Disease". *International Journal of Clinical and Experimental Medicine* 7 (10): 3461–66.
- Fields, Bernard N. 2013. *Fields Virology*. Editado por Jeffrey I Cohen, Diane E Griffin, Robert A Lamb, Malcom A Martin, Vicent R Racaniello, y Bernard Roizman. Sexta. Vol. 1. Philadelphia: Lippincott Williams & Wilkins, a Wolters Kluwer business.
- Finlay, B. Brett, y Grant McFadden. 2006. "Anti-Immunology: Evasion of the Host Immune System by Bacterial and Viral Pathogens". *Cell* 124 (4): 767–82. doi:10.1016/j.cell.2006.01.034.
- Foster, Charles B., Neil Friedman, John Carl, y Giovanni Piedimonte. 2015. "Enterovirus D68: A Clinically Important Respiratory Enterovirus". *Cleveland Clinic Journal of Medicine* 82 (1): 26–31. doi:10.3949/ccjm.82a.14166.
- Frank, Daniel N., Leah M. Feazel, Mary T. Bessesen, Connie S. Price, Edward N. Janoff, y Norman R. Pace. 2010. "The Human Nasal Microbiota and Staphylococcus aureus Carriage". *PLoS ONE* 5 (5). doi:10.1371/journal.pone.0010598.
- Furuse, Yuki, Natthawan Chaimongkol, Michiko Okamoto, Tadatsugu Imamura, Mariko Saito, Raita Tamaki, Mayuko Saito, Tohoku-RITM Collaborative Research Team, Socorro P. Lupisan, y Hitoshi Oshitani. 2015. "Molecular Epidemiology of Enterovirus d68 from 2013 to 2014 in Philippines". *Journal of Clinical Microbiology* 53 (3): 1015–18. doi:10.1128/JCM.03362-14.
- "Galaxy". 2015. Consultado julio 7. <https://usegalaxy.org/>.
- Galindo-Fraga, Arturo, Ana A. Ortiz-Hernández, Alejandra Ramírez-Venegas, Rafael Valdez Vázquez, Sarbelio Moreno-Espinosa, Beatriz Llamosas-Gallardo, Santiago Pérez-Patrigeon, et al. 2013. "Clinical Characteristics and Outcomes of Influenza and Other Influenza-like Illnesses in Mexico City". *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases* 17 (7): e510–17. doi:10.1016/j.ijid.2013.01.006.
- Ghebremedhin, B. 2014. "Human adenovirus: Viral pathogen with increasing importance". *European Journal of Microbiology & Immunology* 4 (1): 26–33. doi:10.1556/EuJMI.4.2014.1.2.
- Goddard, Amanda F., Benjamin J. Staudinger, Scot E. Dowd, Amruta Joshi-Datar, Randall D. Wolcott, Moira L. Aitken, Corinne L. Fligner, y Pradeep K. Singh. 2012. "Direct sampling of

- cystic fibrosis lungs indicates that DNA-based analyses of upper-airway specimens can misrepresent lung microbiota". *Proceedings of the National Academy of Sciences of the United States of America* 109 (34): 13769–74. doi:10.1073/pnas.1107435109.
- Gong, Hong-Li, Yi Shi, Liang Zhou, Chun-Ping Wu, Peng-Yu Cao, Lei Tao, Chen Xu, Dong-Sheng Hou, y Yue-Zhu Wang. 2013. "The Composition of Microbiome in Larynx and the Throat Biodiversity between Laryngeal Squamous Cell Carcinoma Patients and Control Population". *PLoS ONE* 8 (6). doi:10.1371/journal.pone.0066476.
- Greninger, Alexander L., Eunice C. Chen, Taylor Sittler, Alex Scheinerman, Nareg Roubinian, Guixia Yu, Edward Kim, et al. 2010. "A Metagenomic Analysis of Pandemic Influenza A (2009 H1N1) Infection in Patients from North America". *PLoS ONE* 5 (10). doi:10.1371/journal.pone.0013381.
- Greninger, Alexander L., Samia N. Naccache, Kevin Messacar, Anna Clayton, Guixia Yu, Sneha Somasekar, Scot Federman, et al. 2015. "A Novel Outbreak Enterovirus D68 Strain Associated with Acute Flaccid Myelitis Cases in the USA (2012-14): A Retrospective Cohort Study". *The Lancet. Infectious Diseases* 15 (6): 671–82. doi:10.1016/S1473-3099(15)70093-9.
- Grover, Corrinne E., Armel Salmon, y Jonathan F. Wendel. 2012. "Targeted Sequence Capture as a Powerful Tool for Evolutionary Analysis". *American Journal of Botany* 99 (2): 312–19. doi:10.3732/ajb.1100323.
- Gurda, Brittney L., Kristin N. Parent, Heather Bladec, Robert S. Sinkovits, Michael A. DiMattia, Chelsea Rence, Alejandro Castro, et al. 2010. "Human Bocavirus Capsid Structure: Insights into the Structural Repertoire of the Parvoviridae". *Journal of Virology* 84 (12): 5880–89. doi:10.1128/JVI.02719-09.
- Hammond, Sarah P., Lisa S. Gagne, Shannon R. Stock, Francisco M. Marty, Rebecca S. Gelman, Wayne A. Marasco, Mark A. Poritz, y Lindsey R. Baden. 2012. "Respiratory Virus Detection in Immunocompromised Patients with FilmArray Respiratory Panel Compared to Conventional Methods". *Journal of Clinical Microbiology* 50 (10): 3216–21. doi:10.1128/JCM.00538-12.
- Hanshaw, Alissa S, Marie E Jetté, y Susan L Thibeault. 2014. "Characterization and comparison of bacterial communities in benign vocal fold lesions". *Microbiome* 2 (diciembre). doi:10.1186/2049-2618-2-43.
- Hauser, Philippe M., Thomas Bernard, Gilbert Greub, Katia Jatton, Marco Pagni, y Gaudenz M. Hafen. 2014. "Microbiota Present in Cystic Fibrosis Lungs as Revealed by Whole Genome Sequencing". *PLoS ONE* 9 (3). doi:10.1371/journal.pone.0090934.
- Heikkinen, Terho, y Asko Järvinen. 2003. "The Common Cold". *The Lancet* 361 (9351): 51–59. doi:10.1016/S0140-6736(03)12162-9.
- Hilty, Markus, Conor Burke, Helder Pedro, Paul Cardenas, Andy Bush, Cara Bossley, Jane Davies, et al. 2010. "Disordered Microbial Communities in Asthmatic Airways". *PloS One* 5 (1): e8578. doi:10.1371/journal.pone.0008578.
- Hino, Shigeo, y Hironori Miyata. 2007. "Torque Teno Virus (TTV): Current Status". *Reviews in Medical Virology* 17 (1): 45–57. doi:10.1002/rmv.524.
- Holden, Matthew T. G., Edward J. Feil, Jodi A. Lindsay, Sharon J. Peacock, Nicholas P. J. Day, Mark C. Enright, Tim J. Foster, et al. 2004. "Complete genomes of two clinical *Staphylococcus aureus* strains: Evidence for the rapid evolution of virulence and drug resistance". *Proceedings of the National Academy of Sciences of the United States of America* 101 (26): 9786–91. doi:10.1073/pnas.0402521101.

