



**ESPE**  
UNIVERSIDAD DE LAS FUERZAS ARMADAS  
INNOVACIÓN PARA LA EXCELENCIA

**DEPARTAMENTO DE CIENCIAS DE LA VIDA Y DE LA  
AGRICULTURA**

**CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA**

**TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL  
TÍTULO DE INGENIERA EN BIOTECNOLOGÍA**

**TEMA: “DETECCIÓN E IDENTIFICACIÓN DE VIRUS  
FITOPATÓGENOS EN CULTIVOS DE NARANJILLA  
(*Solanum quitoense*) Y LIMÓN MEYER (*Citrus x meyeri*),  
MEDIANTE SECUENCIACIÓN MASIVA PARALELA.”**

**AUTOR: RAMOS RODRÍGUEZ, KARLA PAOLA**

**DIRECTOR: FLORES FLOR, FRANCISCO JAVIER PhD.**

**SANGOLQUÍ**

**2019**



**DEPARTAMENTO DE CIENCIAS DE LA VIDA Y AGRICULTURA**  
**CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA**

**CERTIFICACIÓN**

Certifico que el trabajo de titulación, “**DETECCIÓN E IDENTIFICACIÓN DE VIRUS FITOPATÓGENOS EN CULTIVOS DE NARANJILLA (*Solanum quitoense*) Y LIMÓN MEYER (*Citrus x meyeri*), MEDIANTE SECUENCIACIÓN MASIVA PARALELA**” fue realizado por la señorita *Ramos Rodríguez Karla Paola* el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

**Sangolquí, junio del 2019.**

  
.....  
**FRANCISCO JAVIER FLORES FLOR, PhD**

**DIRECTOR**



**DEPARTAMENTO DE CIENCIAS DE LA VIDA Y AGRICULTURA**  
**CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA**

**AUTORÍA DE RESPONSABILIDAD**

Yo, **Ramos Rodríguez Karla Paola**, declaro que el contenido, ideas y criterios del trabajo de titulación: *Detección e identificación de virus fitopatógenos en cultivos de Naranja (Solanum quitoense) y Limón Meyer (Citrus x meyeri), mediante secuenciación masiva paralela* es de mi autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente el contenido de la investigación mencionada es veraz.

**Sangolquí, junio del 2019.**

**Karla Paola Ramos Rodríguez**

C.C.: 172133793-7



**DEPARTAMENTO DE CIENCIAS DE LA VIDA Y AGRICULTURA**

**CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA**

**AUTORIZACIÓN**

Yo, **Ramos Rodríguez Karla Paola**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: *Detección e identificación de virus fitopatógenos en cultivos de Naranjilla (Solanum quitoense) y Limón Meyer (Citrus x meyeri), mediante secuenciación masiva paralela* en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

**Sangolquí, junio del 2019.**



**Karla Paola Ramos Rodríguez**

C.C.: 172133793-7

## DEDICATORIA

A Dios por darme la vida y guiar cada uno de mis pasos, por la sabiduría que me ha sabido dar para enfrentar las adversidades y por ser mi fortaleza siempre.

A mis padres Carlos y Violeta por su amor y apoyo incondicional que me han brindado en cada momento a pesar de los tropiezos y dificultades, por su confianza y comprensión a lo largo de todos estos años, y sobre todo por el esfuerzo que han hecho para ayudarme a cumplir mis sueños y metas.

A mis hermanos Gisse y John por su ser mi ejemplo, por todo el cariño, apoyo, preocupación y sobretodo paciencia que han sabido tenerme todos estos años, en especial durante mi vida universitaria. A mi sobrino Emilio Santiago por ser la alegría y el motor de mi vida.

A mi novio Dany por su amor infinito, por ser mi compañero y mi apoyo incondicional día a día, por estar siempre a mi lado dándome ánimos para levantarme y seguir adelante.

A mis abuelitos Ángel (+), Victoria (+), Sixto (+) y Rosario (+) mis ángeles que siempre me cuidan desde el cielo.

A mis compañeros fieles de cuatro patas Chikis (+), Toby (+), Max y Fede por su cariño y compañía en las noches de desvelo, por alegrar mi vida y aliviar con su ternura los días difíciles.

Este trabajo va dedicado a todos ustedes con todo el amor y cariño.

## **AGRADECIMIENTOS**

Agradezco a mis padres Carlos y Violeta por su preocupación, cariño, consejos, y ayuda, ustedes son mi ejemplo. Muchos de los logros que he conseguido a lo largo de mi vida se los debo a ustedes, incluyendo este, que a pesar que no ha sido fácil llegar a donde estoy nunca perdieron la fe en mí y me dieron su apoyo incondicional. Por eso y más mi eterno agradecimiento a ustedes.

Quiero Agradecer a mi hermana Gisse por apoyarme sin importar los tropiezos que he tenido y darme ánimos para continuar y nunca renunciar; por enseñarme a ser fuerte y salir siempre adelante a pesar de los problemas y adversidades que se pueden presentar, por estar siempre pendiente de mi a pesar de la distancia. Pero sobretodo, gracias por darme el mejor regalo de la vida, mi sobrino Emilio quien llegó alegrar mí vida. A mi hermano John que a pesar de las diferencias que tenemos, siempre ha estado a mi lado dándome su apoyo y ayuda, gracias por soportar mi mal genio y a pesar de lo difícil e insoportable que a veces puedo llegar a ser, nunca me has abandonado y has estado siempre pendiente todos estos años. Gracias hermanos, gracias mis “ramitos” por su cariño, ejemplo y apoyo infinito, son los mejores hermanos mayores que Dios me pudo dar.

Mi gratitud y agradecimiento al Grupo de Investigación en Microbiología y Ambiente (GIMA) de la carrera de Ingeniería en Biotecnología de la Universidad de las Fuerzas Armadas- ESPE por facilitarme los equipos y reactivos usados en mi

investigación. A la MSc. Alma Koch por abrirme las puertas del Laboratorio de Microbiología Molecular y permitirme trabajar en él, por inspirarme e inculcar en mí el amor por la ciencia y la microbiología y por toda la ayuda y apoyo incondicional brindado en mis años de formación. Al laboratorio ID-GEN por facilitarme el uso de sus instalaciones y equipos. A mi director el Dr. Francisco Flores por darme la oportunidad de trabajar bajo su tutela, por confiar en mi trabajo y brindarme su ayuda para resolver los inconvenientes presentados durante la investigación, por enseñarme que hay retos en la vida que por más difíciles que parezcan todo es posible con esfuerzo y sacrificio.

Mi más sincero agradecimiento al Programa de Fruticultura del Instituto Nacional de Investigaciones Agropecuarias (INIAP) - Granja Experimental Tumbaco, de manera especial al Ing. William Viera y al Proyecto “Biocontrol for Sustainable Farming Systems, Ecuador” del Ministerio de Asuntos Exteriores y Comercio (MFAT) de Nueva Zelanda por haberme dado la oportunidad, confianza, apoyo y facilidades necesarias para la ejecución de esta investigación.

Agradezco a la Ing. Fiama Guevara por su tiempo, paciencia, apoyo y consejos; por compartir sus conocimientos y principalmente por TODA la ayuda brindada dentro y fuera del laboratorio. A Jesy por haber sido una excelente compañera pero sobretodo amiga todos estos años, por haberme tenido paciencia y soportar mis momentos de estrés e histeria, por su apoyo en los instantes de desesperación. Gracias chicas por su amistad y apoyo, por esas conversaciones de

bus, los almuerzos compartidos, por su ayuda en los muestreos y por todas las locuras y experiencias vividas; definitivamente sin ustedes esto no hubiese sido posible.

A Dany, mi compañero y amor de mi vida, quien es y ha sido mi apoyo incondicional cada día, especialmente en los momentos de decline y cansancio. Gracias por tu paciencia, comprensión y por todo lo que hemos compartido juntos; eres mi inspiración y mi motivación.

También quiero agradecer a Ari, Belén, Dany y Alvaro mis compañeros y amigos de laboratorio por los momentos divertidos que compartimos y por su apoyo brindado dentro y fuera del laboratorio.

A mis amigos por confiar y creer en mí, por todos los momentos compartidos y por haber hecho de mi etapa universitaria un trayecto de vivencia que nunca olvidaré.

Gracias a Dios y a la vida por este nuevo triunfo y meta cumplida, gracias a todas las personas que me apoyaron directa e indirectamente y creyeron en la realización de esta tesis.



## ÍNDICE DE CONTENIDOS

DEDICATORIA.....	iv
AGRADECIMIENTOS.....	v
ÍNDICE DE TABLAS .....	xiii
ÍNDICE DE FIGURAS.....	xv
RESUMEN.....	xvii
ABSTRACT.....	xviii
CAPÍTULO I.....	1
INTRODUCCIÓN.....	1
1.1.    Formulación del Problema.....	1
1.2.    Justificación del Problema .....	2
1.3.    Objetivos de la Investigación.....	4
1.3.1.  Objetivo general .....	4
1.3.2.  Objetivos específicos.....	4
CAPÍTULO II.....	5
MARCO TEÓRICO.....	5
2.1.    Virus fitopatógenos .....	5
2.1.1.  Generalidades.....	5

2.1.2.	Modo de infección .....	6
2.1.3.	Mecanismos de transmisión .....	7
2.1.3.1.	Transmisión horizontal .....	7
2.1.3.2.	Transmisión vertical .....	7
2.1.5.	Propagación .....	8
2.1.6.	Control y manejo de enfermedades virales .....	8
2.2.	Efecto de la nutrición mineral en la infección viral .....	9
2.3.	Métodos tradicionales de detección de virus fitopatógenos.....	10
2.3.1.	Ensayo inmunoabsorbente ligado a enzimas (ELISA).....	11
2.3.2.	Reacción en Cadena de la Polimerasa (PCR) y PCR de transcripción inversa (RT- PCR).....	12
2.4.	Metagenómica de virus en plantas.....	13
2.5.	Secuenciación masiva paralela .....	15
2.5.1.	Preparación de Muestras .....	16
2.5.2.	Construcción de librería .....	17
2.5.3.	Amplificación clonal de fragmentos de ADNc presentes en la librería .....	19
2.5.4.	Secuenciación.....	20
2.6.	Aplicación secuenciación masiva paralela en estudios de detección e identificación de virus en plantas .....	21

2.7.	Limon Meyer ( <i>Citrus x meyeri</i> ) .....	24
2.7.1.	Cultivo y producción nacional de limón Meyer .....	24
2.7.2.	Principales enfermedades de cítricos .....	25
2.7.2.1.	Enfermedades causadas por hongos .....	25
2.7.2.2.	Enfermedades causadas por bacterias .....	26
2.7.2.3.	Enfermedades causadas por nematodos.....	27
2.7.2.4.	Enfermedades causadas por virus .....	27
2.8.	Naranjilla ( <i>Solanum quitoense</i> ) .....	28
2.8.1.	Cultivo y producción nacional de naranjilla .....	29
2.8.2.	Variedades e híbridos.....	29
2.8.3.	Principales enfermedades en naranjilla.....	30
2.8.3.1.	Enfermedades causadas por hongos .....	30
2.8.3.2.	Enfermedades causadas por bacterias .....	31
2.8.3.3.	Enfermedades causadas por nematodos.....	31
2.8.3.4.	Enfermedades causadas por virus .....	32
CAPÍTULO III .....		33
MATERIALES Y MÉTODOS .....		33
3.1.	Zona de estudio.....	33
3.1.1.	Análisis de laboratorio .....	33

3.2.	Periodo de duración de la Investigación.....	34
3.3.	Metodología.....	34
3.3.1.	Obtención de muestras de hojas sintomáticas de naranjilla y limón Meyer .....	34
3.3.2.	Limpieza y desinfección de muestras de tejido vegetal sintomático .....	34
3.3.3.	Extracción de ARN total de hojas sintomáticas.....	35
3.3.4.	Síntesis de ADNc .....	36
3.3.5.	Reacción en cadena de la Polimerasa (PCR) .....	37
3.3.5.1.	Electroforesis .....	39
3.3.6.	Detección de virus fitopatógenos mediante secuenciación de ADN complementario .....	40
3.3.7.	Identificación de virus fitopatógenos mediante secuenciación masiva paralela – Illumina .....	40
3.3.7.1.	Análisis de datos .....	41
3.3.7.2.	Comparación de ensambladores de genomas virales .....	43
3.3.7.3.	Detección de virus fitopatógenos mediante el uso de la herramienta bioinformática automatizada VirFind.....	44
	RESULTADOS.....	46
4.2.	Extracción de ARN total y transcripción inversa .....	47
4.3.	Amplificación de muestras de ADNc mediante PCR convencional .....	48

4.4.	Análisis de resultados de secuenciación Sanger.....	50
4.5.	Análisis de muestras de ARN total mediante secuenciación masiva paralela .....	50
4.5.1.	Control de calidad de las muestras .....	51
4.5.2.	Creación de la Librería .....	53
4.5.3.	Secuenciación masiva paralela .....	53
4.6.	Análisis bioinformático de lecturas NGS .....	56
4.9.	Identificación de virus fitopatógenos .....	65
4.10.	Detección de virus fitopatógenos a través de VirFind .....	71
	CAPÍTULO V .....	76
	DISCUSIÓN.....	76
	CAPÍTULO VI .....	84
	CONCLUSIONES.....	84
	CAPÍTULO VII.....	87
	RECOMENDACIONES.....	87
	CAPÍTULO VIII.....	89
	BIBLIOGRAFÍA .....	89

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> <i>Reactivos empleados en la preparación de solución de DNasa I para la purificación de ARN total.....</i>	36
<b>Tabla 2.</b> <i>Primers utilizados para la detección de virus.....</i>	37
<b>Tabla 3.</b> <i>Reactivos empleados en la reacción de cadena de la polimerasa.....</i>	38
<b>Tabla 4.</b> <i>Condiciones de PCR para amplificación con primers genéricos.....</i>	39
<b>Tabla 5.</b> <i>Valores de concentración y pureza de muestras de ADNc de naranjilla y limón Meyer.....</i>	48
<b>Tabla 6.</b> <i>Resultados del análisis de secuencias con la herramienta BLASTN.....</i>	50
<b>Tabla 7.</b> <i>Resultados control de calidad muestras de RNA total.....</i>	51
<b>Tabla 8.</b> <i>Resultado del control de calidad de la librería NGS.....</i>	53
<b>Tabla 9.</b> <i>Datos de la Virus fitopatógenos identificados en muestra de naranjilla con SPAdes y ABySS y diferentes databases.....</i>	54
<b>Tabla 10.</b> <i>Análisis comparativo de ensambladores de novo en muestras de limón Meyer.....</i>	62
<b>Tabla 11.</b> <i>Análisis comparativo de ensambladores de novo en muestras de limón Meyer y naranjilla.....</i>	63
<b>Tabla 12.</b> <i>Virus fitopatógenos identificados en muestra de naranjilla con SPAdes, ABySS y diferentes databases.....</i>	64
<b>Tabla 13.</b> <i>Virus fitopatógenos identificados en muestra de limón Meyer con SPAdes, ABySS y diferentes databases.....</i>	65
<b>Tabla 14.</b> <i>Resultados del análisis de contigs con la herramienta BLASTN en</i>	

<i>muestra de limón Meyer</i> .....	68
<b>Tabla 15.</b> <i>Resultados del análisis de contigs con la herramienta BLASTN en muestra de naranjilla</i> .....	70
<b>Tabla 16.</b> <i>Resultados del análisis de contigs con la herramienta BLASTN en muestra de limón Meyer</i> .....	72
<b>Tabla 17.</b> <i>Resultados del análisis de contigs con la herramienta BLASTN en muestra de naranjilla</i> .....	75

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Descripción procedimiento de preparación de la muestra .....	18
<b>Figura 2.</b> Preparación librería de secuenciación .....	19
<b>Figura 3.</b> Generación de clusters por amplificación clonal de puente.....	20
<b>Figura 4.</b> Diagrama de trabajo de VirFind .....	45
<b>Figura 5.</b> Hojas de plantas de naranjilla con sintomatología vírica .....	46
<b>Figura 6.</b> Hojas de plantas de limón Meyer con sintomatología vírica .....	47
<b>Figura 7.</b> Electroforesis en gel de agarosa al 2% donde se visualiza productos PCR amplificadas con los cebadores Ilar2F5 – Ilar2R9 para Ilarvirus .....	49
<b>Figura 8.</b> Electroforesis en gel de agarosa al 2% donde se visualiza productos de PCR amplificadas con los cebadores Ilar1F5 – Ilar1R7 para Ilarvirus .....	49
<b>Figura 9.</b> Gráfica de las bandas ribosómicas 28S y 18S obtenida en el análisis de calidad de la muestra de naranjilla .....	52
<b>Figura 10.</b> Gráfica de las bandas ribosómicas 28S y 18S obtenida en el análisis de calidad de la muestra de limón Meyer .....	52
<b>Figura 11.</b> Contenido de guanina-citosina y adenina-timina de muestras de naranjilla y limón Meyer. Fuente: (Macrogen Inc., 2018). .....	54
<b>Figura 12.</b> Calidad de las secuencias NGS de naranjilla por base. ....	55
<b>Figura 13.</b> Calidad de las secuencias NGS de limón Meyer por base. ....	55



<b>Figura 14.</b> Contenido de guanina y citosina (GC) por secuencia .....	56
<b>Figura 15.</b> Gráficas de calidad de las secuencias después de usar Trimomatic y BBtools .....	57
<b>Figura 16.</b> Contenido de guanina y citosina (GC) por secuencia, después de usar Trimomatic y BBtools .....	57
<b>Figura 17.</b> Longitud del contig más largo (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras De naranjilla. ....	59
<b>Figura 18.</b> Longitud del contig más largo (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras de limón Meyer .....	59
<b>Figura 19.</b> Número total de bases en los ensamblajes (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras de .....	60
<b>Figura 20.</b> Número total de bases en los ensamblajes (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras de naranjilla .....	61
<b>Figura 21.</b> Virus detectados e identificados en muestra de tejido vegetal de limón Meyer. ....	66
<b>Figura 22.</b> Virus detectados e identificados en muestra de tejido vegetal .....	69
<b>Figura 23.</b> Virus detectados e identificados en muestra de .....	72
<b>Figura 24.</b> Virus detectados e identificados en muestra de tejido vegetal .....	73

## RESUMEN

La naranjilla y el limón son frutales de amplio cultivo en todas las regiones del Ecuador, que han presentado síntomas de infecciones virales que afectan la producción agrícola. En el país son pocos los estudios y reportes de virus fitopatógenos que afectan a estos cultivos, por lo que este trabajo tiene por objetivo detectar e identificar virus fitopatógenos en cultivos de Naranjilla (*Solanum quitoense*) y Limón Meyer (*Citrus x meyeri*) mediante secuenciación masiva paralela a partir de ARN total. Se recolectaron muestras de hojas con síntomas de virosis. El análisis bioinformático de los *reads* obtenidos (92 345 458 *reads* para limón Meyer y 88 747 638 *reads* para naranjilla) se realizó a través de un *pipeline* establecido con herramientas disponibles públicamente para la detección de virus. El ensamblaje de novo de las muestras secuenciadas se evaluó con los ensambladores Velvet, SPAdes y ABySS. El análisis de los datos de secuenciación identificó en plantas de naranjilla la presencia de *Tobacco virus 2*, *Potato leafroll virus*, *Lily symptomless virus*, *Tomato torrado virus*. Mientras que en plantas de limón Meyer se identificó la presencia de *Citrus tristeza virus*, *Citrus vein enation*, *Potato leafroll virus* y *Lily symptomless virus*. Además, se identificó en los dos cultivos mediante análisis de homología la presencia de *Potato yellowing virus* con porcentajes de identidad de 97.14% (limón Meyer) y 98.73% (naranjilla) con la secuencia del segmento 3 de RNA de PYV reportado en el GenBank. Se realizó una confirmación de PYV a través de primers genéricos para *Ilarvirus* y PCR convencional.