- Hoskins, J., W. E. Alborn, J. Arnold, L. C. Blaszcak, S. Burgett, B. S. DeHoff, S. T. Estrem, et al. 2001. "Genome of the Bacterium *Streptococcus Pneumoniae* Strain R6". *Journal of Bacteriology* 183 (19): 5709–17. doi:10.1128/JB.183.19.5709-5717.2001.
- Huang, Yvonne J., Eugenia Kim, Michael J. Cox, Eoin L. Brodie, Ron Brown, Jeanine P. Wiener-Kronish, y Susan V. Lynch. 2010. "A Persistent and Diverse Airway Microbiota Present during Chronic Obstructive Pulmonary Disease Exacerbations". *OMICS : a Journal of Integrative Biology* 14 (1): 9–59. doi:10.1089/omi.2009.0100.
- Huang, Yvonne J., Sanjay Sethi, Timothy Murphy, Snehal Nariya, Homer A. Boushey, y Susan V. Lynch. 2014. "Airway Microbiome Dynamics in Exacerbations of Chronic Obstructive Pulmonary Disease". *Journal of Clinical Microbiology* 52 (8): 2813–23. doi:10.1128/JCM.00035-14.
- Huebinger, Ryan M., Ming-Mei Liu, Scot E. Dowd, Fernando A. Rivera-Chavez, John Boynton, Curtis Carey, Kenneth Hawkins, et al. 2013. "Examination with Next-Generation Sequencing Technology of the Bacterial Microbiota in Bronchoalveolar Lavage Samples after Traumatic Injury". *Surgical Infections* 14 (3): 275–82. doi:10.1089/sur.2012.095.
- Hui, David S., y Alimuddin Zumla. 2015. "Emerging Respiratory Tract Viral Infections". *Current Opinion in Pulmonary Medicine* 21 (3): 284–92. doi:10.1097/MCP.000000000000153.
- Illumina. 2015. "16S Metagenomic Sequencing Library Preparation". Consultado junio 28. http://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparati.html.
- Ishikawa, Morio, Somboon Tanasupawat, Kazuyuki Nakajima, Hajime Kanamori, Shihomi Ishizaki, Kayo Kodama, Akiko Okamoto-Kainuma, Yukimichi Koizumi, Yasushi Yamamoto, y Kazuhide Yamasato. 2009. "Alkalibacterium Thalassium Sp. Nov., Alkalibacterium Pelagium Sp. Nov., Alkalibacterium Putridalgicola Sp. Nov. and Alkalibacterium Kapii Sp. Nov., Slightly Halophilic and Alkaliphilic Marine Lactic Acid Bacteria Isolated from Marine Organisms and Salted Foods Collected in Japan and Thailand". *International Journal of Systematic and Evolutionary Microbiology* 59 (Pt 5): 1215–26. doi:10.1099/ij.s.0.65602-0.
- Jalali, Saakshi, Samantha Kohli, Chitra Latka, Sugandha Bhatia, Shamsudheen Karuthedath Vellarikal, Sridhar Sivasubbu, Vinod Scaria, y Srinivasan Ramachandran. 2015. "Screening Currency Notes for Microbial Pathogens and Antibiotic Resistance Genes Using a Shotgun Metagenomic Approach". *PLoS ONE* 10 (6). doi:10.1371/journal.pone.0128711.
- Jeukens, Julie, Brian Boyle, Irena Kukavica-Ibrulj, Myriam M. Ouellet, Shawn D. Aaron, Steve J. Charette, Joanne L. Fothergill, Nicholas P. Tucker, Craig Winstanley, y Roger C. Levesque. 2014. "Comparative Genomics of Isolates of a *Pseudomonas aeruginosa* Epidemic Strain Associated with Chronic Lung Infections of Cystic Fibrosis Patients". *PLoS ONE* 9 (2). doi:10.1371/journal.pone.0087611.
- Johnson, Coreen L., y James Versalovic. 2012. "The Human Microbiome and Its Potential Importance to Pediatrics". *Pediatrics* 129 (5): 950–60. doi:10.1542/peds.2011-2736.
- Karched, Maribasappa, Radhika G. Bhardwaj, y Sirkka E. Asikainen. 2015. "Coaggregation and Biofilm Growth of *Granulicatella* Spp. with *Fusobacterium Nucleatum* and *Aggregatibacter Actinomycetemcomitans*". *BMC Microbiology* 15: 114. doi:10.1186/s12866-015-0439-z.
- Kehl, S. C., K. J. Henrickson, W. Hua, y J. Fan. 2001. "Evaluation of the Hexaplex Assay for Detection of Respiratory Viruses in Children". *Journal of Clinical Microbiology* 39 (5): 1696–1701. doi:10.1128/JCM.39.5.1696-1701.2001.
- Kidd, Aaron. 2015. "Rash illness, commonly affecting children, spreading in Tokyo". *Stars and Stripes*. Consultado julio 13. <http://www.stripes.com/news/pacific/rash-illness-commonly-affecting-children-spreading-in-tokyo-1.357059>.