### PALABRAS CLAVE:

- VIRUS
- FITOPATÓGENOS
- SECUENCIACIÓN MASIVA PARALELA

## ABSTRACT

Naranjilla and lemon are fruit trees that are widely cultivated in Ecuador, they have shown symptoms of viral infections that affect yields. In the country there are still few studies and reports of phytopathogenic viruses that affect these crops, so this work aims to detect and identify phytopathogenic viruses in naranjilla (*Solanum quitoense*) and Meyer lemon (*Citrus x meyeri*) by massive parallel sequencing of total RNA. Leaf samples were collected from Meyer lemon and naranjilla plants with symptoms of virosis. The bioinformatic analysis of the paired end reads obtained by sequencing (92 345 458 reads for Meyer lemon and 88 747 638 reads for naranjilla) was carried out through a pipeline established for the study which included publicly available tools for virus detection. The de novo assembly of two samples sequenced in this study was evaluated with Velvet, SPAdes and ABySS assemblers. The analysis of the sequencing data indicated in naranjilla plants the presence of *Tobacco virus 2*, *Potato leafroll virus*, *Lily symptomless virus* and *Tomato torrado virus*. While in Meyer lemon plants the presence of *Citrus tristeza virus*, *Citrus vein enation*, *Potato leafroll virus*, *Lily symptomless virus* was detected. The presence of *Potato yellowing virus* with an identity percentage of 97.14% (Meyer lemon) and 98.73% (naranjilla) with the sequence of segment 3 of PVY RNA reported in GenBank was identified in the two plant species. Confirmation of *Potato yellowing virus* was performed through generic primers for *Ilarvirus* and conventional PCR.

### KEYWORDS:

- **VIRUS**
- **PHYTOPATHOGENS**
- **MASSIVE PARALLEL SEQUENCING**

# CAPÍTULO I

## INTRODUCCIÓN

### 1.1. Formulación del Problema

Se han descrito un gran número de virus fitopatógenos nuevos en los últimos años; la base de datos GenBank cuenta con más de 2000 genomas completos de virus y viroides; sin embargo, el 60-95% de las secuencias identificadas en estudios de diagnóstico viral realizados en plantas no cuentan con homólogos detectables (Mesa, 2017; Roossinck, Martin, & Roumagnac, 2015), debido a que algunos de los patógenos virales presentan características genómicas novedosas que han llevado a la propuesta de la creación de nuevos géneros y la revisión del estado actual de la taxonomía de virus (Maliogka et al., 2018).

En Ecuador se han realizado varios estudios de identificación y detección de virus en diversos cultivos con importancia comercial. En los últimos años en varias zonas del país se han observado tanto en naranjilla como en limón Meyer sintomatología viral como ampollamiento, mosaico, deformación de tejido foliar y clorosis; sin embargo, son escasos los estudios e investigaciones que se han realizado usando Next Generation Sequencing (NGS) como técnica de detección de virus fitopatógenos (Alvarez-Quinto, Cornejo-Franco, & Quito-Avila, 2017). La tecnología de secuenciación de próxima generación ha permitido diagnosticar una

mayor cantidad de virus en plantas hospederas, descubrir nuevos virus no descritos previamente; mediante el desarrollado de cadenas de procesos computacionales denominadas *pipelines*, que permiten el análisis de datos de NGS usando herramientas disponibles públicamente, aun así, esta tecnología actualmente es poco utilizada en el país, probablemente debido a que el análisis de los datos obtenidos de la secuenciación es una tarea compleja que requiere de experticia o por la necesidad de usar recursos computacionales avanzados

## **1.2. Justificación del Problema**

La naranjilla es una fruta originaria del Ecuador cultivada principalmente en el oriente y en la zona Andina en las provincias de Chimborazo, Cotopaxi, Imbabura, Pichincha y Tungurahua (INIAP, 2018). El limón es un frutal ampliamente cultivado en el Ecuador, se lo puede encontrar en los valles secos del Litoral (Portoviejo, Echeandía), valles cálidos Interandinos (Puerto Quito, Chota, Guayllabamba, San José de Minas, Tumbaco), y algunos sectores de la Amazonía (Puyo, Nueva Loja) (Medina, 2007). Los frutos de naranjilla y limón Meyer tienen un muy buen potencial de exportación hacia mercados internacionales durante todo el año. Su cultivo es de importancia económica por su rentabilidad y aceptación en el mercado, constituyéndose en una de las principales fuentes de ingresos económicos para el sector productivo del Ecuador (Bolaños, 2007).

Dentro de la industria agrícola, se han producido grandes pérdidas

económicas ocasionadas por enfermedades virales en plantas. Una detección temprana de patógenos y el monitoreo del estado fisiológico de las plantas son necesarios para reducir la propagación de enfermedades en los cultivos (Roossinck et al., 2015). Sin embargo en nuestro país se han llevado a cabo pocos estudios de detección e identificación de virus en plantas, lo cual demuestra el escaso conocimiento sobre patógenos virales que afectan diferentes cultivos en el Ecuador. Roossinck y colaboradores (2015) señalan la importancia de realizar estudios más completos que permitan detectar e incluso descubrir nuevos virus no reportados previamente; a fin de poder establecer un control oportuno que ayude a evitar pérdidas económicas para el sector agrícola (Guevara, 2017).

El método de secuenciación de nueva generación (NGS) o secuenciación masiva paralela es ágil, efectivo y permite detectar e identificar infecciones virales en plantas de forma directa en diferentes estados de infección. A partir de la obtención masiva de datos y con ayuda de herramientas bioinformáticas de análisis es posible obtener genomas completos del virus (Mesa, 2017). El uso de esta tecnología puede optimizar la eficiencia y confiabilidad de los resultados obtenidos, los mismos que servirán posteriormente para el control de enfermedades virales tanto a nivel nacional como internacional (Barba, Czosnek, & Hadidi, 2013).

Un paso crítico en el uso de NGS para el descubrimiento de virus o el análisis de la composición del viroma es la herramienta bioinformática requerida para identificar eficientemente las secuencias virales de una mezcla compleja de

secuencias huésped (Zhao et al., 2017), por ello es importante establecer un *pipeline* bioinformático que permita realizar un análisis de detección e identificación de virus fitopatógenos eficiente, aplicable a cualquier cultivo de interés económico para el país.

### **1.3. Objetivos de la Investigación**

#### **1.3.1. Objetivo general**

Detectar e identificar virus fitopatógenos en cultivos de Naranjilla (*Solanum quitoense*) y Limón Meyer (*Citrus x meyeri*), mediante secuenciación masiva paralela.

#### **1.3.2. Objetivos específicos**

- Detectar la presencia de virus fitopatógenos en plantas de limón Meyer y naranjilla utilizando primers genéricos.
- Identificar virus en plantas de limón Meyer y naranjilla mediante secuenciación masiva paralela.
- Establecer un pipeline bioinformático que permita extraer la mayor cantidad de secuencias correspondientes a virus fitopatógenos a partir de datos de secuenciación masiva paralela.

## **CAPÍTULO II**

### **MARCO TEÓRICO**

#### **2.1. Virus fitopatógenos**

Los virus de plantas son patógenos de importancia económica que han causado pérdidas estimadas en varios miles de millones de dólares por año en todo el mundo. Son capaces de difundirse amplia y rápidamente a cultivos aledaños y están clasificados como el segundo patógeno de plantas más importante después de los hongos (Hull, 2009a; Jeong, Ju, & Noh, 2014).

Los virus y viroides son los más pequeños entre todos los organismos conocidos. El diámetro típico de un virus de planta esférico es de ~ 30 nm , mientras que algunos de los virus filamentosos alcanzan una longitud de ~ 2000 nm o 2  $\mu$ m (Gergerich & Dolja, 2006).

##### **2.1.1. Generalidades**

Los virus son partículas patógenas intracelulares de pequeño tamaño que pueden ser observados solo a través de un microscopio electrónico (Williams, Boehm, & Qu, 2017). Todos los virus son parásitos obligados que no pueden replicarse por sí mismos y dependen de la maquinaria celular de un huésped vivo para hacerlo (Cooper, 2007; Gergerich & Dolja, 2006).



La mayoría de los virus de plantas son partículas relativamente simples formadas por una cubierta proteica llamada cápside que encierra el genoma del ácido nucleico. Su genoma está compuesto de ADN (ácido desoxirribonucleico) o ARN (ácido ribonucleico) que puede ser de una sola cadena (ss) o de doble cadena (ds); la mayoría de los virus fitopatógenos son de ARN monocatenario. Las partículas de virus vienen en una variedad de formas y tamaños. Sin embargo, la mayoría de los virus en plantas tienen forma de varilla o isométricas. Los virus de plantas difieren también en el movimiento dentro del huésped, los síntomas que producen, los métodos de infección, la replicación y diseminación. A menudo, los virus residen en su huésped sin causar ninguna enfermedad o síntomas de la enfermedad. Muchos de estos patógenos latentes aún no se han descubierto (Cooper, 2007; Gergerich & Dolja, 2006).

### **2.1.2. Modo de infección**

Los virus fitopatógenos son parásitos obligados, que frecuentemente provocan la muerte de su huésped. Las infecciones virales tienen lugar a través de la inoculación de plantas con virus ensamblados o transcripciones de ARN infecciosos, que codifican proteínas desconocidas que afectan los procesos celulares del hospedero a medida que ocurre la replicación viral en el tejido de la planta en el lugar por el cual penetró el virus (Shih & Doran, 2009; Stevens, 1983).

### **2.1.3. Mecanismos de transmisión**

La transmisión es una propiedad esencial de los virus, existen dos mecanismos comunes de transmisión usados por los virus fitopatógenos: transmisión horizontal y transmisión vertical (Hull, 2009a).

#### **2.1.3.1. Transmisión horizontal**

El virus es transmitido por una fuente externa, se produce por: la práctica de determinados métodos de reproducción vegetativa artificiales (injertos) usados por los agricultores, organismos que transportan virus de una planta a otra llamados vectores (bacterias, hongos, nematodos, ácaros, insectos, entre otros), inoculación mecánica y a través del polen de plantas infectadas (Hull, 2009b).

#### **2.1.3.2. Transmisión vertical**

El virus se hereda de una planta madre a su progenie y tiene lugar tanto en la reproducción asexual como en la sexual. En la reproducción asexual, las plantas nuevas que se desarrollan a partir de tallos, raíces, bulbos, etc. de la planta madre infectada adquieren el virus. En la reproducción sexual, la transmisión viral es ocasionada por el uso de una semilla infectada (Hull, 2009b).

#### **2.1.4. Importancia de la transmisión**

Los fitovirus dependen principalmente de los diferentes mecanismos de transmisión, para difundirse de un huésped susceptible a otro y poder atravesar la pared celular de la planta (es muy resistente y rígida), e introducirse directamente en el citosol de las células vivas para su supervivencia y continuidad (Whitfield & Rotenberg, 2015).

#### **2.1.5. Propagación**

El virus inicialmente se propaga lentamente de una célula a otra mediante pequeños conductos llamados plasmodesmos que permiten la unión entre células adyacentes. El virus logra llegar al sistema vascular de la planta, en donde comienza a transportarse y distribuirse rápidamente en el hospedero. La propagación del virus en el interior de la planta está estrechamente relacionada con factores como el tipo de virus, temperatura ambiental, variedad y edad del cultivo, entre otros (Rosales et al., 2011).

#### **2.1.6. Control y manejo de enfermedades virales**

Una vez infectada la planta; no se puede volver a recuperar ya que, a diferencia de otros organismos patógenos, los virus fitopatógenos no pueden ser eliminados. Por lo que frecuentemente se emplean métodos de prevención como: la

producción de materiales de siembra libres de virus, uso de semillas certificadas, eliminación de plantas con síntomas virales y el desarrollo de cultivos resistentes. Sin embargo, para que los métodos mencionados anteriormente tengan buenos resultados y exista un manejo óptimo de las enfermedades virales, es importante el uso de métodos sensibles, específicos y eficientes para detectar e identificar el agente viral y poder determinar aspectos como su forma de transmisión, su sobrevivencia en ausencia de cultivo, entre otros (Valarezo Concha, Valarezo Cely, Mendoza, & Alvarez, 2014).

## **2.2. Efecto de la nutrición mineral en la infección viral**

Los nutrientes minerales son importantes tanto para el crecimiento, desarrollo, resistencia y tolerancia de la planta como para la supervivencia de patógenos y son factores fundamentales en las interacciones que suceden entre planta-patógeno (Spann & Schumann, 2009).

Teniendo en cuenta que la infección por virus se limita a las células vivas, la nutrición puede afectar tanto la patogenicidad como la replicación del virus. Usualmente, las condiciones nutricionales que benefician el crecimiento de las plantas también suelen incrementar la capacidad de replicación del virus (Pennazio & Roggero, 1997).

No obstante, la relación entre la nutrición mineral y las enfermedades virales no está clara y puede ser muy compleja, puesto que, un buen estado nutricional puede provocar la aparente disminución o desaparición de los síntomas de la enfermedad viral y a su vez una replicación activa del virus. Esto ocurre debido a que no existe una correlación entre la tasa de multiplicación del virus y la gravedad de los síntomas, por lo que se puede llegar a pensar que la enfermedad viral ha disminuido basándose en la ausencia de síntomas por el uso de nutrientes minerales, cuando la enfermedad no ha sido superada y los síntomas solo están ocultos (Marschner, 2006; Pennazio & Roggero, 1997).

Actualmente, no son muchos los estudios existentes que establezcan la relación entre los nutrientes minerales y las infecciones virales; sin embargo, se ha observado que al combinar fósforo y nitrógeno se duplica la concentración de virus de la savia y aumenta el virus total por planta (Bawden & Kassanis, 1950); así mismo, el silicio a pesar de no ser un nutriente vegetal, es capaz de inhibir la capacidad de nutrición de ciertos insectos (áfidos), ocasionado que la infección viral disminuya. El impacto que puede provocar un nutriente en las plantas infectadas por virus depende de la edad y especie de la planta, factores medioambientales, tipo de virus, nutrientes disponibles (Velasco, 2000).

### **2.3. Métodos tradicionales de detección de virus fitopatógenos**

La detección temprana y precisa de virus de plantas es muy importante para

poder controlarlos, el diagnóstico de la enfermedad viral por síntomas es más difícil que otros patógenos (Jeong et al., 2014) y no hay métodos directos disponibles para controlar los virus. Por lo tanto, los métodos de detección e identificación de virus, tanto en plantas como en vectores, desempeñan un papel fundamental en el manejo de la enfermedad viral (Naidu & Hughes, 2018). Se han desarrollado muchos métodos para detectar virus de plantas, como la observación microscópica, técnicas serológicas (ELISA), métodos moleculares (PCR, NGS), entre otros (Jeong et al., 2014). Sin embargo, una sola prueba de diagnóstico no puede proporcionar información adecuada sobre la identidad de un virus, generalmente es necesario una combinación de métodos para un diagnóstico inequívoco.

Los métodos moleculares se usan para el diagnóstico de muchas enfermedades virales, cuando la información genética de los virus está disponible. Se emplean como método alternativo a las técnicas serológicas, debido a su alta precisión y sensibilidad (Gergerich & Dolja, 2006).

### **2.3.1. Ensayo inmunoabsorbente ligado a enzimas (ELISA)**

ELISA (Enzyme Linked ImmunoSorbent Assay) es un método de detección sensible, preciso, rápido y eficaz cuando se tiene que analizar grandes cantidades de muestras (Ismail, 1997). La introducción de ELISA revolucionó el diagnóstico de virus en plantas al simplificar la detección y reducir el tiempo requerido para la obtención de resultados concluyentes (Boonham et al., 2014).

Este tipo de ensayo se basan en el reconocimiento específico de antígenos víricos por anticuerpos y puede proporcionar datos cuantitativos y cualitativos en poco tiempo (Rowhani & Falk, 1995). Este método no requiere el fraccionamiento de la muestra por electroforesis en gel, se basa en la propiedad de las proteínas para unirse fácilmente a una superficie plástica (Boonham et al., 2014).

### **2.3.2. Reacción en Cadena de la Polimerasa (PCR) y PCR de transcripción inversa (RT- PCR)**

La PCR (*polymerase chain reaction*) es una técnica científica que amplifica millones de copias idénticas a partir de una secuencia de ADN particular dentro de un pequeño tubo de reacción. Se realiza en tres pasos, desnaturalización por encima de 94 °C, *annealing* de cebadores entre 50-75 °C (depende de los cebadores) y elongación a 72 °C (Jeong et al., 2014). La PCR de transcripción inversa (*Reverse Transcription Polymerase Chain Reaction*, RT-PCR) es una técnica molecular sensible, específica, económica y confiable, que utiliza ARNm como plantilla inicial para la transcripción a ADN complementario monocatenario (ADNc) mediante la enzima transcriptasa inversa (Jeong et al., 2014). Las moléculas de ADNc se usan como plantillas para una reacción de PCR convencional (Carter, Shieh, Farra, & Harris, 2015).

Estos métodos se han utilizado en investigaciones basadas en biología molecular como clonación, genotipificación, secuenciación, mutagénesis, entre otros

y se usan en la detección de virus fitopatógenos en el laboratorio; son la base de métodos de diagnóstico utilizados junto con otros métodos de detección (Jeong et al., 2014).

#### **2.4. Metagenómica de virus en plantas**

La metagenómica involucra el estudio de poblaciones virales en muestras ambientales a través de la secuenciación de ácidos nucleicos, mediante el método de secuenciación masiva paralela que produce conjuntos de datos muy grandes (Mutuku et al., 2018). En los últimos años, se han realizado varios estudios de tipo metagenómico que han explorado la biodiversidad de virus en plantas (Stobbe & Roossinck, 2014) y han permitido descubrir nuevos virus y revelar las relaciones mutualistas entre virus y plantas (Roossinck et al., 2015). En 2018 Mutuku y colaboradores mediante un análisis metagenómico determinaron la presencia de *Bean common mosaic necrosis virus* (BCMNV) y *Phaseolus vulgaris alphaendornaviruses 1 and 2* (PvEV1 and 2) en plantas de fréjol (*Phaseolus vulgaris*) en Kenia. En el mismo año, Wamaitha y colaboradores realizaron un estudio metagenómico para identificar los virus asociados con necrosis letal del maíz y como resultado pudieron ensamblar genomas completos y parciales de cuatro virus diferentes *Maize chlorotic mottle virus* (MCMV) *Sugarcane mosaic virus* (SCMV) *Maize streak virus* (MSV) and *Maize yellow dwarf virus-RMV* (MYDV-RMV) (Wamaitha et al., 2018).



Los estudios metagenómicos de virus de plantas se han denominado ecogenómicos, debido a que se analiza poblaciones virales en plantas individuales de los cultivos y se pueden muestrear repetidamente, a diferencia de los estudios metagenómicos tradicionales que utilizan muestras ambientales (Roossinck et al., 2015; Stobbe & Roossinck, 2014)

La metagenómica viral consta de tres procesos principales: muestreo y recolección de virus en tejidos o muestras ambientales; secuenciación de ácidos nucleicos (empleando principalmente técnicas NGS); y finalmente el análisis bioinformático que servirá para manipular e interpretar de manera adecuada los datos obtenidos y así poder proporcionar información sobre la comunidad viral asociada con la muestra y la diversidad dentro de cada especie viral (Pantaleo & Chiumenti, 2018).

El enfoque más simple para la metagenómica viral es utilizar directamente el ARN total extraído de la planta; sin embargo, la principal deficiencia de este enfoque es que a pesar de que se obtiene un alto porcentaje de lecturas de secuencia, no todas son derivadas del patógeno viral (Roossinck et al., 2015). La eliminación de secuencias no virales de la muestra antes de la secuenciación y la correcta manipulación de los datos de secuencia obtenidos, incrementan la probabilidad de identificar virus fitopatogénos en un conjunto de datos de secuencias metagenómicas (Stobbe & Roossinck, 2014).

La metagenómica viral a través de la tecnología de secuenciación de alto rendimiento y el desarrollo de herramientas bioinformáticas para el ensamblaje, han revolucionado la tasa de descubrimiento de virus, la naturaleza de los descubrimientos y han proporcionado información sobre nuevos virus (Alavandi & Poornima, 2012; Dutilh, Reyes, Hall, & Whiteson, 2017).

Los virus fitopatógenos identificados a través de datos de secuenciación metagenómica se pueden clasificar en tres categorías: virus conocidos-conocidos: especies de virus de los cuales ya se tiene información y se sabe que se encuentran en el medio ambiente estudiado; virus desconocidos-conocidos: nuevas especies de virus o aislados de una familia conocida, o virus conocidos que no se han sido encontrados previamente en el sitio de estudio y; virus desconocido-desconocido: virus que son completamente nuevos y que comparten poca o ninguna similitud de secuencia con otros virus conocidos (Stobbe & Roossinck, 2014).

## **2.5. Secuenciación masiva paralela**

Secuenciación masiva paralela (*Massive Parallel Sequencing*, MPS) también denominada secuenciación de nueva generación o de segunda generación (*Next Generation Sequencing*, NGS), es una tecnología que ha hecho posible detectar, identificar y descubrir nuevos virus en varias plantas sin anticuerpos o conocimiento previo de las secuencias de virus. El genoma viral completo podría secuenciarse a partir de plantas sintomáticas o asintomáticas a través de la secuenciación de

próxima generación de ácidos nucleicos totales, incluidos los ARN pequeños (Prabha, Baranwal, & Jain, 2013).

Las plataformas disponibles comercialmente como Illumina, tienen la capacidad de secuenciar completamente los genomas individuales con altos niveles de cobertura, generando millones de lecturas en una sola ejecución. Los datos NGS son particularmente ventajosos para el estudio de la variación estructural (SV) porque ofrece la sensibilidad para detectar variantes de diferentes tipos y tamaños (Koboldt, Larson, Chen, Ding, & Wilson, 2012). Con la llegada de NGS, se ha incrementado la capacidad de secuenciación y se ha facilitado enormemente los esfuerzos para identificar virus (Zhao et al., 2013).

Cada plataforma NGS tiene características específicas, pero el esquema de trabajo que utilizan es conceptualmente parecido y comprende los siguientes pasos: preparación de muestras, construcción de librerías a partir de ARN, amplificación clonal de los fragmentos de ADNc en la librería, secuenciación, generación de datos, análisis primario/secundario (Zoni & Martínez, 2015).

### **2.5.1. Preparación de Muestras**

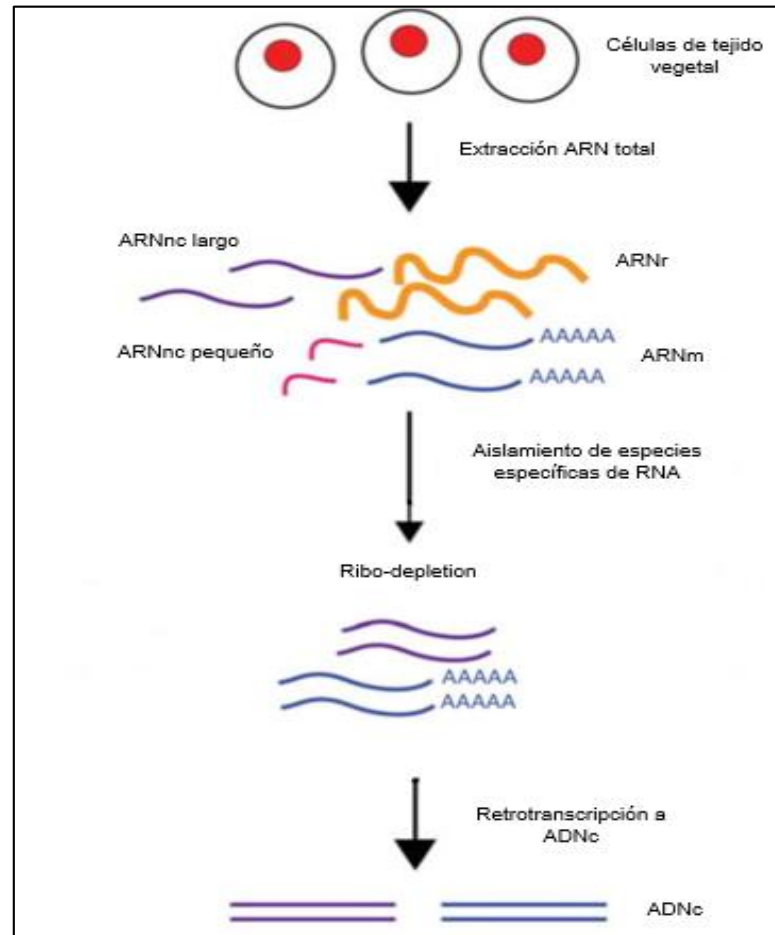
El ARN extraído de la muestra vegetal es sometido a un control de calidad (QC), para garantizar que tiene la calidad necesaria para producir una biblioteca para la secuenciación. La calidad del ARN de la muestra se mide empleando un

bioanalizador, que genera un número de integridad del ARN (RIN) en el rango de 1 a 10 (siendo 10 una muestra de alta calidad con menor degradación). El RIN determina la integridad de la muestra empleando electroforesis en gel y el análisis de los ratios de las bandas ribosómicas 28S a 18S (Kukurba & Montgomery, 2015).

El conjunto total de ARN incluye ARN ribosomal (ARNr), ARN mensajero (ARNm) y distintos tipos de ARN no codificante (ARNnc). En la mayoría de las células, la mayor parte de las moléculas de ARN son ARNr, que si no se eliminan antes de la construcción de la biblioteca, consumirán la mayor parte de las lecturas de secuenciación, lo que reducirá la cobertura de la secuencia y limitará la detección de otros ARN menos abundantes. De manera que, emplear el método de *Ribo-depletion* es una opción recomendada para eliminar el ARNr (Figura 1) (Kukurba & Montgomery, 2015) .

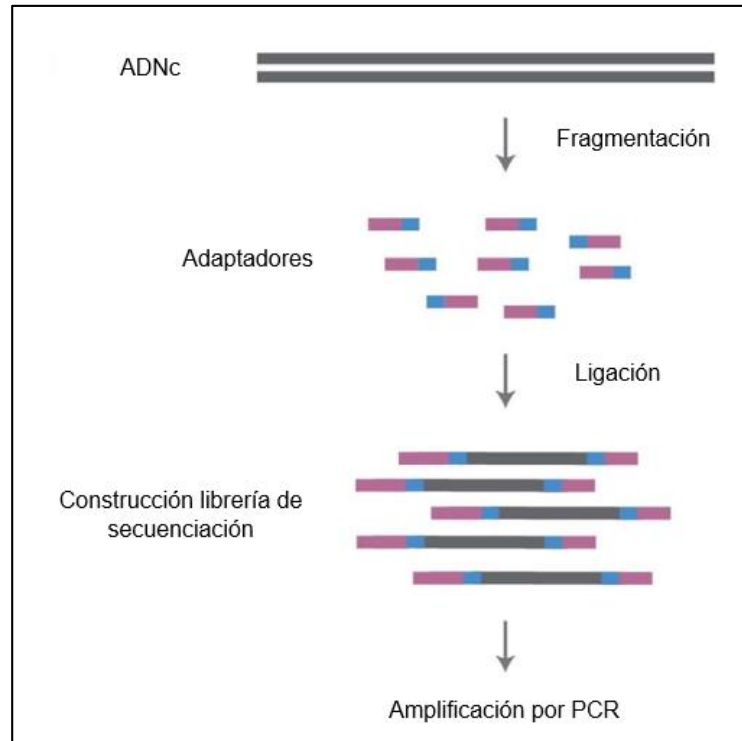
### **2.5.2. Construcción de librería**

La construcción de la librería de secuenciación involucra el aislamiento de moléculas de ARN deseadas, la transcripción inversa del ARN a ADNc como se puede ver en la Figura 1; y la fragmentación aleatoria de moléculas de ADNc de la muestra para crear cadenas más pequeñas y la unión de adaptadores especializados (cuya secuencia es característica de cada tecnología de secuenciación comercializada) en los dos extremos del fragmento 5' y 3' (Figura 2) (Kukurba & Montgomery, 2015; Macrogen, 2018a).



**Figura 1.** Descripción procedimiento de preparación de muestra y retrotranscripción a ADNc.  
Adaptado de: (Kukurba & Montgomery, 2015)

Los fragmentos ligados al adaptador son amplificados por PCR, para crear grandes cantidades de grupos de ADN e intensificar la señal que es recibida desde el secuenciador, para que esta sea lo suficientemente fuerte para ser detectada con precisión y exactitud. Finalmente los productos obtenidos son purificados en gel (Illumina Inc., 2017; Macrogen, 2018a).



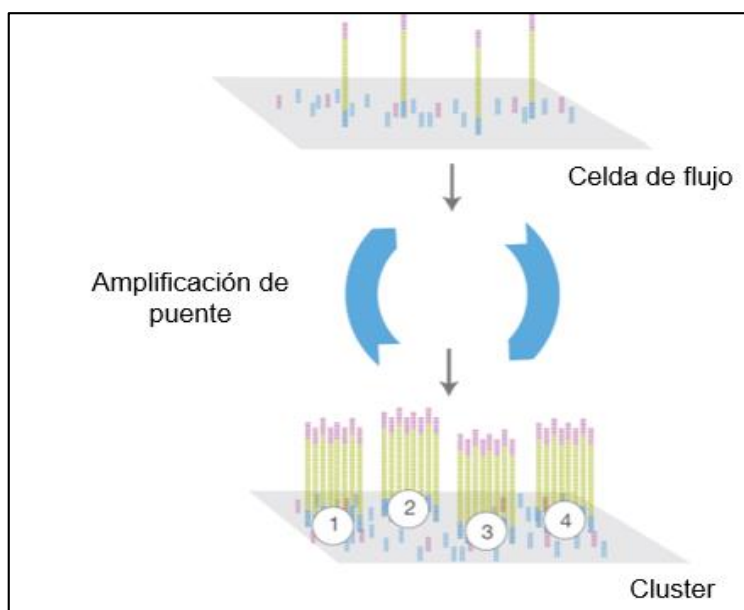
**Figura 2.** Preparación librería de secuenciación  
Adaptado de: (Illumina Inc., 2017)

En la preparación de librerías suele emplearse un método alternativo llamado *tagmentation*, el cual combina las reacciones de fragmentación y ligación en un solo paso aumentando considerablemente la eficiencia del proceso de desarrollo de la librería. La creación de librerías de ARN puede variar según la plataforma de NGS utilizada (Illumina Inc., 2017; Macrogen, 2018a).

### 2.5.3. Amplificación clonal de los fragmentos de ADNc presentes en la librería

La librería es cargada en una celda de flujo, donde la superficie de los adaptadores de los fragmentos se hibridan con la superficie de oligos

complementarios de la celda de flujo. Cada fragmento unido se amplifica clonalmente mediante amplificación de puente y se agrupan en *clusters* (Figura 3). Finalmente los *cluster* están listos como plantillas para la secuenciación (Illumina Inc., 2017; Macrogen, 2018a).



**Figura 3.** Generación de clusters por amplificación clonal de puente. Adaptado de: (Illumina Inc., 2017)

#### 2.5.4. Secuenciación

La tecnología de secuenciación por síntesis (SBS) de Illumina utiliza un método propio fundamentado en terminador reversible, que encuentra bases individuales mientras se van incorporando a las hebras de la plantilla de ADNc. La secuenciación de extremo pareado (*paired end*, PE) es un importante avance de la tecnología de secuenciación masiva paralela, que involucra la secuenciación de los

dos extremos de los fragmentos de ADNc en la librería y la alineación de las lecturas *forward* y *reverse* como lecturas pareadas (Illumina Inc., 2017). Los datos de secuenciación se analizan a través del uso de herramientas bioinformáticas.

## **2.6. Aplicación de secuenciación masiva paralela en estudios de detección e identificación de virus en plantas**

En el año 2000 inició la era de las tecnologías NGS y en el año 2004 se comercializó la primera plataforma de secuenciación. A partir de estos acontecimientos se ha ido expandiendo la industria de NGS con diferentes compañías que comercializan varios modelos de plataformas de secuenciación (Barba et al., 2013). El descubrimiento de nuevos virus se ha incrementado rápidamente después de la introducción de esta tecnología (Pecman et al., 2017).

Las tecnologías NGS combinadas con bioinformática se utilizaron con éxito en virología de plantas desde 2009 cuando Kreuze y sus colaboradores realizaron un experimento para evaluar a través de NGS de miRNAs la ruta de silenciamiento de RNA y el efecto de los virus sobre ella. Al procesar los datos obtenidos observaron que se podría usar esta misma información para realizar diagnóstico viral, es así como identificaron y ensamblaron el genoma completo de dos virus (*Sweetpotato feathery mottle virus*, SPFMV) y *Sweetpotato chlorotic stunt virus*, SPCSV) extraídos de plantas de camote (*Ipomoea batatas*) (Barba et al., 2013; Mesa, 2017). En el mismo año, otro estudio pionero del diagnóstico de virus fue el realizado por Adams



y colaboradores usando NGS, quienes a partir de la secuenciación de cDNA obtenido del RNA extraído de tejido vegetal de tomate, obtuvieron la secuencia completa del genoma de *Pepino mosaic virus* (PepMV). Utilizando la misma metodología del estudio en tomate, buscaron identificar en amaranto redondo (*Gomphrena globosa*) un patógeno desconocido y lograron detectar, identificar y ensamblar el genoma de un nuevo miembro del género *Cucumovirus* (*Gayfeather mild mottle virus*) (Mesa, 2017).

A partir de estos estudios, la tecnología de secuenciación de próxima generación ha permitido diagnosticar virus en plantas hospederas, descubrir e identificar nuevos virus como *Citrus chlorotic dwarf-associated virus* (Loconsole et al., 2012), *Raspberry latent virus* (Quito-Avila, Jelkmann, Tzanetakis, Keller, & Martin, 2011), *Tomato necrotic stunt virus* (Li et al., 2012), *Naranjilla chlorotic mosaic virus* (Green et al., 2017), *Pepper yellow leaf curl virus* y *Eggplant mild leaf mottle virus* (Dombrovsky, Glanz, Pearlsman, Lachman, & Antignus, 2010) y obtener genomas completos de varios virus fitopatógenos (Barba et al., 2013; Mesa, 2017). Los resultados obtenidos mediante la tecnología de secuenciación masiva paralela, han revelaron la solidez de las tecnologías NGS para realizar metagenómica viral en plantas detectando e identificando virus conocidos y nuevos (Guevara, 2017).

Investigadores han desarrollado *pipelines* que abordan las limitaciones del análisis de datos de NGS e incluyen herramientas disponibles públicamente para la detección de virus. Los *pipelines* desarrollados utilizan un flujo de trabajo, que incluye

el mapeo de lecturas de secuencia contra genomas de referencia de virus existentes, o el ensamblaje de *novο* de lecturas y la posterior identificación de *contigs* ensamblados que se alinean con las secuencias de virus presentes en las bases de datos. Este último tiene la ventaja de descubrir nuevos virus, los dos métodos de ensamblaje han demostrado ser exitosos para confirmar y caracterizar la diversidad genética dentro de la especie, sin embargo, consumen bastante tiempo y requieren recursos computacionales extensos y un procesamiento previo de los datos (Macrogen, 2018b; Visser, Burger, & Maree, 2016).

Algunos de los *pipelines* desarrollados para la detección de virus a partir de datos NGS también han sido compilados en una sola canalización y empaquetados en una interfaz fácil de usar, dando lugar a software que ayudan a realizar el análisis de *reads* NGS a científicos e investigadores de forma rápida, económica y eficiente. Zhao y colaboradores (2017) diseñó VirusSeeker una línea de análisis basada en BLAST para el descubrimiento de virus eucarióticos y el análisis de la composición del viroma. Ho & Tzanetakis (2014) construyeron VirFind, una herramienta de bioinformática diseñada específicamente para la detección y descubrimiento de virus, capaz de mapear y filtrar las lecturas del host, entregar archivos de lecturas de virus con información taxonómica y los informes correspondientes de Blastn y Blastx. VirusDetect es un canal automatizado desarrollado en el año 2017 por Zheng y colaboradores para analizar eficientemente los conjuntos de datos de ARN pequeño (sRNA), es altamente sensible y eficiente en la identificación de virus.

## **2.7. Limon Meyer (*Citrus x meyeri*)**

*Citrus x meyeri* es un árbol pequeño de hoja perenne, perteneciente a la familia Rutaceae que puede alcanzar una altura entre 3 y 6 m y una capacidad de fructificación de 40 años aproximadamente. El desarrollo, rendimiento y calidad tanto de los cultivos como de los productos depende de varios factores como el tipo de suelo, humedad relativa y condiciones ambientales adecuadas. La temperatura óptima para que el cultivo de este cítrico tenga un buen crecimiento vegetativo se encuentran a una temperatura promedio de 18 °C y 30 °C, con una altitud de hasta 1800 msnm (Valarezo Concha et al., 2014).

### **2.7.1. Cultivo y producción nacional de limón Meyer**

El limón Meyer es cultivado en todas las zonas del país, puesto que la ubicación geográfica y las condiciones climáticas de las diferentes regiones del Ecuador favorecen el cultivo de este cítrico. De acuerdo con los datos obtenidos en el año 2000 a través del Censo Nacional Agropecuario, la mayor producción de este cítrico se centra en la zona Costera del país con el 57%, seguido por la región Sierra y la Amazonía con una producción de 39% y 1% respectivamente (León, 2009).

De acuerdo con la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC) realizada en el año 2017 por el Instituto nacional de estadística y censos (INEC), la superficie de cultivos de limón plantada es de 8.152 ha y la

superficie cosechada es de 5.487 ha, obteniéndose una producción de 30.702 toneladas métricas y 29.315 toneladas métricas en ventas. A partir del año 2008, las exportaciones de este cítrico incrementaron considerablemente, siendo Estados Unidos y Colombia los principales mercados internacionales que adquieren el producto ecuatoriano (León, 2009).

### **2.7.2. Principales enfermedades de cítricos**

Los cítricos, sin importar la edad de la planta, son propensos a enfermedades causadas por agentes fitopatógenos como: hongos, nematodos, bacterias y virus que atacan raíces, tallo, hojas, flores, frutos. Provocan el debilitamiento de las plantas y disminuyen el rendimiento de las cosechas (Valarezo Concha et al., 2014).

#### **2.7.2.1. Enfermedades causadas por hongos**

En los cultivos de cítricos se han reportados diferentes enfermedades tales como el conocido mal del semillero que es provocado por un complejo de hongos que habitan en el suelo como *Phythium* sp., *Rhizoctonia solani* Kuhn, y *Phytophthora* sp. y afectan a las semillas en estado de germinación. La gomosis o pudrición del tallo y raíces es causada por hongos de las especies *Phytophthora* sp. Produce la podredumbre de las raíces, tallo, y es capaz de alcanzar incluso las partes altas de la planta como los frutos. La 'muerte regresiva' es una enfermedad que se presenta en plantas de cualquier edad, se caracteriza por presentar síntomas como clorosis,

decaimiento de la planta, necrosis, defoliaciones y es provocada por *Fusarium oxysporum*. El agente causal de la antracnosis de los cítricos o caída prematura de flores y fruto es el hongo *Colletotrichum gloesporioides*, que ataca fundamentalmente a los brotes jóvenes de la planta produciendo manchas necróticas. Se han identificado además enfermedades como mancha grasienta, afelpado de los cítricos, fumagina, secamiento de los cítricos, mal rosado, secamiento de ramas, mancha foliar o mancha café, melanosis, tizón cuyos agentes causales son *Mycosphaerella citri*, *Septobasidium pseudopedicellatum*, *Capnodium citri*, *Ceratocystis fimbriata*, *Corticium salmonicolor*, *Botryodiplodia theobromae*, *Alternaria* (especies *alternata*, *limicola* y *tenuissima*), *Diaporthe citri* Wolf, *Botrytis cinerea* Pers. respectivamente (Castro, Timmer, Leguizamón, Müller, & Corrales. José Antonio, 2000; Valarezo Concha et al., 2014).