- Kiley, James P., y Elisabet V. Caler. 2014. "The Lung Microbiome. A New Frontier in Pulmonary Medicine". *Annals of the American Thoracic Society* 11 Suppl 1 (enero): S66–70. doi:10.1513/AnnalsATS.201308-285MG.
- Kiss, Gabriella, Jens M. Holl, Grant M. Williams, Eric Alonas, Daryll Vanover, Aaron W. Lifland, Manasa Gudheti, et al. 2014. "Structural Analysis of Respiratory Syncytial Virus Reveals the Position of M2-1 between the Matrix Protein and the Ribonucleoprotein Complex". *Journal of Virology* 88 (13): 7602–17. doi:10.1128/JVI.00256-14.
- Kraus-Haas, Martina, Martin Mielke, y Arne Simon. 2015. "[Update on outbreaks reported from neonatal intensive care units: *Serratia marcescens*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Pseudomonas aeruginosa*]". *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 58 (3): 308–22. doi:10.1007/s00103-014-2114-1.
- Lagier, J.-C., F. Armougom, M. Million, P. Hugon, I. Pagnier, C. Robert, F. Bittar, et al. 2012. "Microbial Culturomics: Paradigm Shift in the Human Gut Microbiome Study". *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* 18 (12): 1185–93. doi:10.1111/1469-0691.12023.
- Landa-Cardaña, Adriana, Jaime Morales-Romero, Rebeca García-Roman, Ana Georgina Cobián-Güemes, Ernesto Méndez, Cristina Ortiz-Leon, Felipe Pitalúa-Cortés, Silvia Ivonne Mora, y Hilda Montero. 2012. "Clinical Characteristics and Genetic Variability of Human Rhinovirus in Mexico". *Viruses* 4 (2): 200–210. doi:10.3390/v4020200.
- La Roche. 2015a. "MagNA Pure LC 2.0 Instrument". Consultado junio 27. <https://lifescience.roche.com/shop/products/magna-pure-lc-20-instrument>.
- La Roche. 2015b. "Transcriptor First Strand cDNA Synthesis Kit". Consultado junio 28. <https://lifescience.roche.com/shop/products/transcriptor-first-strand-cdna-synthesis-kit>.
- Lecuit, Marc, y Marc Eloit. 2013. "The Human Virome: New Tools and Concepts". *Trends in Microbiology* 21 (10): 510–15. doi:10.1016/j.tim.2013.07.001.
- Ledergerber, Christian, y Christophe Dessimoz. 2011. "Base-Calling for next-Generation Sequencing Platforms". *Briefings in Bioinformatics*, enero, bbq077. doi:10.1093/bib/bbq077.
- Lemon, Katherine P., Vanja Klepac-Ceraj, Hilary K. Schiffer, Eoin L. Brodie, Susan V. Lynch, y Roberto Kolter. 2010. "Comparative Analyses of the Bacterial Microbiota of the Human Nostril and Oropharynx". *mBio* 1 (3). doi:10.1128/mBio.00129-10.
- Liljeroos, Lassi, Magdalena Anna Krzyzaniak, Ari Helenius, y Sarah Jane Butcher. 2013. "Architecture of Respiratory Syncytial Virus Revealed by Electron Cryotomography". *Proceedings of the National Academy of Sciences of the United States of America* 110 (27): 11133–38. doi:10.1073/pnas.1309070110.
- Liu, Lin, Yinhu Li, Siliang Li, Ni Hu, Yimin He, Ray Pong, Danni Lin, Lihua Lu, y Maggie Law. 2012. "Comparison of next-Generation Sequencing Systems". *Journal of Biomedicine & Biotechnology* 2012: 251364. doi:10.1155/2012/251364.
- Li, Yan, Hongmei Guo, Xijun Wang, Yang Lu, Chunyu Yang, y Pishan Yang. 2015. "Coinfection with *Fusobacterium nucleatum* Can Enhance the Attachment and Invasion of *Porphyromonas gingivalis* or *Aggregatibacter actinomycetemcomitans* to Human Gingival Epithelial Cells". *Archives of Oral Biology* 60 (9): 1387–93. doi:10.1016/j.archoralbio.2015.06.017.
- Lobo, Ingrid. 2015. "Basic Local Alignment Search Tool (BLAST) | Learn Science at Scitable". Consultado julio 7. <http://www.nature.com/scitable/topicpage/basic-local-alignment-search-tool-blast-29096>.