#### **2.7.2.2. Enfermedades causadas por bacterias**

Las enfermedades descritas en cítricos ocasionadas por bacterias son tres, el cáncer o bacteriosis una enfermedad muy severa causada por *Xanthomonas axonopodis* pv. Citr que produce la defoliación y caída de frutos. La clorosis variagada de los cítricos (CVC) afecta principalmente a plantas jóvenes, la bacteria causante de esta enfermedad es *Xylella fastidiosa* y se aloja en el xilema de la planta por lo que se transmite a través de vectores como los insectos. Huanglongbing de los cítricos (HLB) es quizá la enfermedad que más daño ocasiona a los cultivos de cítricos debido a que provoca la degeneración del floema y daña el transporte de

savia elaborada cuando la planta está infectada, su agente causal es una bacteria Gram negativa y se transmite a través de insectos (Castro et al., 2000; Luis, Llauger, & Collazo, s. f.).

### **2.7.2.3. Enfermedades causadas por nematodos**

En las zonas de cultivo de cítricos se ha identificado el nematodo *T. semipenetrans* del género *Tylenchulus* Cobb como organismo causante de la Muerte Lenta o Paloteo Lento en los árboles de cítricos. Otros nematodos que atacan estos cultivos y que han sido reportados son *Radopholus citrophilus*, *Pratylenchus* spp (especies *P. vulnus*, *P. coffeae*, *P. brachyurus*) que atacan portainjetos y provocan declinamiento de la planta; *Belonolaimus longicaudatus* y *Belonolaimus longicaudatus* que impiden el crecimiento de raicillas (Castro et al., 2000).

### **2.7.2.4. Enfermedades causadas por virus**

Las enfermedades virales descritas en cítricos a nivel mundial son varias; sin embargo, en el Ecuador no existe un reporte detallado de los virus fitopatógenos de cítricos existentes. En la guía técnica sobre el manejo de los cítricos en el litoral ecuatoriano del INIAP (Valarezo Concha et al., 2014) se menciona que se ha observado síntomas víricos de mosaico, reducción de limbo, descortezamiento, clorosis en plantaciones de cítricos. Las enfermedades víricas con más incidencia son dos, la tristeza de los cítricos causada por el virus del mismo nombre (CTV)

perteneciente al género de los *Closterovirus* y transmitido mecánicamente o por vectores como el pulgón negro; y la psorosis, una enfermedad transmitida por injertos que se ha observado en la zona costera del país, el agente causal es el virus de la psorosis de los cítricos (*CPsV*) que causa la reducción de la producción y vigor de la planta, descamaciones tanto en el tronco como en las ramas de los árboles infectados (Castro et al., 2000; Peña, Pérez, López, & Batista, s. f.; Valarezo Concha et al., 2014).

Otros virus fitopatógenos causantes de enfermedades víricas de cítricos de importancia son el *virus de la leprosis de los cítricos* (CiLV) responsables de causar leprosis o clorosis zonada, *Citrus tatter leaf virus* (CTLV), *Citrus variegation virus* (CVV), *Citrus yellow mosaic virus* (CYMV), *Citrus leaf rugose virus* (CiLRV), *Citrus leaf blotch virus* (CLBV), *Indian citrus ringspot virus* (ICRSV), *Citrus yellow vein clearing virus* (CYVCV) y *Satsuma dwarf virus* (SDV) (Castro et al., 2000; Jooste, 2017).

## **2.8. Naranjilla (*Solanum quitoense*)**

La naranjilla es una fruta exótica con excelentes propiedades nutricionales. De origen andino, cultivada principalmente en Ecuador y Colombia. Tiene gran importancia comercial en el mercado nacional y presenta gran potencial en el mercado internacional. Son plantas arbustivas de la familia de las Solanáceas, cuyas características difieren dependiendo de la variedad cultivada (Viteri Diaz et al., 2009).

### **2.8.1. Cultivo y producción nacional de naranjilla**

En nuestro país el cultivo y producción de las variedades e híbridos de naranjilla se realiza a una altitud de 800 - 1500 msnm y una temperatura que fluctúa entre 17°C y 24°C (Bucheli, 2013). La provincia de Napo, Pastaza y Morona Santiago en la Amazonía son las zonas con mayor producción de naranjilla a nivel nacional, seguida en menor porcentaje por zonas de la región Andina (Torres, Vargas, García, Arteaga, & Navarrete, 2017) . En el año 2010 la producción nacional de esta fruta fue de 20005 toneladas métricas y un rendimiento de 5.49 t/ha (Silva, Gómez, Viera, Viteri, & Ron, 2016). Sin embargo, en el año 2015 de acuerdo con datos estadísticos del Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP) la producción a nivel nacional de naranjilla incrementó a 44.360 toneladas métricas y el rendimiento fue de 13.52 t/ha; con lo cual se evidencia un incremento importante en la producción de este frutal en los últimos años (Dávila, 2016).

### **2.8.2. Variedades e híbridos**

En Ecuador existen diferentes variedades de naranjilla, siendo las más frecuentes y típicas la Variedad agria (*Solanum quitoense* Lam var. *quitoense*), Variedad Baeza dulce (*Solanum quitoense* Lam var. *quitoense*), Variedad espinosa (*Solanum quitoense* Lam var. *septentrionale*) y la Variedad INIAP-Quitoense (*Solanum quitoense* Lam var. *quitoense*). Existen 3 tipos de híbridos comerciales, Híbrido Puyo cultivado en la región Amazónica obtenida del cruzamiento de una



especie silvestre conocida como naranjilla jibara del Oriente (*S. sessiliflorum*) y naranjilla común. Híbrido INIAP Palora producida del cruzamiento entre la Variedad Baeza (*Solanum quitoense* Lam var. *quitoense*) y una variedad de *S. sessiliflorum*; y finalmente el Híbrido Mera producido en las zonas Napo y Pastaza (Revelo et al., 2010).

En nuestro país las variedades que más se cultivan son la naranjilla común (*Solanum quitoense* Lam.), Variedad Baeza dulce (*Solanum quitoense* Lam var. *quitoense*), al igual que los híbridos Puyo e INIAP-Palora, siendo este último el híbrido de mayor comercialización (Fiallos, 2000; Guayasamín, 2015).

### **2.8.3. Principales enfermedades en naranjilla**

La producción de naranjilla se encuentra afectada por agentes patógenos como: hongos, nematodos, virus y bacterias que causan síntomas de manchas, pudriciones, mosaicos, marchitamientos, entre otros que influyen en la productividad y calidad de la fruta (Bucheli, 2013; Revelo et al., 2010).

#### **2.8.3.1. Enfermedades causadas por hongos**

En los cultivos de naranjilla se han reportado enfermedades provocadas por hongos tales como el tizón tardío que ataca a la naranjilla común. Esta enfermedad, causada por *Phytophthora infestans*, ataca al tallo, hojas y brotes jóvenes de la

planta, sin embargo, los híbridos Puyo y Palora presentan resistencia. La antracnosis es causada por el hongo *Colletotrichum gloeosporioides* transmitida a través del viento; ataca a frutos, tallos y brotes jóvenes de naranjilla común e híbridos (con menor severidad) provocando manchas oscuras. *Fusarium Oxysporum* Schlecht es el patógeno causante de fusariosis en plantas susceptibles de naranjilla común e híbrido Puyo ocasionando el amarillamiento y pérdida de hojas en las plantas. Otras enfermedades como la pudrición y la mancha clorótica de la hoja provocadas por *Sclerotinia sclerotiorum* y *Cladosporium* sp. Link. también han sido reportadas en naranjilla (Revelo et al., 2010).

#### **2.8.3.2. Enfermedades causadas por bacterias**

Se ha observado síntomas de marchitez, pudrición, mal olor y muerte de plantas de naranjilla producidos por bacterias como *Ralstonia solanacearum* y *Pectobacterium solanacearum*, causantes de enfermedades como la Marchitez bacteriana y la Pudrición bacteriana respectivamente (Revelo et al., 2010).

#### **2.8.3.3. Enfermedades causadas por nematodos**

En los cultivos de naranjilla se ha detectado la presencia de una especie de nematodo llamado *Meloidogyne incógnita*, que ataca a todas las variedades susceptibles de naranjilla cultivadas. Produce síntomas tales como clorosis, marchites, reducción del tamaño y número de raíces, generación de nudos en la raíz

que obstaculiza la asimilación de agua y nutrientes. También se han identificado otros tipos de nematodos como: *Meloidogyne javanica* y *Meloidogyne arenaria* (Dávila, 2016; Revelo et al., 2010).

#### **2.8.3.4. Enfermedades causadas por virus**

En Colombia se han observado síntomas de virosis como amarillamiento intervenal, clorosis, mosaicos y deformación causados por *Cucumber mosaic virus* (CMV), *Potato yellow vein virus* (PYVV) y *Alstroemeria necrotic streak virus* (ANSV) (Gallo, Toro, Jaramillo, Gutiérrez, & Marín, 2018). En el año 2017, Quito-Avila y colaboradores caracterizaron un nuevo *Tymovirus*, *Naranjilla chlorotic mosaic virus* (NarCMV) (nombre tentativo) causante de atrofia y mosaico clorótico en naranjilla (*Solanum quitoense*)(Green et al., 2018).

## **CAPÍTULO III**

### **MATERIALES Y MÉTODOS**

#### **3.1. Zona de estudio**

La recolección de muestras de naranjilla (*Solanum quitoense*) y limón Meyer (*Citrus x meyeri*) que presentaron síntomas de virosis (mosaico, amarillamiento intervenal, ampollamiento, manchas cloróticas, deformación de brotes y tejido foliar) se realizó en la Granja experimental del Instituto Nacional de Investigaciones Agropecuarias (INIAP) de la Provincia de Pichincha, ubicado en la Av. Interoceánica km 15 y Eloy Alfaro, sector Tumbaco con coordenadas de latitud 00°12'56"S y longitud 78°24'44"O a una altura de 2348 m.s.n.m.

##### **3.1.1. Análisis de laboratorio**

El trabajo de investigación se llevó a cabo en el Laboratorio de Investigación de Microbiología Molecular de la carrera de Ingeniería en Biotecnología de la Universidad de las Fuerzas Armadas-ESPE, ubicado en la Av. General Rumiñahui S/N - Sector "La Colina" parroquia Sangolquí, cantón Rumiñahui, Provincia de Pichincha, Ecuador; con coordenadas de latitud 0°18'46.8"S y longitud 78°26'43.0"O a una altura de 2518.640 m.s.n.m.

### **3.2. Periodo de duración de la Investigación**

La investigación inició el 04 de junio del 2018 y finalizó el 6 de mayo del 2019.

### **3.3. Metodología**

#### **3.3.1. Obtención de muestras de hojas sintomáticas de naranjilla y limón Meyer**

Se identificaron plantas enfermas mediante observación y se recolectaron muestras de hojas con síntomas relacionados con enfermedades virales como ampollamiento, deformación de brotes y tejido foliar, mosaico y clorosis en cultivos de naranjilla (*Solanum quitoense*), mientras que en limón Meyer (*Citrus x meyeri*) se observaron síntomas de ampollamiento, amarillamiento intervenal, deformación de brotes y clorosis. Se almacenaron en bolsas de papel estériles y se transportaron manteniendo una cadena de frío hasta el laboratorio de Microbiología de la Universidad de las Fuerzas Armadas-ESPE.

#### **3.3.2. Limpieza y desinfección de muestras de tejido vegetal sintomático**

Se llevó a cabo un proceso de desinfección y lavado superficial de las muestras vegetales para eliminar residuos de tierra y polvo, posteriormente se realizaron cortes en forma de discos del tejido vegetal sintomático con saca bocados

y se almacenaron a  $-80^{\circ}\text{C}$  en tubos plásticos de 15 mL con gel de sílice y papel absorbente para garantizar su conservación hasta su posterior uso.

### **3.3.3. Extracción de ARN total de hojas sintomáticas**

El ARN total se extrajo de aproximadamente 150 mg de tejido vegetal sintomático, utilizando el kit de extracción de PureLink® RNA Mini Kit (Ambion, LifeTechnologies), siguiendo las instrucciones del fabricante (ANEXO 1). Las muestras fueron maceradas con buffer de lisis (incluido en el kit) en un homogenizador (Mini-BeadBeater, BioSpec, USA) por lapsos de 1 min hasta obtener una maceración completa. La extracción se realizó en una cámara de flujo laminar manteniendo siempre la cadena de frío de las muestras, además se usó un removedor de RNasas RNaseZap (Ambion The RNA Company, USA) para eliminar contaminación y evitar la degradación de ARN.

Durante el proceso de purificación de ARN, se realizó una digestión con DNase I (PureLink™, Invitrogen™, USA) para eliminar el ADN presente en las muestras. Se añadió 10  $\mu\text{L}$  de solución de DNasa I a una concentración de 1u/  $\mu\text{L}$  (Tabla 1) en la membrana de cada columna que contenía el ARN adherido, se incubó a temperatura ambiente por 15 min y se removió la DNasa I añadiendo 350  $\mu\text{L}$  de Buffer de lavado I en la columna y centrifugando a 12,000 g por 15 s. Finalmente se continuó con el procedimiento de lavado y elución de ARN descrito por los fabricantes del kit de extracción.

**Tabla 1***Reactivos empleados en la preparación de DNasa I para purificación de ARN total.*

<b>Reactivo</b>	<b>Concentración</b>	<b>Volumen</b>
DNase I Buffer	10X	1 $\mu$ L
Agua libre de ARNasas	-	8 $\mu$ L
DNase I, Amplification Grade	1 U/ $\mu$ l	1 $\mu$ L
<b>Volumen Total</b>		<b>10 <math>\mu</math>L</b>

Las extracciones obtenidas de ARN total fueron conservadas a largo plazo a una temperatura de  $-80^{\circ}\text{C}$ .

### 3.3.4. Síntesis de ADNc

La síntesis de ADN complementario (ADNc) se realizó a partir de 5  $\mu$ L de ARN total extraído previamente. Se sintetizó una cadena de ADN complementaria a partir de un cebador que usó ARNm como plantilla, utilizando la transcriptasa inversa del virus de la leucemia murina de Moloney (M-MLV Reverse Transcriptase) (Invitrogen, USA), según las instrucciones del fabricante (ANEXO 2).

Se obtuvo un volumen final de reacción de 20  $\mu$ l en un tubo de microcentrífuga libre de nucleasas. La calidad de las muestras de ADNc sintetizadas fueron cuantificadas en el espectrofotómetro NanoDrop 2000 UV-Vis, finalmente las muestras fueron almacenadas a  $-20^{\circ}\text{C}$ .

### 3.3.5. Reacción en cadena de la Polimerasa (PCR)

Se amplificó el ADNc obtenido mediante PCR convencional, usando cebadores genéricos para virus del género *Illavirus* diseñados para el gen de ARN polimerasa dependiente de ARN (Tabla 2).

**Tabla 2**

*Primers utilizados para la detección de virus en naranjilla y limón Meyer.*

Género	Primer	Secuencia (5'-3')	T <sup>°</sup> m (°C)	Tamaño del fragmento amplificado
Illavirus	llar2F5	TCRAYRTTYGAYAARTCNCA	46.1	~380 pb
	llar2R9	GGTTGRTTTRTGHGGRAAYTT	50.2	
	llar1F5	GCNCGWTGYGGDAARWCNAC	56.4	~300 pb
	llar1R7	AMDGGWAYYTGYTYNGTRTCACC	55.7	

Fuente: (Untiveros, Perez-Egusquiza, & Clover, 2010).

Las reacciones de PCR se efectuaron a un volumen final de reacción de 20 µL, empleando el kit GoTaq® Green Master Mix (Tabla 3).



**Tabla 3***Reactivos empleados en la reacción de cadena de la polimerasa (PCR).*

<b>Reactivo</b>	<b>Concentración</b>	<b>Volumen</b>
Agua libre de ARNasas	-	4.5 $\mu$ L
GoTaq® Green Master Mix	2X	12.5 $\mu$ L
Cebador delantero	20 $\mu$ M	1 $\mu$ L
Cebador reverso	10 $\mu$ M	1 $\mu$ L
ADNc	400 - 600 (ng/ $\mu$ L)	1 $\mu$ L

En la Tabla 4 se muestran las condiciones y parámetros de PCR empleadas para la amplificación que se llevó a cabo en el termociclador ProFlex™ PCR System (Applied Biosystems™) durante 40 ciclos para cada uno de los pares de cebadores empleados.

**Tabla 4***Condiciones de PCR para amplificación con primers genéricos.*

<b>Ciclos PCR</b>	<b>Primers Ilar2F5/Ilar2R9</b>	<b>Primers Ilar1F5/Ilar1R7</b>
Desnaturalización inicial	94°C por 5 min	94°C por 5 min
Desnaturalización	94°C por 30 s	94°C por 30 s
Hibridación	44°C por 1 min	47.7°C por 30 s
Extensión	72°C por 1 min	72°C por 30 s
Extensión final	72°C por 10 min	72°C por 10 min
Número de ciclos	40	40

Fuente: (Untiveros et al., 2010).

**3.3.5.1. Electroforesis**

Los productos de PCR se visualizaron mediante electroforesis en gel de agarosa al 2% preparado con 40 mL de buffer TAE 1X y 10 µL de visualizador SYBR GOLD (Invitrogen, USA).

Se usó el marcador molecular de 100 bp DNA Ladder (Thermo Scientific). La electroforesis se realizó a 100 V durante 30 min, posteriormente se colocó el gel en el equipo High Performance UV – transiluminador y se observó la presencia de bandas.

### **3.3.6. Detección de virus fitopatógenos mediante secuenciación de ADN complementario**

Los productos de PCR de los virus detectados en las muestras de hojas sintomáticas fueron secuenciados por el método Sanger en la compañía ©Macrogen Inc. en Corea del Sur. Se remitió un total de 16 µL por cada amplicón obtenido de la reacción PCR y 15 µL de cada cebador a una concentración de 10 nM.

Se utilizó el software bioinformático Geneious v11.1.5 para analizar y ensamblar las secuencias obtenidas de la secuenciación. Se realizó una búsqueda de homólogos utilizando la herramienta BLASTN (Nucleotide Basic Local Alignment Search Tool) del NCBI (National Center for Biotechnology Information).

### **3.3.7. Identificación de virus fitopatógenos mediante secuenciación masiva paralela – Illumina**

Las extracciones obtenidas de ARN total, fueron precipitadas con 2 volúmenes de etanol 100% y 0,1 volúmenes de acetato de sodio 3M (NaOAc, pH 5,5) para mejorar las condiciones de conservación y fueron enviadas a la compañía ©Macrogen Inc. en Corea del Sur para ser secuenciadas por el sistema de secuenciación NovaSeq 6000 de la plataforma Illumina, se envió un total mínimo de 50 µL con una concentración de 3 – 5 µg por cada muestra de ARN extraído.

Se realizó un control de calidad al ARN de las muestras enviadas y posteriormente se llevó a cabo un proceso de ribosomal RNA depletion, para eliminar los ARNr de la muestra reduciendo la contaminación de ribosomas e incrementando el porcentaje de lecturas asignadas para ARNm y ARN viral. La librería de secuenciación se preparó usando TruSeq Stranded Total RNA LT Sample Prep Kit (Plant) (Macrogen, 2018a).

Se ejecutó la secuenciación masiva paralela usando la tecnología de secuenciación por síntesis (SBS, *Sequencing by Synthesis*). Se generaron secuencias cortas (*reads*) mediante secuenciación de extremo pareado (*paired end*) de cada extremo de los fragmento de ADNc (Macrogen, 2018a).

#### **3.3.7.1. Análisis de datos**

Las secuencias cortas (*reads*) conseguidas a partir de los extremos de ADNc fueron enviados por ©Macrogen Inc. en formato FastQ, y se utilizaron como archivos de entrada para el *pipeline* bioinformático para detección y ensamblaje de genomas de virus fitopatógenos.

Los *reads* fueron analizados mediante un *pipeline* que empleó varias herramientas bioinformáticas. Se utilizó el software Trimomatic-0.38 para remover las secuencias de los adaptadores usados en la creación de la librería de secuenciación y recortar las lecturas de baja calidad (Phred <30). Posteriormente,

se realizó la eliminación de lecturas duplicadas para naranjilla y para limón Meyer a través de la herramienta dedupe incluido en el paquete BBTools.

Las secuencias pertenecientes al genoma de la planta (*host*) fueron eliminadas a través de la herramienta Bowtie2 - 2.3.3.1, que se utilizó para crear una base de datos con el genoma de referencia más cercano a las muestras de estudio (mandarina para limón Meyer y papa para naranjilla). Los *reads* libres de duplicados fueron alineados con las secuencias de referencia. El archivo de extensión .sam (Sequence Alignment Map) obtenido con los datos de secuencia, tanto alineados como no alineados; fue transformado a un archivo de extensión .bam (versión binaria comprimida de almacenamiento) mediante el conjunto de programas Samtools – 1.7. A continuación, se filtraron las lecturas no alineadas y se empleó la herramienta bioinformática Bedtools2 para separar las lecturas de extremo pareado en dos archivos FASTQ.

Las *reads* filtrados fueron ensamblados usando SPAdes-3.11.1, Velvet\_1.2.10 y ABySS-2.1.5 con diferentes valores de k-mers (21, 33, 55 y 77). Las lecturas ensambladas en *contigs* fueron comparadas con secuencias de genomas virales de bases de datos BLAST (Altschul, Gish, Miller, Myers, & Lipman, 1990) utilizando los registros RefSeq de virus de GenBank (Viral Genome) y un database no curada (All virus) (<http://www.ncbi.nlm.nih.gov/genome/viruses/>) con un porcentaje de identidad de 70% y un query coverage de 80%.

El resultado del alineamiento se visualizó a través de la herramienta Krona tools–2.7 (<https://github.com/marbl/Krona/wiki/Installing>).

Para determinar secuencias falsas positivas y comprobar los resultados del análisis de identificación de virus fitopatógenos obtenidos, se importaron los *contigs* generados a Geneious v11.1.5 y se llevó a cabo búsquedas Blastn .

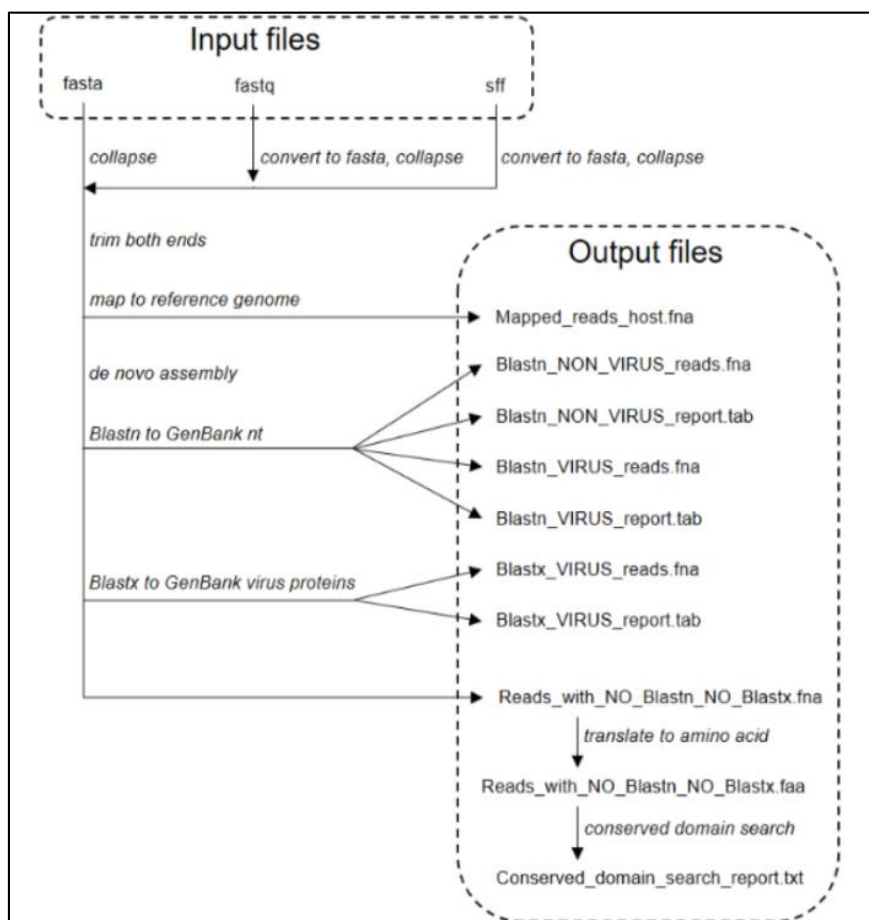
### **3.3.7.2. Comparación de ensambladores de genomas virales**

En este estudio, se evaluaron y compararon los resultados del ensamblaje de las secuencias virales realizado mediante los ensambladores de genomas SPAdes-3.11.1, Velvet\_1.2.10 y ABySS-2.1.5 usando tamaños de *k-mers* de 21 (K21), 33 (K33), 55 (K55), 77 (K77); los cuales fueron comparados a través de la herramienta QUASt-4.6.3.

Se utilizó RStudio, un entorno dedicado al cálculo estadístico y la elaboración de gráficos, para comparar mediante gráficos de barra agrupados los datos obtenidos en el análisis QUASt.

### 3.3.7.3. Detección de virus fitopatógenos mediante el uso de la herramienta bioinformática automatizada VirFind

Los datos de RNA-Seq obtenidos de la secuenciación masiva paralela, fueron analizados con la herramienta bioinformática automatizada VirFind que se encuentra disponible públicamente y realiza la detección e identificación de virus a partir de datos obtenidos de secuenciación masiva paralela. El pipeline bioinformático usado por VirFind acepta archivos de entrada en formatos sff (454), fastq (Illumina) y fasta, e incluye varias herramientas de manipulación de secuencias (Figura 4). Inicialmente, los archivos de entrada son transformados a formato fasta, para luego proceder a recortar un número  $n$  de nucleótidos de los dos extremos ( $n$  es determinado por el usuario). A través de la herramienta Bowtie2, realiza el mapeo de los *reads* con un genoma de referencia (proporcionado por el usuario previamente) para eliminar las secuencias del hospedero, las secuencias no mapeadas obtenidas en este proceso son ensambladas mediante el ensamblador de genomas Velvet. Finalmente, VirFind compara los *contigs* resultantes con secuencias de genomas virales disponibles en la base de datos GenBank mediante BLASTN, las secuencias no detectadas por BLASTN son mapeadas con las secuencias de proteínas de virus del GenBank a través de BLASTX. Los archivos generados son proporcionados al usuarios, quien utilizarán la información de los archivos Blastn y Blastx para determinar si existen virus presentes en su muestra (Ho & Tzanetakis, 2014a).



**Figura 4.** Diagrama de trabajo de VirFind para detección de virus, utilizando datos de secuenciación masiva paralela.

Fuente: (Ho & Tzanetakis, 2014b).

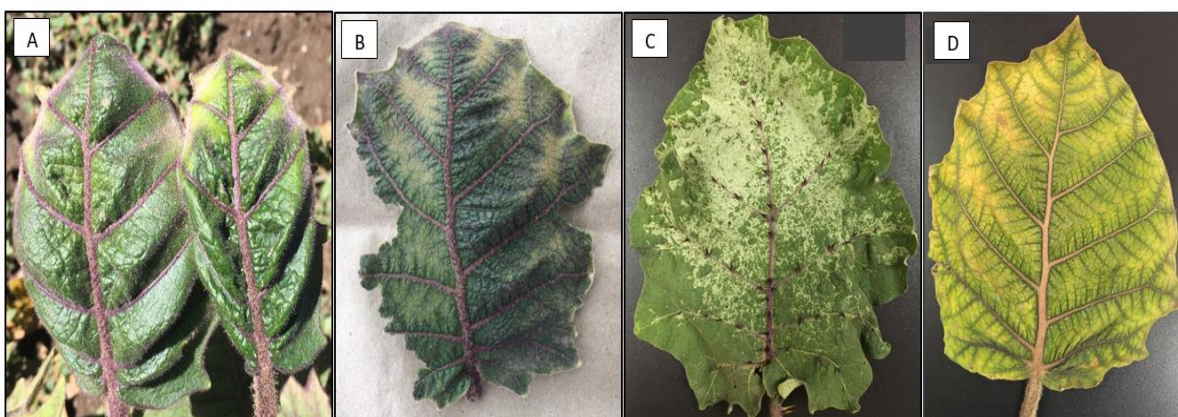


## CAPÍTULO IV

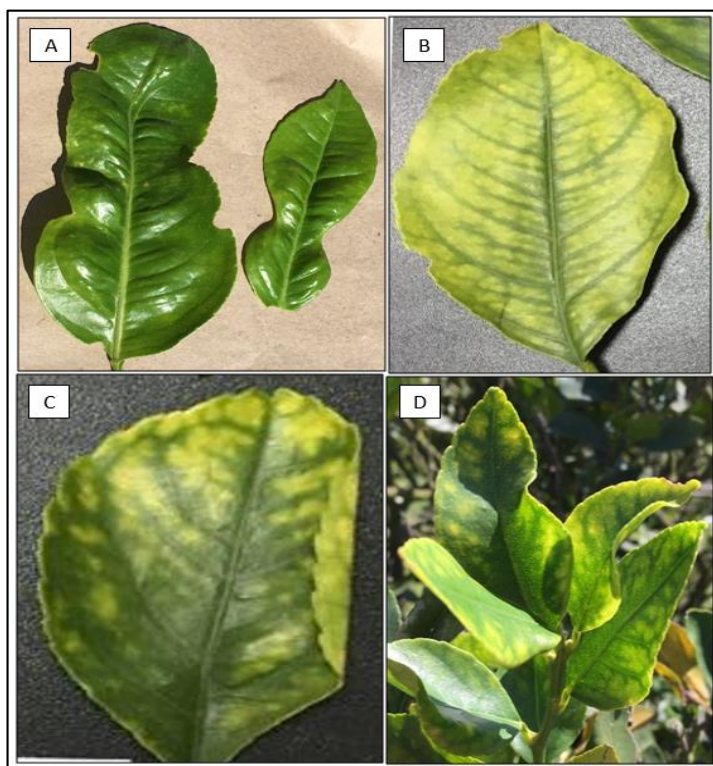
### RESULTADOS

#### 4.1. Caracterización de la zona de estudio y obtención de muestras

Se establecieron dos zonas de muestreo dentro de la Granja experimental del Instituto Nacional de Investigaciones Agropecuarias (INIAP) una para cultivos de naranjilla y otra para cultivos de limón Meyer. Se realizó un total de seis muestreos, tres en cada zona, donde se recolectaron muestras aleatorias de tejido vegetal (hojas) con síntomas víricos como ampollamiento, mosaico, amarillamiento intervenal, manchas cloróticas, deformación de brotes y tejido foliar en los cultivos de naranjilla (*Solanum quitoense*) (Figura 5) y limón Meyer (*Citrus x meyeri*) (Figura 6). La temperatura durante los muestreos osciló entre 17 y 22°C.



**Figura 5.** Hojas de plantas de naranjilla con sintomatología vírica. A) Deformación de brotes y ampollamiento; B) Deformación de tejido foliar y mosaico; C) Clorosis; D) Amarillamiento intervenal.



**Figura 6.** Hojas de plantas de limón Meyer con sintomatología vírica. A) Deformación de brotes y ampollamiento; B) Amarillamiento intervenal; C) y D) Clorosis.

#### 4.2. Extracción de ARN total y transcripción inversa

Con el ARN total extraído se realizó la retrotranscripción para obtener ADN complementario (ADNc). La pureza de las muestras fue valorada utilizando la relación de absorbancias  $A_{260}/280$  y  $A_{260}/230$ .

En la Tabla 5 se presenta los valores de concentración y pureza obtenidos. Las muestras con valores de concentración superior a  $300 \text{ ng}/\mu\text{L}$  y radios de absorbancia dentro del rango de 1.6 a 1.8 (Pureza aceptable) y 1.8 a 2 (Pureza

óptima) para el radio A260/280 (relación más estable), fueron utilizadas en las reacciones de PCR posteriores.

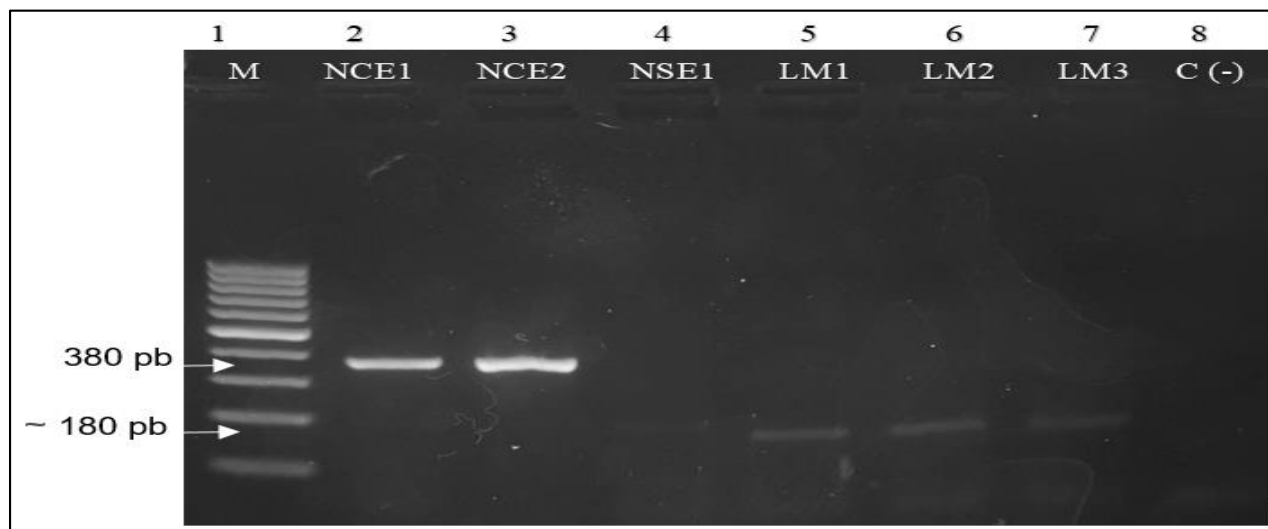
**Tabla 5**

*Valores de concentración y pureza de muestras ADNc de naranjilla y limón Meyer.*

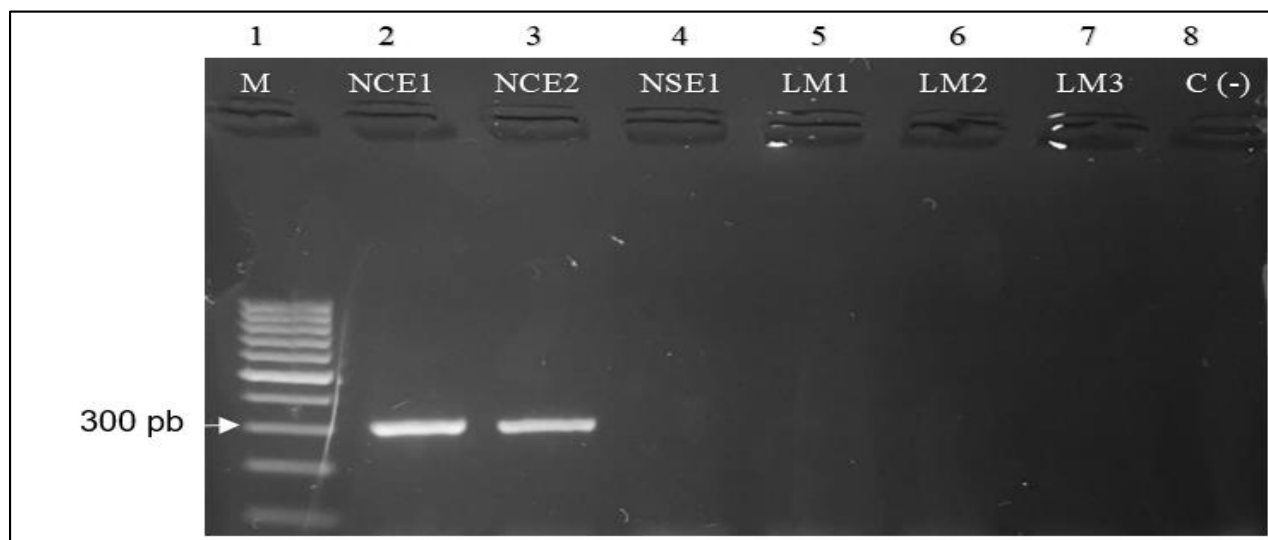
Muestra	Concentración	Absorbancia	Absorbancia 260/230
	ng/μL	260/280	
NCE 1	1466.4	1.765	2.050
NCE 2	887.18	1.645	1.834
NSE1	443.25	1.634	1.822
LM1	307.57	1.688	1.917
LM2	1733.5	1.605	1.937
LM3	11.456	1.709	1.148

#### 4.3. Amplificación de muestras de ADNc mediante PCR convencional

Se realizó la amplificación de las muestras de ADNc usando los pares de primers Ilar2F5/Ilar2R9, Ilar1F5/Ilar1R7, descritos anteriormente. Se obtuvieron fragmentos de tamaño esperado de 380 pb y 300 pb para cada par de cebadores en las muestras de naranjilla (NCE1, NCE2) (Figura 7 y Figura 8); mientras que para las muestras de limón (LM1, LM2, LM3) no se obtuvieron los resultados esperados, únicamente se obtuvieron bandas inespecíficas de aproximadamente 180 pb con uno de los pares de primers como se puede observar en la Figura 7.



**Figura 7.** Electroforesis en gel de agarosa al 2% donde se visualiza productos PCR de muestras de tejido vegetal con sintomatología vírica, amplificadas con los cebadores Ilar2F5 – Ilar2R9 para Ilarvirus. Pocillo 1: Marcador 100 bp (Promega), Pocillo 2 -4: Muestras de naranja, Pocillo 5 -7: Muestras de limón Meyer y Pocillo 8: control negativo.



**Figura 8.** Electroforesis en gel de agarosa al 2% donde se visualiza productos de PCR amplificadas con los cebadores Ilar1F5 – Ilar1R7 para Ilarvirus. Pocillo 1: Marcador 100 bp (Promega). Pocillo 2 -4: Muestras de naranja, Pocillo 5 -7: Muestras de limón Meyer y Pocillo 8: control negativo.

#### 4.4. Análisis de resultados de secuenciación Sanger

La búsqueda de homólogos utilizando la herramienta BLASTN (Nucleotide BLAST) permitió identificar en muestras de naranjilla al virus PYV (*Potato yellowing virus*) con un porcentaje de identidad de 97.63% y 97.65% como se muestra en la Tabla 6. En muestras de hojas sintomáticas de limón Meyer no se pudo identificar virus del género *Illavirus*, la búsqueda de homólogos de los bandas inespecíficas obtenidas presentó un porcentaje de identidad de 97.99% para ADN de cloroplasto.