- Lozupone, Catherine, Adela Cota-Gomez, Brent E. Palmer, Derek J. Linderman, Emily S. Charlson, Erica Sodergren, Makedonka Mitreva, et al. 2013. "Widespread Colonization of the Lung by *Tropheryma whipplei* in HIV Infection". *American Journal of Respiratory and Critical Care Medicine* 187 (10): 1110–17. doi:10.1164/rccm.201211-2145OC.
- Lüsebrink, Jessica, Felix Wittleben, Verena Schildgen, y Oliver Schildgen. 2009. "Human Bocavirus – Insights into a Newly Identified Respiratory Virus". *Viruses* 1 (1): 3–12. doi:10.3390/v1010003.
- Lynch, Susan V. 2014. "Viruses and Microbiome Alterations". *Annals of the American Thoracic Society* 11 Suppl 1 (enero): S57–60. doi:10.1513/AnnalsATS.201306-158MG.
- Lysholm, Fredrik, Anna Wetterbom, Cecilia Lindau, Hamid Darban, Annelie Bjerkner, Kristina Fahlander, A. Michael Lindberg, Bengt Persson, Tobias Allander, y Björn Andersson. 2012. "Characterization of the Viral Microbiome in Patients with Severe Lower Respiratory Tract Infections, Using Metagenomic Sequencing". *PLoS ONE* 7 (2). doi:10.1371/journal.pone.0030875.
- Madan, J. C., D. C. Koestler, B. A. Stanton, L. Davidson, L. A. Moulton, M. L. Housman, J. H. Moore, et al. 2012. "Serial Analysis of the Gut and Respiratory Microbiome in Cystic Fibrosis in Infancy: Interaction between Intestinal and Respiratory Tracts and Impact of Nutritional Exposures". *mBio* 3 (4). doi:10.1128/mBio.00251-12.
- Mardis, Elaine R. 2008. "Next-Generation DNA Sequencing Methods". *Annual Review of Genomics and Human Genetics* 9: 387–402. doi:10.1146/annurev.genom.9.081307.164359.
- Mardis, Elaine R. 2013. "Next-Generation Sequencing Platforms". *Annual Review of Analytical Chemistry (Palo Alto, Calif.)* 6: 287–303. doi:10.1146/annurev-anchem-062012-092628.
- Martín, Rebeca, Sylvie Miquel, Philippe Langella, y Luis G Bermúdez-Humarán. 2014. "The role of metagenomics in understanding the human microbiome in health and disease". *Virulence* 5 (3): 413–23. doi:10.4161/viru.27864.
- Méndez, Jose Antonio, Jesús Mateos, Alejandro Beceiro, María Lopez, María Tomás, Margarita Poza, y Germán Bou. 2015. "Quantitative Proteomic Analysis of Host-Pathogen Interactions: A Study of *Acinetobacter Baumannii* Responses to Host Airways". *BMC Genomics* 16: 422. doi:10.1186/s12864-015-1608-z.
- Molyneaux, Philip L., Patrick Mallia, Michael J. Cox, Joseph Footitt, Saffron A. G. Willis-Owen, Daniel Homola, Maria-Belen Trujillo-Torralbo, et al. 2013. "Outgrowth of the Bacterial Airway Microbiome after Rhinovirus Exacerbation of Chronic Obstructive Pulmonary Disease". *American Journal of Respiratory and Critical Care Medicine* 188 (10): 1224–31. doi:10.1164/rccm.201302-0341OC.
- Morris, Alison, James M. Beck, Patrick D. Schloss, Thomas B. Campbell, Kristina Crothers, Jeffrey L. Curtis, Sonia C. Flores, et al. 2013. "Comparison of the Respiratory Microbiome in Healthy Nonsmokers and Smokers". *American Journal of Respiratory and Critical Care Medicine* 187 (10): 1067–75. doi:10.1164/rccm.201210-1913OC.
- Naccache, Samia N., Scot Federman, Narayanan Veeraraghavan, Matei Zaharia, Deanna Lee, Erik Samayoa, Jerome Bouquet, et al. 2014. "A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples". *Genome Research* 24 (7): 1180–92. doi:10.1101/gr.171934.113.
- Nakamura, Shota, Cheng-Song Yang, Naomi Sakon, Mayo Ueda, Takahiro Tougan, Akifumi Yamashita, Naohisa Goto, et al. 2009. "Direct Metagenomic Detection of Viral Pathogens in Nasal and Fecal Specimens Using an Unbiased High-Throughput Sequencing Approach". *PLoS ONE* 4 (1). doi:10.1371/journal.pone.0004219.

- Namiki, Toshiaki, Tsuyoshi Hachiya, Hideaki Tanaka, y Yasubumi Sakakibara. 2012. "MetaVelvet: An Extension of Velvet Assembler to de Novo Metagenome Assembly from Short Sequence Reads". *Nucleic Acids Research* 40 (20): e155. doi:10.1093/nar/gks678.
- NCBI, Bordetella. 2015. "Bordetella pertussis". Consultado junio 21. http://www.ncbi.nlm.nih.gov/genome/1008?genome_assembly_id=170243.
- NCBI. 2015. "Moraxella catarrhalis". Consultado junio 21. <http://www.ncbi.nlm.nih.gov/genome/1232>.
- NCBI, Chlamydia. 2015. "Chlamydia pneumoniae". Consultado junio 21. <http://www.ncbi.nlm.nih.gov/genome/171>.
- NCBI, Coxiella. 2015. "Coxiella burnetii". Consultado junio 21. <http://www.ncbi.nlm.nih.gov/genome/543>.
- NCBI, Klebsiella. 2015. "Klebsiella pneumoniae". Consultado junio 21. <http://www.ncbi.nlm.nih.gov/genome/815>.
- NCBI, Mycoplasma. 2015. "Mycoplasma pneumoniae". Consultado junio 21. <http://www.ncbi.nlm.nih.gov/genome/1028>.
- NCBI, Streptococcus. 2015. "Streptococcus pneumoniae". *Streptococcus pneumoniae (ID 176) - Genome - NCBI*. Consultado junio 21. <http://www.ncbi.nlm.nih.gov/genome/176>.
- "Nextera XT DNA Sample Preparation Guide.pdf". 2015. Consultado junio 28. http://www.liai.org/files/nextera_xt_sample_preparation_guide_15031942_c.pdf.
- Nguyen, K. T., E. J. Hansen, y M. A. Farinha. 1999. "Construction of a Genomic Map of Moraxella (Branhamella) Catarrhalis ATCC 25238 and Physical Mapping of Virulence-Associated Genes". *Canadian Journal of Microbiology* 45 (4): 299–303.
- Oh, Julia, Allyson L. Byrd, Clay Deming, Sean Conlan, NISC Comparative Sequencing Program, Heidi H. Kong, y Julia A. Segre. 2014. "Biogeography and Individuality Shape Function in the Human Skin Metagenome". *Nature* 514 (7520): 59–64. doi:10.1038/nature13786.
- Orozco Sanchez, Carmen Olivia, y Joel Armando Vázquez Pérez. 2012. "Detección e identificación de virus de influenza estacional, A/H1N1 pandémica, así como de diferentes virus respiratorios mediante amplificación múltiple en tiempo real empleando la tecnología de arreglos dinámicos de nanofluidos". México. D. F.: Universidad Nacional Autónoma de México. Facultad de Química. http://oreon.dgbiblio.unam.mx/F/8RIS23745AMFD56DUEMIUAAMRV7PB8YP8BIGYNNDV NDBDKCB28-67794?func=find-b&request=Carmen+Olivia+Orozco+Sanchez+&find_code=WRD&adjacent=N&local_base=TES01&x=0&y=0&filter_code_2=WYR&filter_request_2=&filter_code_3=WYR&filter_request_3=.
- Ostos Ortiz, Olga Lucia, y Ruth Mélida Sánchez. 2003. "Chlamydia trachomatis: avances y perspectivas" 1 (1): 1–116.
- Panda, Swagatika, Nirmal Kumar Mohakud, Lindomar Pena, y Subrat Kumar. 2014. "Human Metapneumovirus: Review of an Important Respiratory Pathogen". *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases* 25 (agosto): 45–52. doi:10.1016/j.ijid.2014.03.1394.
- Park, O.-J., H. Yi, J. H. Jeon, S.-S. Kang, K.-T. Koo, K.-Y. Kum, J. Chun, C.-H. Yun, y S. H. Han. 2015. "Pyrosequencing Analysis of Subgingival Microbiota in Distinct Periodontal Conditions". *Journal of Dental Research* 94 (7): 921–27. doi:10.1177/0022034515583531.
- Payne, Brendan, y Richard Bellamy. 2014. "Novel Respiratory Viruses: What Should the Clinician Be Alert For?" *Clinical Medicine (London, England)* 14 Suppl 6 (diciembre): s12–16. doi:10.7861/clinmedicine.14-6-s12.