**Tabla 6**

*Resultados del análisis de secuencias con la herramienta BLASTN.*

<b>Muestra</b>	<b>Género</b>	<b>Especie</b>	<b>Query Coverage</b>	<b>% de Identidad</b>	<b>Acceso GenBank</b>
NCE1	<i>Illavirus</i>	<i>Potato</i>	99 %	97.63	MH937419.1
NCE2		<i>yellowing virus</i>	100%	97.65	MH937419.1

#### 4.5. Análisis de muestras de ARN total mediante secuenciación masiva paralela

Las muestras de ARN total extraídas fueron enviadas a la compañía ©Macrogen Inc., en Corea; la cual se encargó de realizar el análisis de las muestras, creación de librerías, proceso de secuenciación y proporcionó los datos y resultados generados.

#### 4.5.1. Control de calidad de las muestras

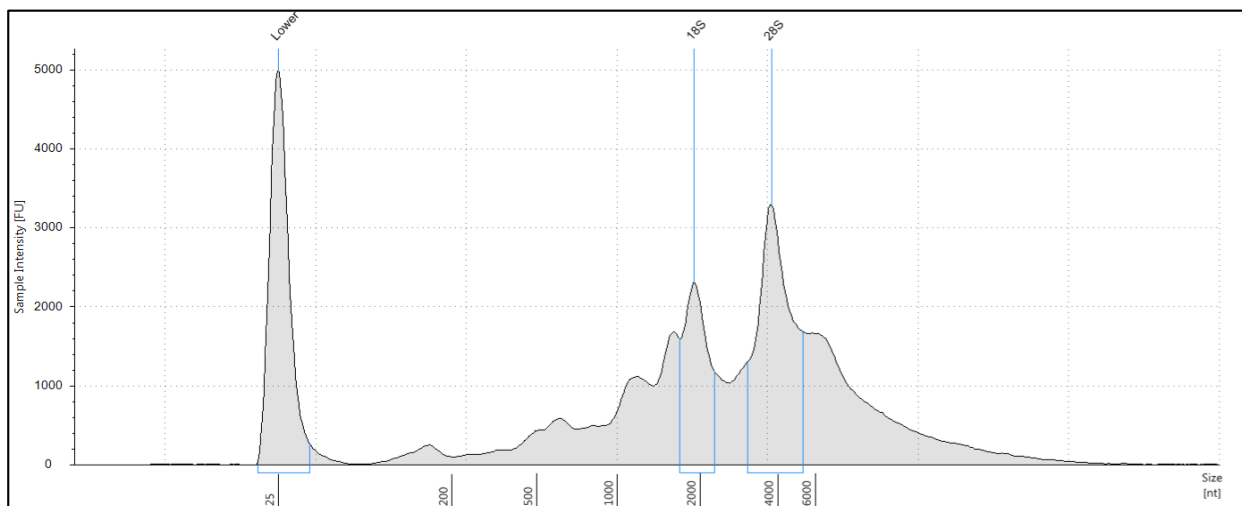
Se realizó un control de calidad (QC, Quality Control) previo, para verificar la calidad de las muestras. En la Tabla 7 se presentan los resultados del QC de ARN, donde se puede observar que el número de integridad del ARN (RIN) (Figura 9 y Figura 10) para las muestras de naranjilla y limón Meyer fue de 5.1 y 4.9 respectivamente, lo que indica que las muestras presentaron una calidad aceptable con una degradación media.

**Tabla 7**

*Resultados del control de calidad realizado a las muestras de RNA total.*

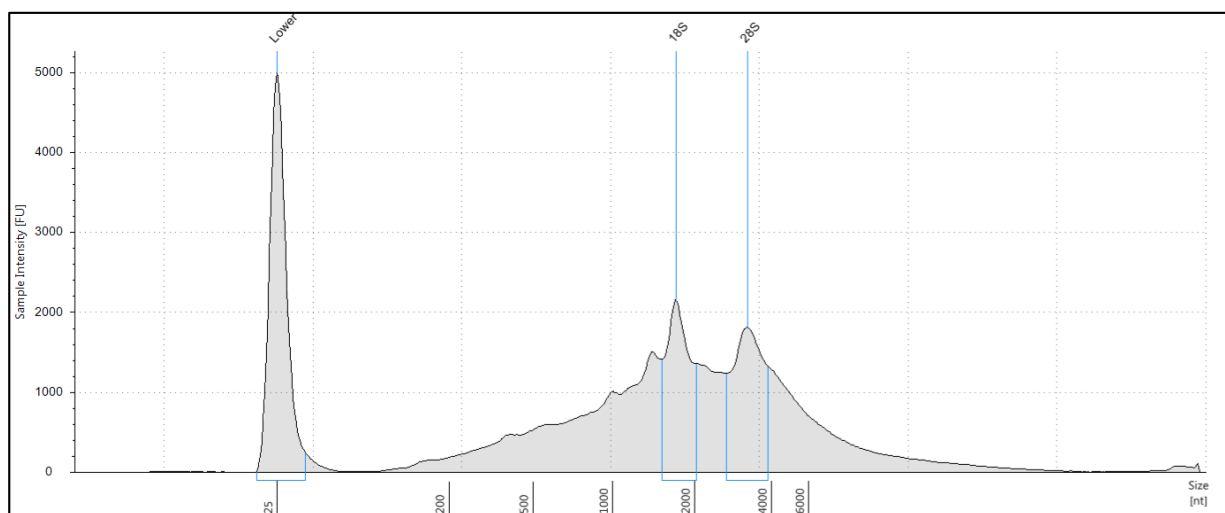
<b>Muestra</b>	<b>Concentración (ng/ <math>\mu</math>L)</b>	<b>Volumen Final (<math>\mu</math>L)</b>	<b>Cantidad Total (<math>\mu</math>g)</b>	<b>RIN</b>
Naranjilla	411.138	25	10.278	5.1
Limón Meyer	88.956	25	2.224	4.9

Fuente: (Macrogen Inc., 2018).



**Figura 9.** Gráfica de las bandas ribosómicas 28S y 18S obtenida en el análisis de calidad de la muestra de naranjilla, mediante el bioanalizador Agilent Technologies 2100 para estimar la integridad total de ARN.

Fuente: Macrogen, Inc. Original Sample QC, 2018.



**Figura 10.** Gráfica de las bandas ribosómicas 28S y 18S obtenida en el análisis de calidad de la muestra de limón Meyer, mediante el bioanalizador Agilent Technologies 2100 para estimar la integridad total de ARN.

Fuente: Macrogen, Inc. Original Sample QC, 2018.

#### 4.5.2. Creación de la Librería

Se llevó a cabo la construcción de la librería de secuenciación usando el TruSeq Stranded Total RNA LT Sample Prep Kit (Plant). Se puede observar en la Tabla 8 el resultado del control de calidad realizado a la librería obtenida proporcionado por MacroGen.

**Tabla 8**  
*Resultado del control de calidad de la librería NGS.*

<b>Muestra</b>	<b>Tipo de Librería</b>	<b>Concentración (ng/ <math>\mu</math>L)</b>	<b>Concentración (nM)</b>	<b>Tamaño (bp)</b>
NaranjaJilla	TruSeq Stranded Total RNA LT	87.91	406.13	333
Limón Meyer	TruSeq Stranded Total RNA LT	44.32	200.55	340

Fuente: (MacroGen Inc., 2018)

#### 4.5.3. Secuenciación masiva paralela

Con la secuenciación de extremo pareado (*paired-end*) de la plataforma Illumina, se obtuvieron en promedio 92'345,458 *reads* para naranjaJilla y 88'747,638 *reads* para limón Meyer de 151 nt cada lectura. En la Tabla 9 se resume la información obtenida de cada muestra como porcentaje de GC y AT (Figura 11) y puntaje de calidad 30.

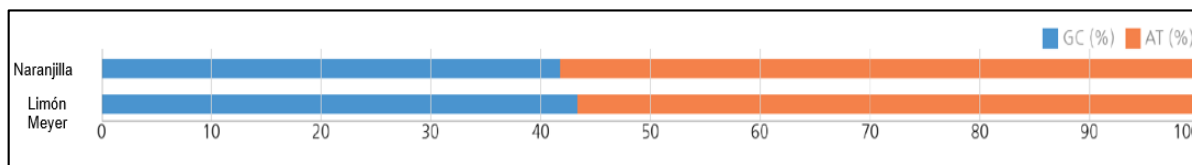


**Tabla 9**

*Datos de la secuenciación masiva paralela realizada por Macrogen Inc.*

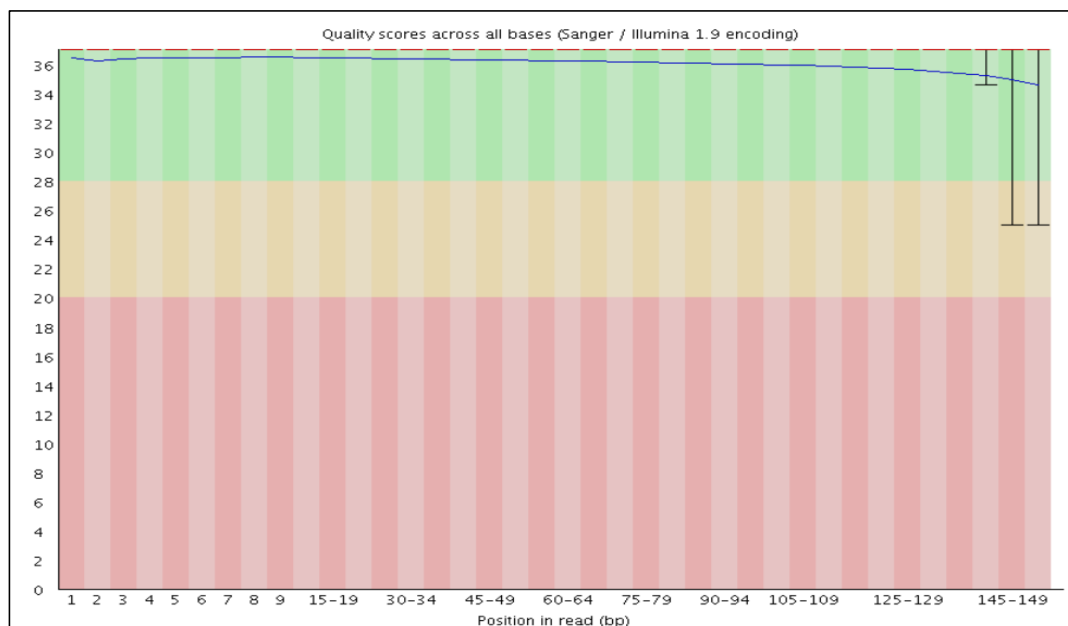
Muestra	Bases Secuenciadas Totales (bp)	Reads Totales	GC%	AT%	Q20(%)	Q30(%)
Naranjilla	13,944,164,158	92'345,458	41.84	58.16	98.42	95.11
Limón Meyer	13,400,893,338	88'747,638	43.34	56.66	98.50	95.39

Fuente: (Macrogen Inc., 2018).

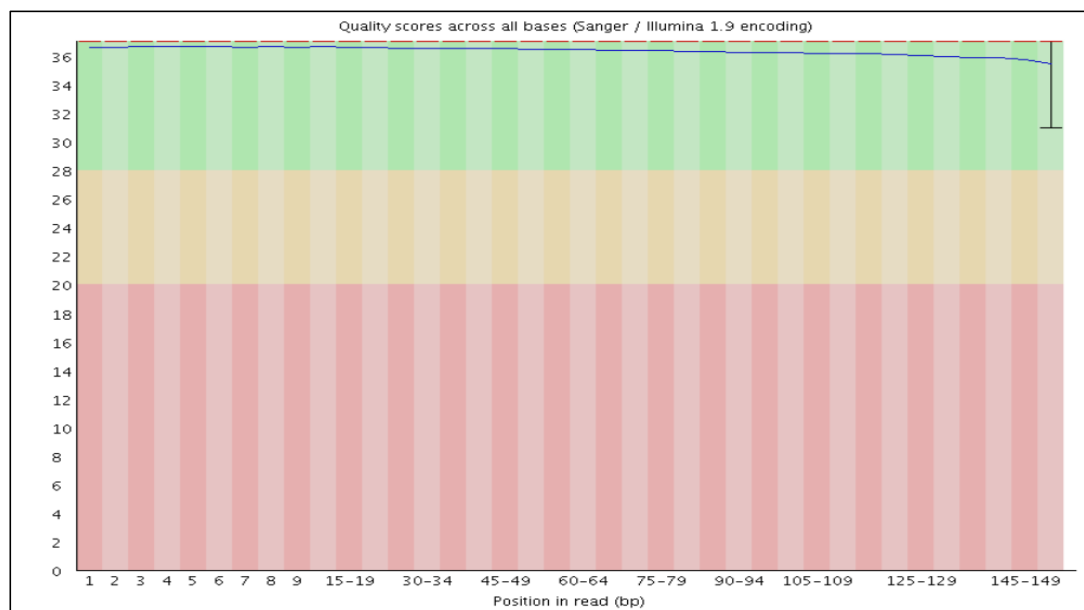


**Figura 11.** Contenido de guanina-citosina y adenina-timina de muestras de naranjilla y limón Meyer. Fuente: (Macrogen Inc., 2018).

En la Figura 12 y 13 se observa la calidad de las secuencias por base analizadas usando la herramienta FastQC v0.11.8. El eje X representa las bases de las lecturas y el eje Y las calidades entre 0 y 36 divididas en zonas de colores. El color verde representa muy buena, el naranja calidad intermedia y el rojo mala calidad. La calidad de las secuencias de las dos muestras se encontró dentro de los parámetros de buena calidad entre 34 y 36 (color verde) para los 151 nt de los *reads*.

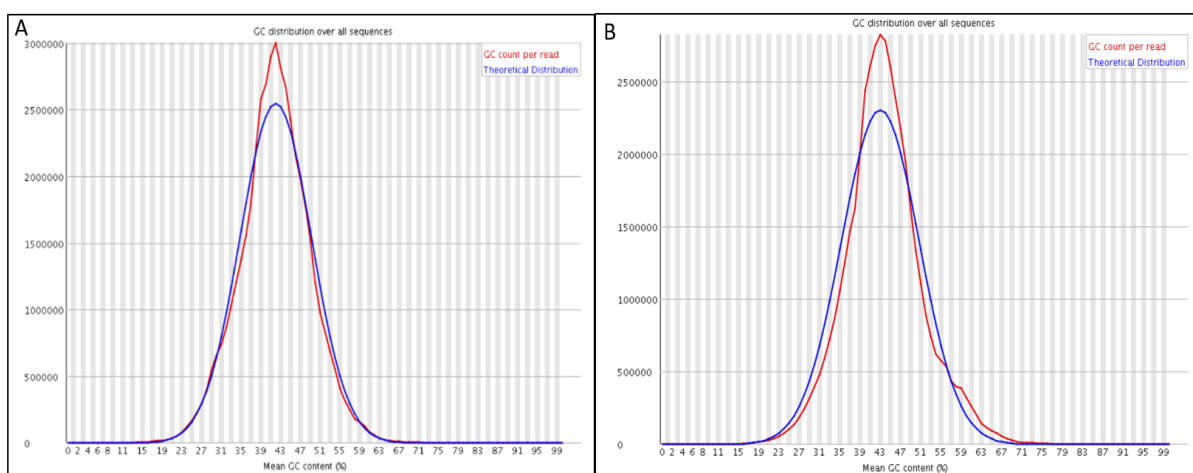


**Figura 12.** Calidad de las secuencias NGS de naranjilla por base.



**Figura 13.** Calidad de las secuencias NGS de limón Meyer por base.

En la Figura 14 se observa en color rojo la curva del contenido de guanina-citosina (GC) y en color azul la curva de distribución normal teórica del contenido GC. En el eje X tenemos el porcentaje medio del contenido de GC y en eje Y el número de *reads*.

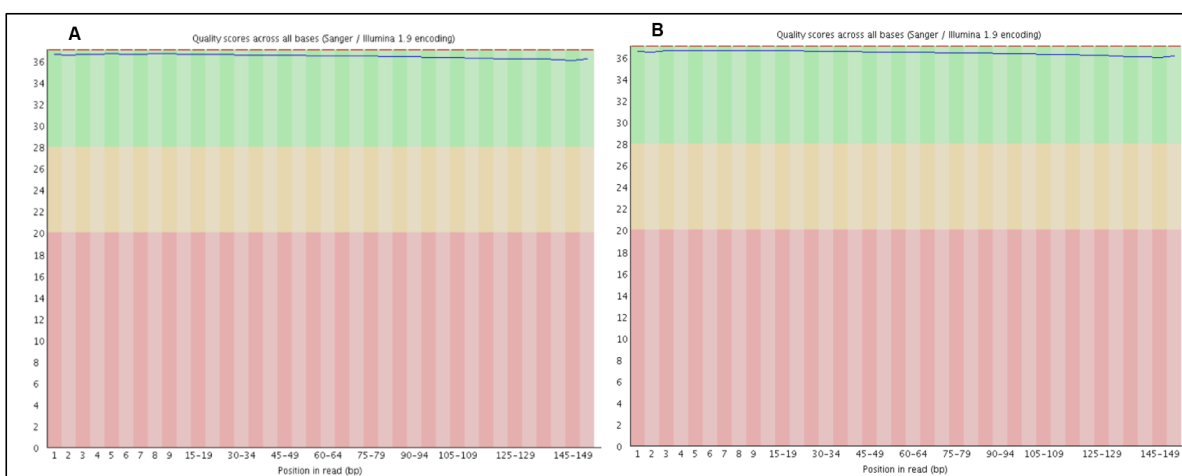


**Figura 14.** Contenido de guanina y citosina (GC) por secuencia. A) Muestra de naranjilla; B) Muestra de limón Meyer.

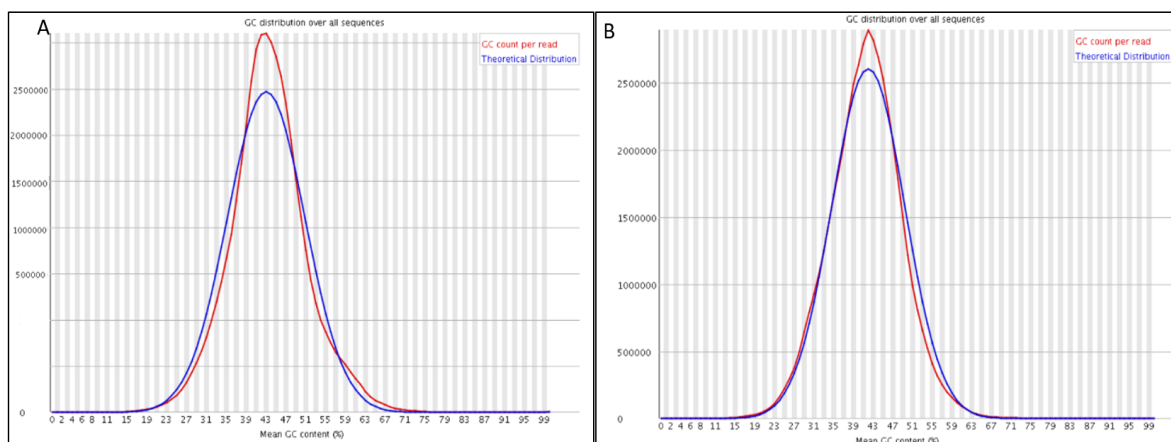
#### 4.6. Análisis bioinformático de lecturas NGS

En el desarrollo del pipeline de identificación de virus fitopatógenos se emplearon varias herramientas y programas bioinformáticos. Se utilizó el software Trimomatic para remover las secuencias de los adaptadores usados en la creación de la librería de secuenciación y recortar las lecturas de baja calidad (Phred <30), obteniéndose 76'156,490 y 69'555,658 *reads* para naranjilla y limón Meyer respectivamente. Una vez realizada la depuración de secuencias de baja calidad y la deduplicación de 27'835,480 *reads* para naranjilla (36.55%) y 23'558,501 *reads*

para limón Meyer (33.87%), se obtuvieron 45'997.157 *reads* (66.13%) para limón Meyer y 48'321,010 *reads* (63.45%) para naranjilla. En las Figuras 15 y 16 se detallan los resultados obtenidos mediante FastQC, en el análisis posterior a la limpieza y eliminación de duplicados de las secuencias; en los que se observa que la calidad y el porcentaje de GC mejoraron considerablemente.



**Figura 15.** Gráficas de calidad de las secuencias después de usar Trimomatic y BBtools. A) Muestra de limón Meyer y B) Muestra de naranjilla.



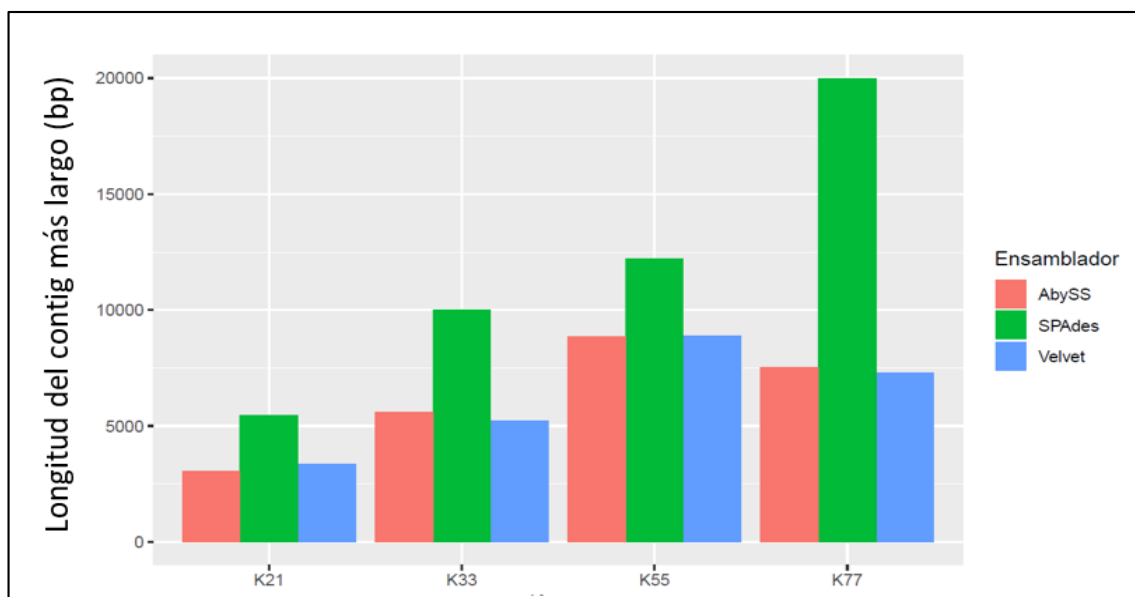
**Figura 16.** Contenido de guanina y citosina (GC) por secuencia, después de usar Trimomatic y BBtools. A) Muestra de limón Meyer; B) Muestra de naranjilla.

Los *reads* libres de duplicados fueron alineados con las secuencias de referencia de los genomas de las plantas hospederas respectivas, observándose una tasa general de alineamiento de 11.14% para naranjilla y 15.33% para limón Meyer. Finalmente se obtuvieron 42'938.049 y 38'964,192 *reads* para naranjilla y limón Meyer respectivamente.

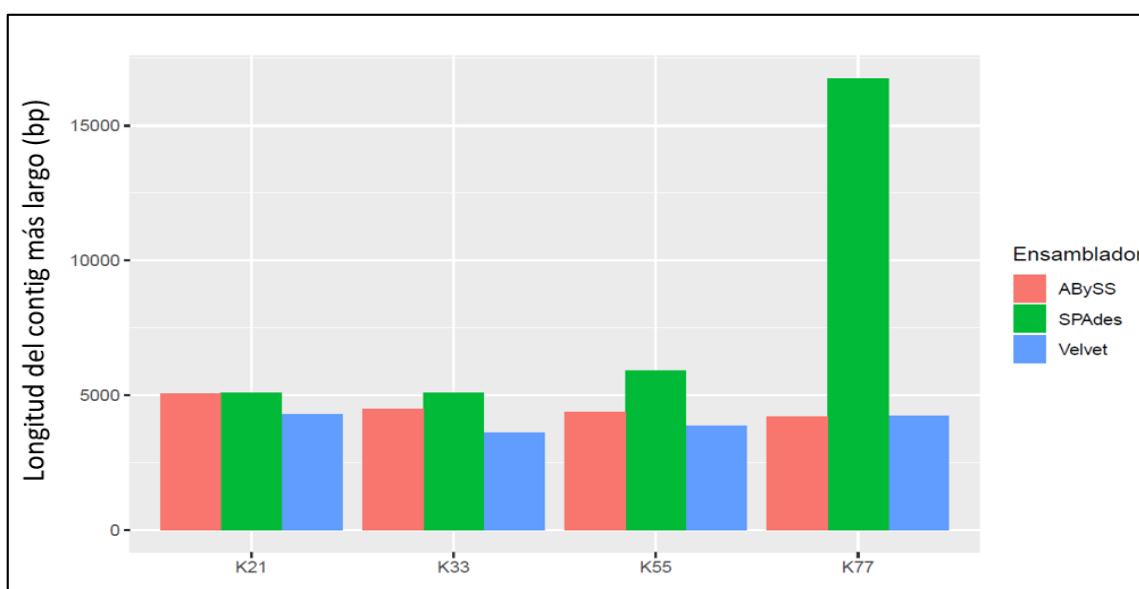
En el ensamblaje realizado con SPAdes-3.11.1 se obtuvieron 32,024 *contigs* para limón Meyer y 134,506 *contigs* para naranjilla. El alineamiento de los *contigs* ensamblados con secuencias de genomas virales de la base de datos del GenBank mostró un total de 25 *contigs* de una longitud de 116 a 16,747 nt para limón Meyer, mientras que para naranjilla se ensamblaron 38 *contigs* con una longitud de 83 a 7,287 nt.

#### **4.7. Comparación de ensambladores**

El análisis comparativo realizado entre los ensambladores (SPAdes, Velvet y ABySS) mostró que SPAdes generó el *contig* más largo con K77 tanto en naranjilla (19,984 bp) (Figura 17) como en limón Meyer (16,747 bp) (Figura 18).

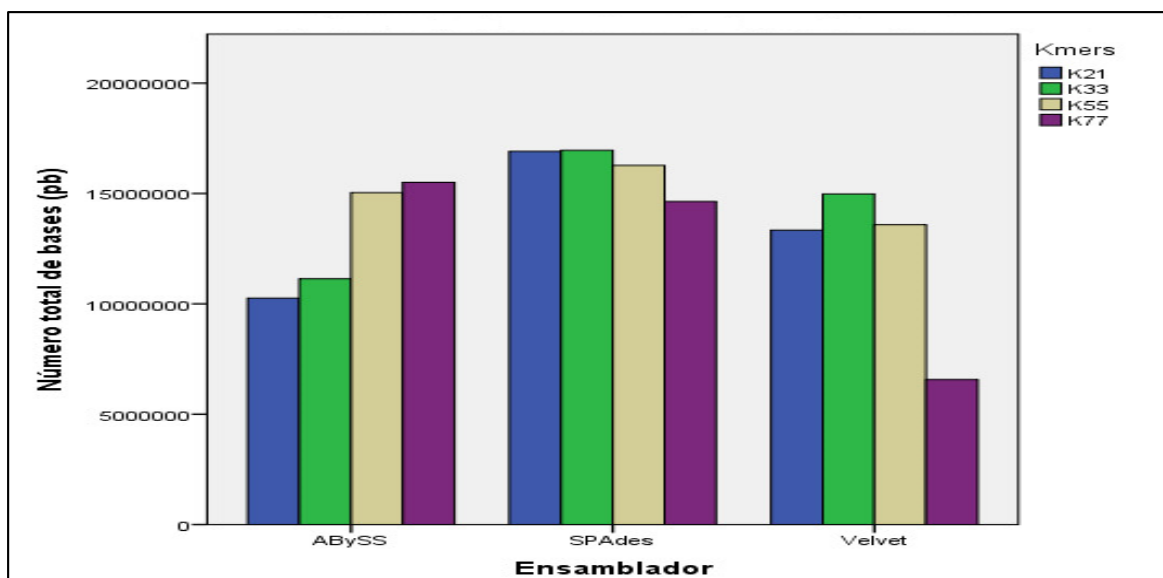


**Figura 17.** Longitud del contig más largo (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras De naranjilla.

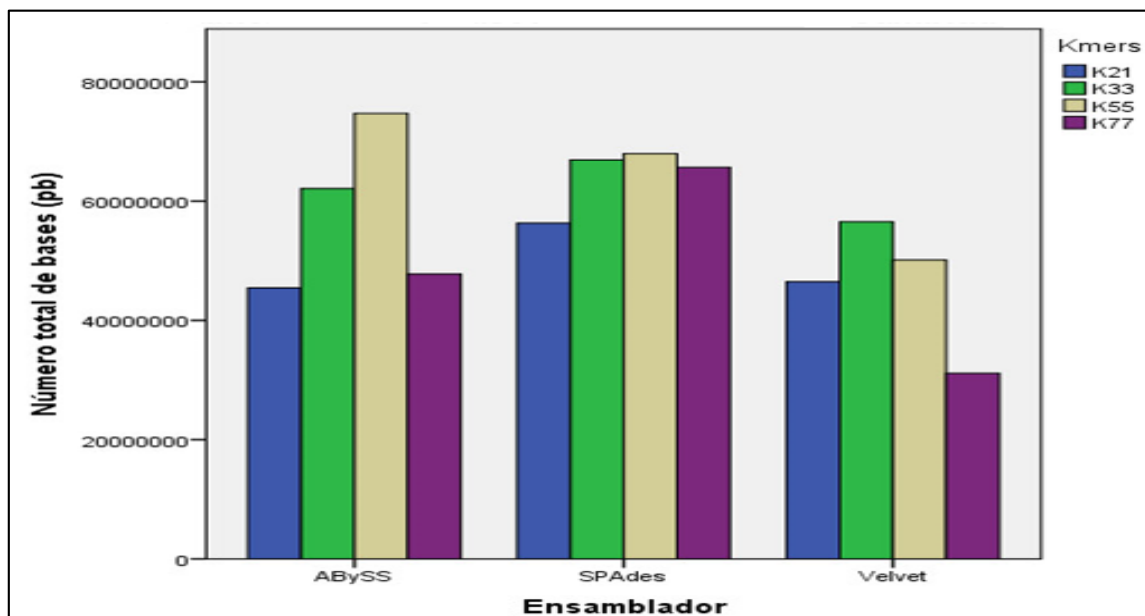


**Figura 18.** Longitud del contig más largo (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras de limón Meyer.

En la Figura 19 y 20 se observa el número total de bases obtenidos en los ensamblajes para las muestras de naranjilla y limón Meyer. Spades presentó el mayor número total de bases con 14'637,417 bp (limón Meyer) y 65'673,696 bp (naranjilla). ABySS tuvo una ligera ventaja sobre Velvet en la muestra de naranjilla, mientras que en limón Meyer tanto ABySS como Velvet presentaron variaciones para los diferentes *k-mers*.



**Figura 19.** Número total de bases en los ensamblajes (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras de limón Meyer.



**Figura 20.** Número total de bases en los ensamblajes (bp) generado con los ensambladores SPAdes, ABySS y Velvet (K21, K33, K55, K77) en muestras de naranjilla.

Velvet y ABySS presentaron *contigs* más diversos; ABySS K55 generó el número de *contigs* totales más alto, 78,553 y 376,985 *contigs* y Velvet K77 el número más bajo con solo 18,232 y 86,826 *contigs* para limón Meyer y naranjilla respectivamente (Tabla 10 y 11).



**Tabla 10***Análisis comparativo de ensambladores de novo en muestras de limón Meyer*

<b>Ensamblador</b>	<b>Tamaño Max de contig (bp)</b>	<b>Tamaño Min de contig (bp)</b>	<b>Número de contigs</b>	<b>Longitud total (bp)</b>	<b>N50</b>	<b>L50</b>
ABySS 21	5,073	28	43,106	10'258,468	269	10,886
ABySS 33	4,500	41	45,578	11'135,816	279	11,179
ABySS 55	4,379	55	78,553	15'038,296	202	20,154
ABySS 77	4,209	87	72,128	15'503,237	208	22,379
SPAdes	16,747	116	32,024	14'637,417	542	7,301
Velvet 21	4,287	26	67,539	13'338,203	210	20,398
Velvet 33	3,624	39	73,910	14'978,328	214	22,312
Velvet 55	3,869	51	57,855	13'584,655	244	17,731
Velvet 77	4,233	77	18,232	6'567,546	373	5,742

Longitud total, número total de bases del ensamblaje; N50, longitud para la cual la colección de todos los *contigs* de esa longitud o más cubre al menos la mitad del ensamblaje; L50, número de *contigs* igual o mayor que N50.

**Tabla 11***Análisis comparativo de ensambladores de novo en muestras de naranjilla*

Ensamblador	Tamaño Max de contig (bp)	Tamaño Min de contig (bp)	Número de contigs	Longitud total (bp)	N50	L50
ABySS 21	3,060	31	230,819	45'448,870	206	67,637
ABySS 33	5,599	45	292,614	62'116,021	227	77,160
ABySS 55	8,845	59	376,985	74'707,313	198	104,325
ABySS 77	7,515	93	190,673	47'758,671	257	57,422
SPAdes	19,984	83	134,506	65'673,696	528	29,015
Velvet 21	3,353	28	197,437	46'475,598	266	54,776
Velvet 33	5,238	45	221,792	56'518,370	294	58,160
Velvet 55	8,875	66	171,761	50'123,875	319	48,941
Velvet 77	7,287	77	86,826	31'147,396	361	28,267

Longitud total, número total de bases del ensamblaje; N50, longitud para la cual la colección de todos los *contigs* de esa longitud o más cubre al menos la mitad del ensamblaje; L50, número de *contigs* igual o mayor que N50.

#### 4.8. Comparación de Databases

En la Tabla 12 y 13 se observa el número de virus totales y virus fitopatógenos identificados en muestras de naranjilla y limón Meyer utilizando dos bases de datos diferentes (All Virus y Viral Genome). La base de datos All Virus generó una mayor cantidad de *contigs* a diferencia de Viral Genomes, sin embargo, no todos los *contigs* obtenidos pertenecían a virus y no todos los virus identificados eran fitopatógenos.

All virus identificó una mayor cantidad de virus fitopatógenos en naranjilla, mientras que en limón Meyer no se observó un variación representativa. A pesar de que All virus identificó una mayor cantidad de virus totales y virus fitopatógenos, los mayores porcentaje de virus fitopatógenos reales que se encuentran presentes en las muestras se obtuvieron con la base de datos Viral Genome y el ensamblador SPAdes.

**Tabla 12**

*Virus fitopatógenos identificados en muestra de naranjilla con SPAdes y ABySS y diferentes databases.*

	ABYSS									
	SPAdes		<i>k-mer 21</i>		<i>k-mer 33</i>		<i>k-mer 55</i>		<i>k-mer 77</i>	
	Viral Genomes	All Viruses	Viral Genomes	All Viruses	Viral Genomes	All Viruses	Viral Genomes	All Viruses	Viral Genomes	All Viruses
Número total de <i>contigs</i> generados	38	80	73	142	169	269	167	268	144	261
Número de <i>contigs</i> virales	38	60	73	113	169	251	167	259	144	260
<i>Contigs</i> virales (%)	100	75	100	80	100	93	100	97	100	100
Virus totales identificados	7	9	14	24	8	20	6	11	9	12
Virus fitopatógenos identificados	7	6	8	12	6	9	5	6	6	7
Virus fitopatógenos reales presentes en la muestra (%)	85.71	83.33	87.5	50	83.33	66.67	80	66.67	83.33	57.14

**Tabla 13**

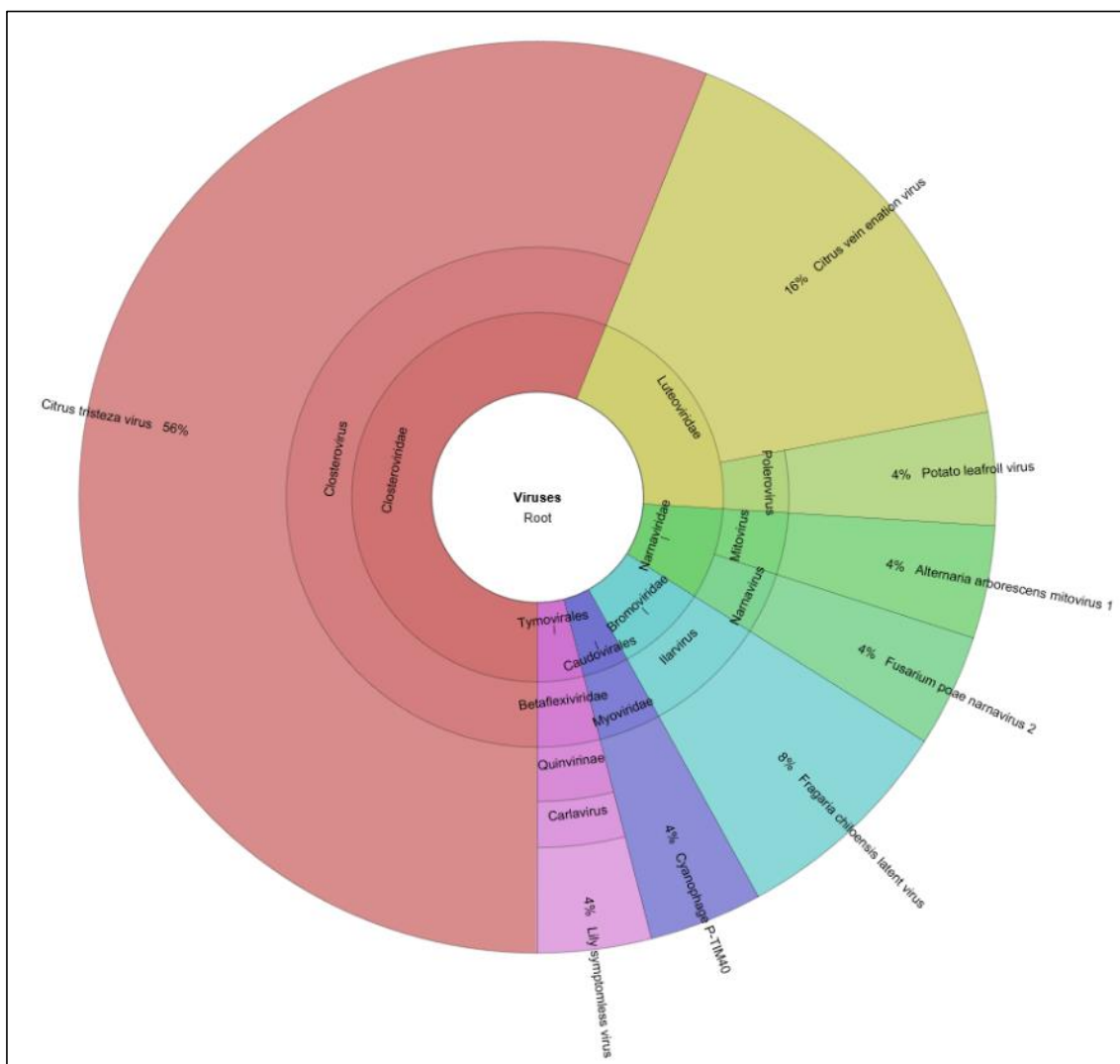
*Virus fitopatógenos identificados en muestra de limón Meyer con SPAdes y ABySS y diferentes databases.*

	ABySS									
	SPAdes		<i>k-mer 21</i>		<i>k-mer 33</i>		<i>k-mer 55</i>		<i>k-mer 77</i>	
	Viral Genomes	All Viruses	Viral Genomes	All Viruses	Viral Genomes	All Viruses	Viral Genomes	All Viruses	Viral Genomes	All Viruses
Número total de <i>contigs</i> generados	25	105	154	609	249	767	177	465	73	133
Número de <i>contigs</i> virales	25	55	154	527	249	636	177	384	73	117
<i>Contigs</i> virales (%)	100	52	100	87	100	83	100	83	100	88
Virus totales identificados	8	13	12	16	8	12	5	12	2	4
Virus fitopatógenos identificados	5	6	3	4	3	3	3	5	2	2
Virus fitopatógenos reales presentes en la muestra (%)	80	66.67	66.66	50	66.66	66.66	66.66	40	100	100

#### 4.9. Identificación de virus fitopatógenos

Del total de *contigs* ensamblados con SPAdes en limón Meyer el 56% (14 *contigs*) corresponden *Citrus tristeza virus*, seguido por *Citrus vein enation* con 16% (4 *contigs*) y *Fragaria chiloensis latent virus* 8% (2 *contigs*). Además se pudo identificar en menor proporción la presencia de virus como *Potato leafroll virus*, *Lily symptomless virus* con porcentajes de 4% (1 *contig*) cada uno. El 8% (2 *contigs*)

pertencen a virus de la familia *Narnaviridae* que infectan exclusivamente a hongos y el 4% restante (1 contig) corresponde a la familia de bacteriófagos *Myoviridae* (Figura 16)



**Figura 21.** Virus detectados e identificados en muestra de tejido vegetal de limón Meyer.

Sin embargo, al importar los *contigs* generados a Geneious v11.1.5 y realizar una búsquedas Blastn se obtuvo que los *contigs* correspondientes a *Fragaria chiloensis latent virus* presentaron un porcentaje de identidad de 97.14% con la secuencia del segmento 3 de RNA de *Potato yellowing virus* reportado en el Genbank (Tabla 14), lo cual nos indicaría la presencia de un falso positivo en la identificación de virus fitopatógenos obtenida. El porcentaje de identidad de nucleótidos obtenido para los otros virus identificados en el análisis se presentan en la Tabla 14.