- PHAC. 2015. "Moraxella". Consultado junio 21. <http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/moraxella-eng.php#endnote2>.
- PHAC, Bordetella. 2015. "Bordetella pertussis". Consultado junio 21. <http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/bordetella-pertussis-eng.php>.
- PHAC, Klebsiella. 2015. "Klebsiella spp. - Pathogen Safety Data Sheets". Consultado junio 21. <http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/klebsiella-eng.php>.
- Picelli, Simone, Åsa K. Björklund, Björn Reinius, Sven Sagasser, Gösta Winberg, y Rickard Sandberg. 2014. "Tn5 Transposase and Tagmentation Procedures for Massively Scaled Sequencing Projects". *Genome Research* 24 (12): 2033–40. doi:10.1101/gr.177881.114.
- Prachayangprecha, Slinporn, Claudia M. E. Schapendonk, Marion P. Koopmans, Albert D. M. E. Osterhaus, Anita C. Schürch, Suzan D. Pas, Annemiek A. van der Eijk, Yong Poovorawan, Bart L. Haagmans, y Saskia L. Smits. 2014. "Exploring the Potential of Next-Generation Sequencing in Detection of Respiratory Viruses". *Journal of Clinical Microbiology* 52 (10): 3722–30. doi:10.1128/JCM.01641-14.
- Pragman, Alexa A., Hyeun Bum Kim, Cavan S. Reilly, Christine Wendt, y Richard E. Isaacson. 2012. "The Lung Microbiome in Moderate and Severe Chronic Obstructive Pulmonary Disease". *PLoS ONE* 7 (10): e47305. doi:10.1371/journal.pone.0047305.
- Pride, David T., Julia Salzman, Matthew Haynes, Forest Rohwer, Clara Davis-Long, Richard A. White, Peter Loomer, Gary C. Armitage, y David A. Relman. 2012. "Evidence of a Robust Resident Bacteriophage Population Revealed through Analysis of the Human Salivary Virome". *The ISME Journal* 6 (5): 915–26. doi:10.1038/ismej.2011.169.
- QIAGEN. 2015. "QIAamp cador Pathogen Mini Handbook - QIAGEN". Consultado junio 27. <https://www.qiagen.com/kr/resources/resourcedetail?id=dd291a2f-f2ff-4768-ad97-8e2462fe58c0&lang=en>.
- Quiñones-Mateu, Miguel E., Santiago Avila, Gustavo Reyes-Teran, y Miguel A. Martinez. 2014. "Deep Sequencing: Becoming a Critical Tool in Clinical Virology". *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology* 61 (1): 9–19. doi:10.1016/j.jcv.2014.06.013.
- Ramakrishnan, Vijay R., Leah M. Feazel, Sarah A. Gitomer, Diana Ir, Charles E. Robertson, y Daniel N. Frank. 2013. "The Microbiome of the Middle Meatus in Healthy Adults". *PLoS ONE* 8 (12). doi:10.1371/journal.pone.0085507.
- "Reference Genome". 2015. Consultado julio 7. ftp://ftp.ncbi.nlm.nih.gov/genbank/genomes/Eukaryotes/vertebrates_mammals/Homo_sapiens/GRCh38/.
- Renwick, Julie, Paul McNally, Bettina John, Todd DeSantis, Barry Linnane, y Philip Murphy. 2014. "The Microbial Community of the Cystic Fibrosis Airway Is Disrupted in Early Life". *PLoS ONE* 9 (12). doi:10.1371/journal.pone.0109798.
- Rinke, Christian, Patrick Schwientek, Alexander Sczyrba, Natalia N. Ivanova, Iain J. Anderson, Jan-Fang Cheng, Aaron Darling, et al. 2013. "Insights into the Phylogeny and Coding Potential of Microbial Dark Matter". *Nature* 499 (7459): 431–37. doi:10.1038/nature12352.
- Roca, B. 2007. "Fiebre Q". *Anales de Medicina Interna* 24 (11): 558–60.
- Roth, Stina B., Jari Jalava, Olli Ruuskanen, Aino Ruohola, y Simo Nikkari. 2004. "Use of an Oligonucleotide Array for Laboratory Diagnosis of Bacteria Responsible for Acute Upper Respiratory Infections". *Journal of Clinical Microbiology* 42 (9): 4268–74. doi:10.1128/JCM.42.9.4268-4274.2004.

- Roux, Simon, Michaël Faubladié, Antoine Mahul, Nils Paulhe, Aurélien Bernard, Didier Debroas, y François Enault. 2011. "Metavir: A Web Server Dedicated to Virome Analysis". *Bioinformatics (Oxford, England)* 27 (21): 3074–75. doi:10.1093/bioinformatics/btr519.
- "SAMtools". 2015. Consultado julio 7. <http://samtools.sourceforge.net/>.
- "samtools(1) manual page". 2015. Consultado julio 7. <http://www.htslib.org/doc/samtools.html>.
- Sánchez, Nallely Concepción Juárez, y Arturo Calderon Vega. 2013. "Revisión y actualización de la literatura sobre legionella pneumophila". http://oreon.dgbiblio.unam.mx/F/B5FD14JECD4XNB4UA1FAEUYNPKCTFUIEGXLGCH6U8M2FI4IIX-29488?func=full-set-set&set_number=004195&set_entry=000001&format=999.
- Sanger Institute. 2015a. "Haemophilus - Wellcome Trust Sanger Institute". Consultado junio 21. http://www.sanger.ac.uk/resources/downloads/bacteria/haemophilus.html#t_2.
- Sanger Institute. 2015b. "SMALT - Wellcome Trust Sanger Institute". Consultado julio 7. <https://www.sanger.ac.uk/resources/software/smalt/>.
- Schildgen, Oliver. 2013. "Human Bocavirus: Lessons Learned to Date". *Pathogens* 2 (1): 1–12. doi:10.3390/pathogens2010001.
- Schomacker, Henrick, Anne Schaap-Nutt, Peter L. Collins, y Alexander C. Schmidt. 2012. "Pathogenesis of acute respiratory illness caused by human parainfluenza viruses". *Current opinion in virology* 2 (3): 294–99. doi:10.1016/j.coviro.2012.02.001.
- Segal, Leopoldo N, Alexander V Alekseyenko, Jose C Clemente, Rohan Kulkarni, Benjamin Wu, Hao Chen, Kenneth I Berger, et al. 2013. "Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation". *Microbiome* 1 (julio): 19. doi:10.1186/2049-2618-1-19.
- "SINAVE.pdf". 2015. Consultado junio 27. <http://www.ssm.gob.mx/portal/page/influenza/LINEAMIENTOS%20DE%20INFLUENZA.pdf>.
- Smet, Birgit De, Mark Mayo, Charlotte Peeters, James E. A. Zlosnik, Theodore Spilker, Trevor J. Hird, John J. LiPuma, et al. 2015. "Burkholderia Stagnalis Sp. Nov. and Burkholderia Territorii Sp. Nov., Two Novel Burkholderia Cepacia Complex Species from Environmental and Human Sources". *International Journal of Systematic and Evolutionary Microbiology*, abril. doi:10.1099/ijs.0.000251.
- Smits, Saskia L., y Albert Dme Osterhaus. 2013. "Virus Discovery: One Step beyond". *Current Opinion in Virology*, abril. doi:10.1016/j.coviro.2013.03.007.
- "SSA". 2015. Consultado junio 16. <http://www.epidemiologia.salud.gob.mx/doctos/boletin/2013/semanas/sem38/pdf/edit3813.pdf>.
- "Staphylococcus.pdf". 2015. Consultado junio 21. <http://www.higiene.edu.uy/cefa/2008/Staphylococcus.pdf>.
- Stressmann, Franziska A., Geraint B. Rogers, Erich R. Klem, Andrew K. Lilley, Scott H. Donaldson, Thomas W. Daniels, Mary P. Carroll, et al. 2011. "Analysis of the Bacterial Communities Present in Lungs of Patients with Cystic Fibrosis from American and British Centers". *Journal of Clinical Microbiology* 49 (1): 281–91. doi:10.1128/JCM.01650-10.
- Sze, Marc A., Pedro A. Dimitriu, Shizu Hayashi, W. Mark Elliott, John E. McDonough, John V. Gosselink, Joel Cooper, Don D. Sin, William W. Mohn, y James C. Hogg. 2012. "The Lung Tissue Microbiome in Chronic Obstructive Pulmonary Disease". *American Journal of Respiratory and Critical Care Medicine* 185 (10): 1073–80. doi:10.1164/rccm.201111-2075OC.

- Taboada, Blanca, Marco A. Espinoza, Pavel Isa, Fernando E. Aponte, María A. Arias-Ortiz, Jesús Monge-Martínez, Rubén Rodríguez-Vázquez, et al. 2014. "Is There Still Room for Novel Viral Pathogens in Pediatric Respiratory Tract Infections?" Editado por Amit Kapoor. *PLoS ONE* 9 (11): e113570. doi:10.1371/journal.pone.0113570.
- Tanaka, Shoji, Mikako Yoshida, Yukio Murakami, Takako Ogiwara, Masao Shoji, Satoko Kobayashi, Sigeru Watanabe, Mamoru Machino, y Seiichiro Fujisawa. 2008. "The Relationship of *Prevotella Intermedia*, *Prevotella Nigrescens* and *Prevotella Melaninogenica* in the Supragingival Plaque of Children, Caries and Oral Malodor". *The Journal of Clinical Pediatric Dentistry* 32 (3): 195–200.
- Tariq, Mohammad A., Francesca L. C. Everest, Lauren A. Cowley, Anthony De Soyza, Giles S. Holt, Simon H. Bridge, Audrey Perry, et al. 2015. "A metagenomic approach to characterize temperate bacteriophage populations from Cystic Fibrosis and non-Cystic Fibrosis bronchiectasis patients". *Frontiers in Microbiology* 6 (febrero). doi:10.3389/fmicb.2015.00097.
- te Velthuis, Aartjan J. W., Jamie J. Arnold, Craig E. Cameron, Sjoerd H. E. van den Worm, y Eric J. Snijder. 2010. "The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent". *Nucleic Acids Research* 38 (1): 203–14. doi:10.1093/nar/gkp904.
- Thermo Fisher Scientific. 2015a. "Applied Biosystems® Veriti® Thermal Cycler". Consultado junio 27. <https://www.lifetechnologies.com/mx/es/home/life-science/pcr/thermal-cyclers-realtime-instruments/thermal-cyclers/veriti-thermal-cycler.html>.
- Thermo Fisher Scientific. 2015b. "DNase I, RNase-free (supplied with MnCl₂) (1 U/L) - Life Technologies". Consultado junio 27. <https://www.lifetechnologies.com/order/catalog/product/EN0525>.
- Thermo Fisher Scientific. 2015c. "NanoDrop Products - Spectrophotometers and Fluorospectrometers - www.nanodrop.com". Consultado junio 28. <http://www.nanodrop.com/>.
- Thermo Fisher Scientific. 2015d. "Platinum® DNA Polymerases". Consultado junio 28. <http://www.lifetechnologies.com/mx/es/home/brands/invitrogen/platinum-dna-polymerase.html>.
- Thermo Fisher Scientific. 2015e. "PureLink Viral RNA/DNA Mini Kit - Life Technologies". Consultado junio 27. <http://www.lifetechnologies.com/order/catalog/product/12280050>.
- Thermo Fisher Scientific. 2015f. "Qubit® 3.0 Fluorometer". Consultado junio 28. <http://www.lifetechnologies.com/mx/es/home/life-science/laboratory-instruments/fluorometers/qubit/qubit-fluorometer.html.html>.
- Tran, Thomas, Bui Trong Chien, Georgina Papadakis, Julian Druce, Chris Birch, Doris Chibo, Truong Phuoc An, et al. 2012. "Respiratory Virus Laboratory Pandemic Planning and Surveillance in Central Viet Nam, 2008-2010". *Western Pacific Surveillance and Response Journal: WPSAR* 3 (3): 49–56. doi:10.5365/WPSAR.2012.3.2.001.
- Tregoning, John S., y Jürgen Schwarze. 2010. "Respiratory Viral Infections in Infants: Causes, Clinical Symptoms, Virology, and Immunology". *Clinical Microbiology Reviews* 23 (1): 74–98. doi:10.1128/CMR.00032-09.
- Tuccio, Alessandra, Giacomo Zanelli, Dhyana Chiara Rodriguez, Maria Luisa Tataranno, Marina Vascotto, y Paolo Balestri. 2014. "Petechial Rash Associated with Parvovirus B19 in Children: Case Report and Literature Review". *Le Infezioni in Medicina: Rivista Periodica Di Etiologia, Epidemiologia, Diagnostica, Clinica E Terapia Delle Patologie Infettive* 22 (3): 250–54.

- Tuthill, Tobias J., Elisabetta Gropelli, James M. Hogle, y David J. Rowlands. 2010. "Picornaviruses". *Current topics in microbiology and immunology* 343: 43–89. doi:10.1007/82_2010_37.
- Twomey, Kate B., Mark Alston, Shi-Qi An, Oisín J. O’Connell, Yvonne McCarthy, David Swarbreck, Melanie Febrer, J. Maxwell Dow, Barry J. Plant, y Robert P. Ryan. 2013. "Microbiota and Metabolite Profiling Reveal Specific Alterations in Bacterial Community Structure and Environment in the Cystic Fibrosis Airway during Exacerbation". *PLoS ONE* 8 (12). doi:10.1371/journal.pone.0082432.
- Ursell, Luke K., William Van Treuren, Jessica L. Metcalf, Meg Pirrung, Andrew Gewirtz, y Rob Knight. 2013. "Replenishing Our Defensive Microbes". *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 35 (9): 810–17. doi:10.1002/bies.201300018.
- van den Bergh, Menno R., Giske Biesbroek, John W. A. Rossen, Wouter A. A. de Steenhuijsen Pipers, Astrid A. T. M. Bosch, Elske J. M. van Gils, Xinhui Wang, et al. 2012. "Associations between Pathogens in the Upper Respiratory Tract of Young Children: Interplay between Viruses and Bacteria". *PLoS One* 7 (10): e47711. doi:10.1371/journal.pone.0047711.
- van der Schaar, Hilde M., Pieter Leyssen, Hendrik J. Thibaut, Armando de Palma, Lonneke van der Linden, Kjerstin H. W. Lanke, Céline Lacroix, et al. 2013. "A Novel, Broad-Spectrum Inhibitor of Enterovirus Replication That Targets Host Cell Factor Phosphatidylinositol 4-Kinase IIIβ". *Antimicrobial Agents and Chemotherapy* 57 (10): 4971–81. doi:10.1128/AAC.01175-13.
- Venkataraman, Arvind, Christine M. Bassis, James M. Beck, Vincent B. Young, Jeffrey L. Curtis, Gary B. Huffnagle, y Thomas M. Schmidt. 2015. "Application of a Neutral Community Model To Assess Structuring of the Human Lung Microbiome". *mBio* 6 (1). doi:10.1128/mBio.02284-14.
- Voronina, Olga L., Marina S. Kunda, Natalia N. Ryzhova, Ekaterina I. Aksenova, Andrey N. Semenov, Anna V. Lasareva, Elena L. Amelina, Alexandr G. Chuchalin, Vladimir G. Lunin, y Alexandr L. Gintsburg. 2015. "The Variability of the Order Burkholderiales Representatives in the Healthcare Units". *BioMed Research International* 2015: 680210. doi:10.1155/2015/680210.
- Wang, Ruixue, y Jeffery K. Taubenberger. 2010. "Methods for Molecular Surveillance of Influenza". *Expert Review of Anti-Infective Therapy* 8 (5): 517–27. doi:10.1586/eri.10.24.
- Wenner, Melinda. 2008. "Virology: The Battle within". *Nature* 451 (7177): 388–89. doi:10.1038/451388a.
- WHO. 2015a. "WHO | Battle against Respiratory Viruses (BRaVe) initiative". WHO. Consultado junio 6. http://www.who.int/influenza/patient_care/clinical/brave/en/.
- WHO. 2015b. "WHO | The top 10 causes of death". WHO. Consultado junio 16. <http://www.who.int/mediacentre/factsheets/fs310/en/>.
- Willner, Dana, Mike Furlan, Matthew Haynes, Robert Schmieder, Florent E. Angly, Joas Silva, Sassan Tammadoni, Bahador Nosrat, Douglas Conrad, y Forest Rohwer. 2009. "Metagenomic Analysis of Respiratory Tract DNA Viral Communities in Cystic Fibrosis and Non-Cystic Fibrosis Individuals". *PLoS ONE* 4 (10). doi:10.1371/journal.pone.0007370.
- Woo, Patrick C. Y., Yi Huang, Susanna K. P. Lau, y Kwok-Yung Yuen. 2010. "Coronavirus Genomics and Bioinformatics Analysis". *Viruses* 2 (8): 1804–20. doi:10.3390/v2081803.
- Wu, Jing, Wei Liu, Lei He, Fuli Huang, Jiazhen Chen, Peng Cui, Yaojie Shen, et al. 2013. "Sputum Microbiota Associated with New, Recurrent and Treatment Failure Tuberculosis". *PLoS ONE* 8 (12). doi:10.1371/journal.pone.0083445.

- Wylie, Kristine M., Kathie A. Mihindukulasuriya, Erica Sodergren, George M. Weinstock, y Gregory A. Storch. 2012. "Sequence Analysis of the Human Virome in Febrile and Afebrile Children". *PLoS ONE* 7 (6). doi:10.1371/journal.pone.0027735.
- Wylie, Kristine M., George M. Weinstock, y Gregory A. Storch. 2012. "Emerging View of the Human Virome". *Translational Research: The Journal of Laboratory and Clinical Medicine* 160 (4): 283–90. doi:10.1016/j.trsl.2012.03.006.
- Xiaoyan, Li, Chen Jinying, Kong Mei, Su Xu, Zou Ming, y Guo Liru. 2014. "Development of High-Throughput Liquid Chips for Respiratory Virus Detection". *Clinical Laboratory* 60 (3): 347–61.
- Yozwiak, Nathan L., Peter Skewes-Cox, Aubree Gordon, Saira Saborio, Guillermina Kuan, Angel Balmaseda, Don Ganem, Eva Harris, y Joseph L. DeRisi. 2010. "Human Enterovirus 109: a Novel Interspecies Recombinant Enterovirus Isolated from a Case of Acute Pediatric Respiratory Illness in Nicaragua". *Journal of Virology* 84 (18): 9047–58. doi:10.1128/JVI.00698-10.
- Zakharkina, Tetyana, Elke Heinzl, Rembert A. Koczulla, Timm Greulich, Katharina Rentz, Josch K. Pauling, Jan Baumbach, et al. 2013. "Analysis of the Airway Microbiota of Healthy Individuals and Patients with Chronic Obstructive Pulmonary Disease by T-RFLP and Clone Sequencing". *PLoS ONE* 8 (7). doi:10.1371/journal.pone.0068302.
- Zaki, Ali M., Sander van Boheemen, Theo M. Bestebroer, Albert D. M. E. Osterhaus, y Ron A. M. Fouchier. 2012. "Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia". *The New England Journal of Medicine* 367 (19): 1814–20. doi:10.1056/NEJMoa1211721.
- Zambon, Maria. 2014. "Influenza and Other Emerging Respiratory Viruses". *Medicine* 42 (1): 45–51. doi:10.1016/j.mpmed.2013.10.017.
- Zemanick, Edith T., J. Kirk Harris, Brandie D. Wagner, Charles E. Robertson, Scott D. Sagel, Mark J. Stevens, Frank J. Accurso, y Theresa A. Laguna. 2013. "Inflammation and Airway Microbiota during Cystic Fibrosis Pulmonary Exacerbations". *PLoS ONE* 8 (4). doi:10.1371/journal.pone.0062917.
- Zerbino, Daniel R., y Ewan Birney. 2008. "Velvet: Algorithms for de Novo Short Read Assembly Using de Bruijn Graphs". *Genome Research* 18 (5): 821–29. doi:10.1101/gr.074492.107.
- Zhao, Jiangchao, Patrick D. Schloss, Linda M. Kalikin, Lisa A. Carmody, Bridget K. Foster, Joseph F. Petrosino, James D. Cavalcoli, et al. 2012. "Decade-long bacterial community dynamics in cystic fibrosis airways". *Proceedings of the National Academy of Sciences of the United States of America* 109 (15): 5809–14. doi:10.1073/pnas.1120577109.
- Zoll, Jan, Janette Rahamat-Langendoen, Inge Ahout, Marien I. de Jonge, Jop Jans, Martijn A. Huijnen, Gerben Ferwerda, Adilia Warris, y Willem J. G. Melchers. 2015. "Direct Multiplexed Whole Genome Sequencing of Respiratory Tract Samples Reveals Full Viral Genomic Information". *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology* 66 (mayo): 6–11. doi:10.1016/j.jcv.2015.02.010.
- Zumla, Alimuddin, Jaffar A. Al-Tawfiq, Virve I. Enne, Mike Kidd, Christian Drosten, Judy Breuer, Marcel A. Muller, et al. 2014. "Rapid Point of Care Diagnostic Tests for Viral and Bacterial Respiratory Tract Infections—needs, Advances, and Future Prospects". *The Lancet Infectious Diseases* 14 (11): 1123–35. doi:10.1016/S1473-3099(14)70827-8.

XII. ANEXO

Tabla A1. Mezcla de reacción 1, para hexámeros (izquierda) y hexámeros unidos a M13 (derecha).

Hexámeros del ensayo comercial		Hexámeros unidos a M13	
Componente	(μL)	Componente	(μL)
Pool (RNA)	10	Pool (RNA)	10
Oligonucleótido	2	Oligonucleótido	2
Hexámero		Hexámero unido a M13	
Agua	1	Agua	1
Total	13	Total	13

Tabla A2.- Mezcla de reacción 2, donde se observa la concentración y volumen de cada componente de la reacción para realizar la retro transcripción.

Componente	Concentración	Volumen (μL)
Mezcla de reacción 1, con Hexámeros y hexámeros M13, respectivamente	--	13
Buffer transcriptor	5x	4
Transcriptor	20 U/ μL	0.5
Inhibidor de RNAsa	40 U/ μL	0.5
dNTPs	10 mM	2
Total	--	20

Tabla A3.- Mezcla de reacción para síntesis de dsDNA con fragmento klenow.

Componente	Concentración	Volumen (μL)
Buffer	10x	2.5
dNTPs	10mM	1
M13 random	20 μM	1
Agua	--	15.5
cDNA/DNA viral	--	5
Total	--	25

Tabla A4. Mezcla de reacción para la amplificación de las secuencias M13

Componente	Concentración	Volumen (μL)
Buffer	10x	2.5
M13 (sin hexamero)	20 μM	0.25
M13 random	20 μM	0.25
Sulfato de magnesio	50mM	0.75
dNTPs	10mM	0.5
Agua	--	15.5
Taq	5 U/ μL	0.25
dsDNA	--	5
Total	--	25

Tabla A5.- Mezcla de reacción para preparar productos de amplificación de 16S ribosomal.

Componente	Concentración	Volumen (μL)
Buffer de PCR sin Mg	10X	2.5
dNTPs	10 mM	0.5
MgCl ₂	50 mM	0.75
Oligonucleotido 1	1 μM	5
Oligonucleotido 2	1 μM	5
Platinum Taq		0.2
DNA bacteria	5 ng	2.5
H ₂ O	--	7.55
Volumen final	--	25