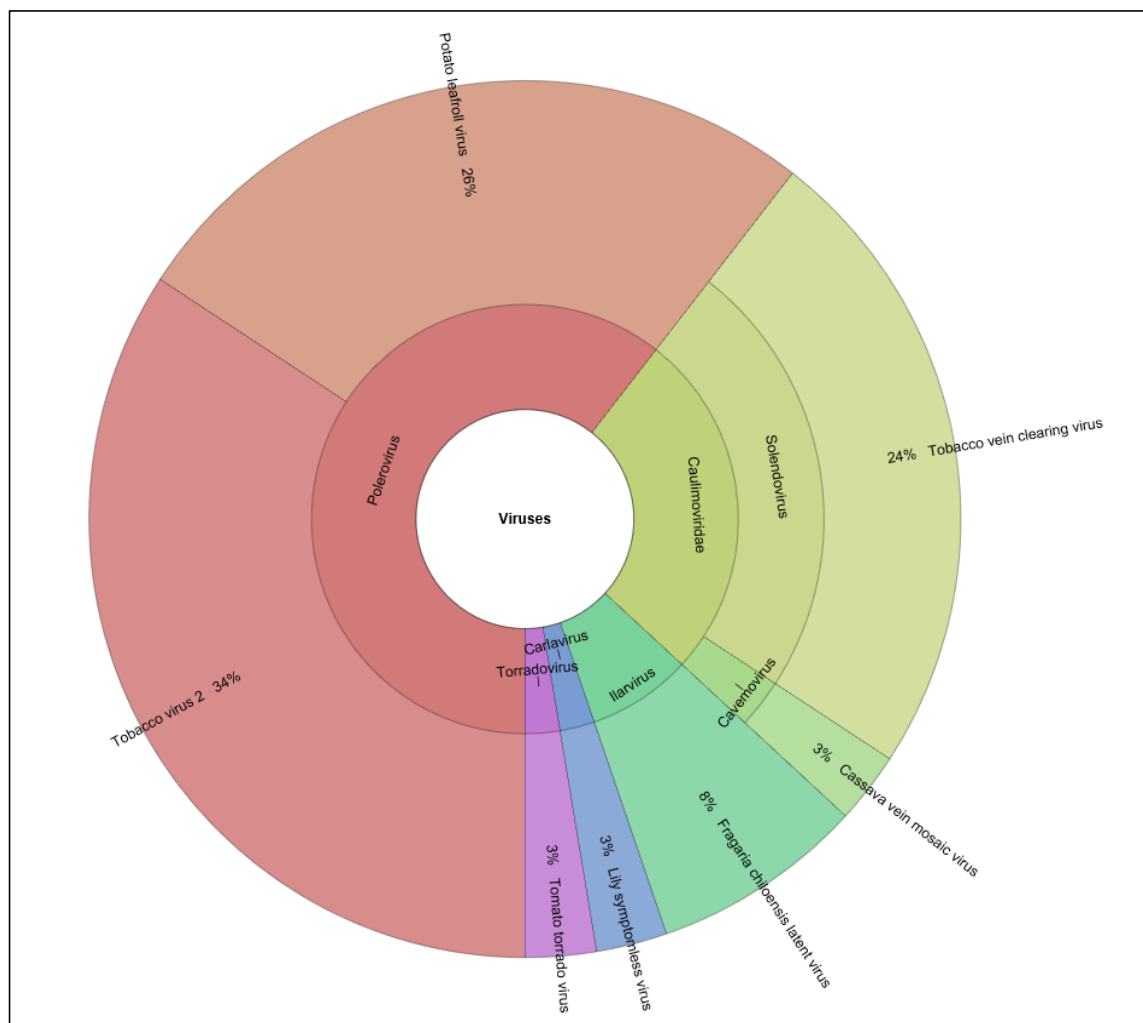
Se obtuvo el genoma casi completo de *Citrus tristeza virus* (19,195 bp) y el genoma completo de *Citrus vein enation* (6,103 bp) mediante el ensamblaje realizado con SPAdes. Mientras que, para el resto de virus fitopatógenos identificados en el análisis se obtuvo identidades superiores a 90% con secuencias de los genomas virales correspondientes presentes en el GenBank.

**Tabla 14**

Resultados del análisis de contigs con la herramienta BLASTN en muestra de limón Meyer.

Género	Especie	Número de hits	Query Coverage	% de identidad de nucleótidos	Número de accesión
<i>Closterovirus</i>	<i>Citrus tristeza virus</i>	14	100%	99.96	KY110738.1
<i>Luteovirus</i>	<i>Citrus vein enation virus</i>	4	97%	98.48	LC433634.1
<i>Polerovirus</i>	<i>Potato leafroll virus</i>	1	100%	93.92	KX073467.2
<i>Ilarvirus</i>	<i>Potato yellowing virus</i>	2	87%	97.14	MH937420.1
<i>Carlavirus</i>	<i>Lily symptomless virus</i>	1	98%	99.57	MK649770.1

En naranjilla se identificaron principalmente virus como *Tobacco virus 2* con 36% (14 contigs), *Potato leafroll virus* con 26% (10 contigs), *Tobacco vein clearing virus* con 21% (8 contigs) y *Fragaria chiloensis latent virus* con 8% (3 contigs). También se detectaron en la muestra otros virus en menor proporción (3%) entre ellos *Cassava vein mosaic virus*, *Lily symptomless virus*, y *Tomato torrado virus* (Figura 17).



**Figura 22.** Virus detectados e identificados en muestra de tejido vegetal de naranjilla.

Al igual que en limón Meyer, se obtuvo que los *contigs* correspondientes a *Fragaria chiloensis latent virus* presentaron un porcentaje de identidad de 98.73% con la secuencia del segmento 3 de RNA de *Potato yellowing virus* reportado en el GenBank (Tabla 15). El porcentaje de identidad de nucleótidos obtenido para *Tobacco virus 2*, *Potato leafroll virus*, *Potato yellowing virus*, *Lily symptomless virus*,



*Tomato torrado virus* fue superior a 95% como se muestra en la Tabla 15, mientras que para *Tobacco vein clearing virus* y *Cassava vein mosaic virus* el análisis de homología presentó porcentaje de identidad de 81.70% y 78.24% respectivamente, valores que indicaron que no existe una alta homología entre la secuencia obtenida en el análisis y la secuencia del virus identificado.

**Tabla 15**

*Resultados del análisis de contigs con la herramienta BLASTN en muestra de naranjilla.*

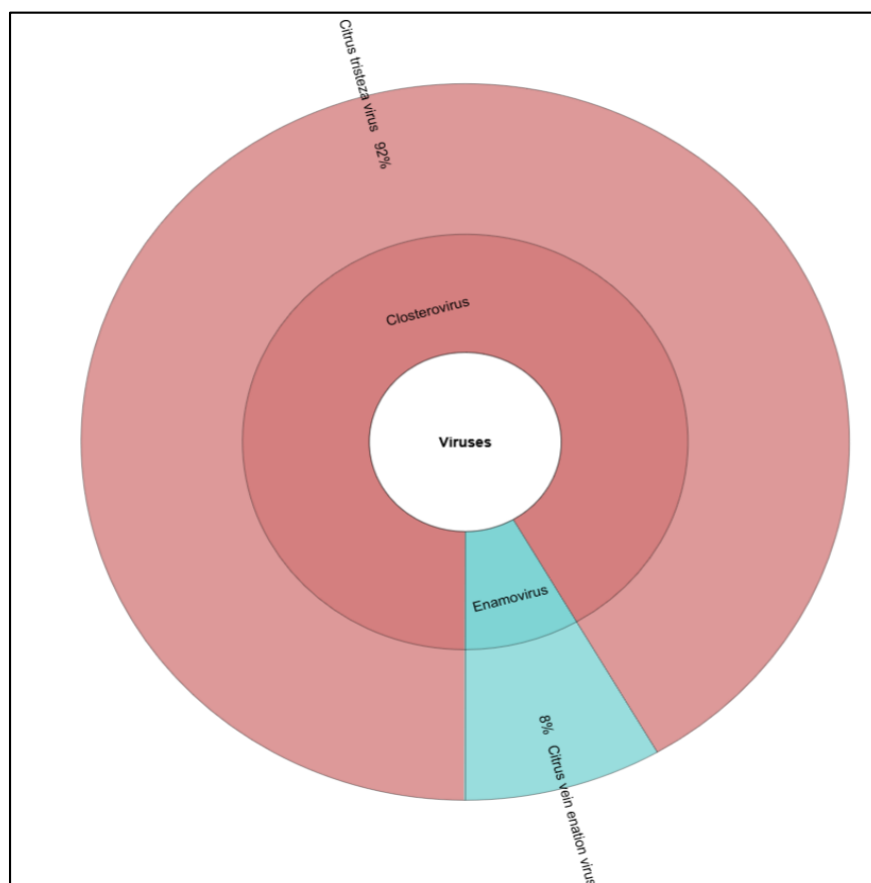
Género	Especie	Número de hits	Query Coverage	% de identidad de nucleótidos	Descripción
<i>Polerovirus</i>	<i>Tobacco virus 2</i>	13	97%	95.55	KY038943.1
<i>Polerovirus</i>	<i>Potato leafroll virus</i>	10	99%	96.72	JQ420902.1
<i>Solendovirus</i>	<i>Tobacco vein clearing virus</i>	9	93%	81.70	AF190123.1
<i>Cavemovirus</i>	<i>Cassava vein mosaic virus</i>	1	89%	78.24	U59751.1
<i>Illavirus</i>	<i>Potato yellowing virus</i>	3	100%	98.73	MH937420.1
<i>Carlavirus</i>	<i>Lily symptomless virus</i>	1	100%	98.80	MK649770.1
<i>Torradovirus</i>	<i>Tomato torrado virus</i>	1	100%	96.17	MH587229.1

No se obtuvo el genoma completo de alguno de los virus fitopatógenos identificados mediante el ensamblaje realizado con SPAdes.

Únicamente se obtuvieron secuencias correspondientes a los genomas de los virus identificados y la secuencia casi completa del segmento 1 de RNA de *Tomato torrado virus* (7,287 bp).

#### **4.10. Detección de virus fitopatógenos a través de VirFind**

Mediante el archivo con secuencias blastn proporcionado por VirFind se identificó secuencias virales en las 2 muestras. En limón Meyer se obtuvieron 24 *contigs* de una longitud de 306 a 19,258 nt en los cuales se identificó únicamente 2 especies de virus fitopatógenos *Citrus tristeza virus* (92%) y *Citrus vein enation* (8%) (Figura 23), con porcentajes de identidad de 99.65% y 98.46% respectivamente (Tabla 16). Se obtuvieron genomas casi completos para *Citrus tristeza virus* (19,245 bp) y *Citrus vein enation*. (5,922 bp).



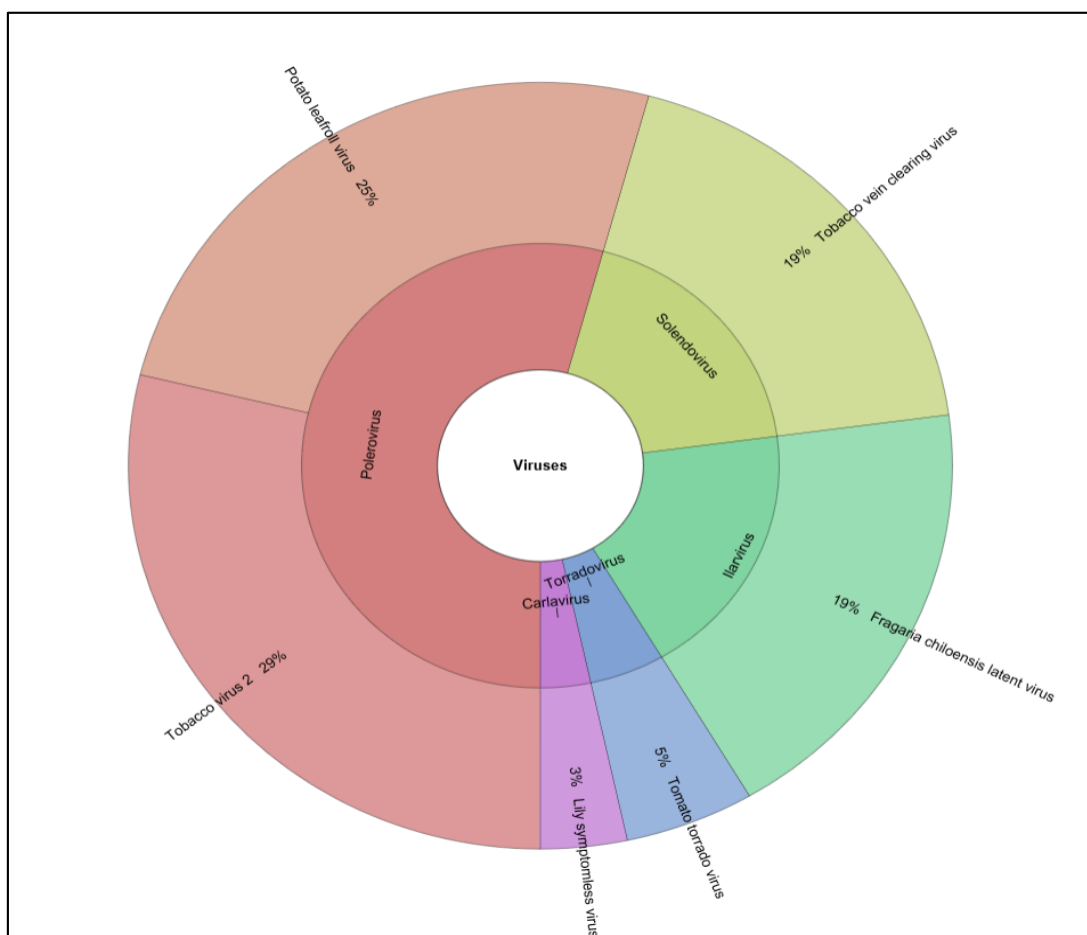
**Figura 23.** Virus detectados e identificados en muestra de tejido vegetal de limón Meyer a través de VirFind.

**Tabla 16**

Resultados del análisis de contigs con la herramienta BLASTN en muestra de limón Meyer.

Género	Especie	Número de hits	Query Coverage	% de identidad de nucleótidos	Número de accesión
<i>Closterovirus</i>	<i>Citrus tristeza virus</i>	22	99%	99.65	KU358530.1
<i>Carlavirus</i>	<i>Citrus vein enation virus</i>	2	100%	98.46	LC433634.1

En naranjilla se obtuvieron 59 *contigs* de una longitud de 83 a 7,636 en los cuales se identificó 6 especies de virus fitopatógenos *Tobacco virus 2* con 29% (17 *contigs*), *Potato leafroll virus* con 25% (15 *contigs*), *Tobacco vein clearing virus* con 19% (11 *contigs*), *Fragaria chiloensis latent virus* con 19% (11 *contigs*). También se detectaron en la muestra otros virus en menor proporción entre ellos *Tomato torrado virus* con 5% (3 *contigs*) y *Lily symptomless virus* con 3% (2 *contigs*) (Figura 24).



**Figura 24.** Virus detectados e identificados en muestra de tejido vegetal de naranjilla a través de VirFind.

En la tabla 17 se observa los porcentajes de identidad obtenidos mediante análisis de homología para los virus detectados en naranjilla a través de VirFind. *Tobacco virus 2* no pudo ser identificado debido a que los *contigs* correspondientes a este virus presentaron un porcentaje de identidad de 96.21% con secuencias del genoma de *Potato leafroll virus* reportado en el GenBank; lo mismo ocurrió en el caso de *Fragaria chiloensis latent virus*, el análisis de homología de los *contigs* obtenidos presentó porcentajes de identidad de 97.17% con la secuencia del genoma del segmento 3 de RNA, 97.65% con la secuencia del genoma del segmento 2 de RNA y 98.66% con la secuencia del genoma del segmento 3 de RNA de *Potato yellowing virus*.

Se obtuvo el genoma completo de *Potato leafroll virus* (5,912 bp) y *Tomato torrado virus* (7,645 bp), además de la secuencia completa del segmento 1 de RNA

**Tabla 17**

Resultados del análisis de contigs con la herramienta BLASTN en muestra de naranjilla.

Género	Especie	Número de hits	Query Coverage	% de identidad de nucleótidos	Descripción
<i>Polerovirus</i>	<i>Tobacco virus 2</i>	-	-	-	-
<i>Polerovirus</i>	<i>Potato leafroll virus</i>	15	99%	96.01	JQ420902.1
<i>Solendovirus</i>	<i>Tobacco vein clearing virus</i>	11	100%	82.29	AF190123.1
<i>Ilarvirus</i>	<i>Potato yellowing virus</i>	11	100%	98.73	MH937420.1
<i>Carlavirus</i>	<i>Lily symptomless virus</i>	2	100%	96.80	MK649770.1
<i>Torradovirus</i>	<i>Tomato torrado virus</i>	3	94%	95.96	MH587229.1

## CAPÍTULO V

### DISCUSIÓN

Los métodos de extracción de ARN y los procesos utilizados antes de realizar una secuenciación masiva paralela, pueden interferir en su eficiencia. El ARN de cadena simple (ssRNA) suele presentar inestabilidad que podría afectar la detección de virus, a diferencia del ARN de doble cadena (dsRNA) que es estable y resistente a la acción de enzimas RNasas, por lo que presenta una tasa de degradación menor (Valverde & De La Torre-Almaraz, 2017). El ARN extraído debe ser de calidad suficiente para poder producir una biblioteca para la secuenciación, la calidad del ARN se mide a través de un número de integridad (RIN) entre 1 y 10. Kukurba & Montgomery (2015) señalan que cuando se obtiene un RIN < 6 el ARN es considerado de baja calidad. Esto; puede afectar los resultados de la secuenciación y generar conclusiones biológicas erróneas. A pesar de que los valores de RIN obtenidos en este caso fueron de 4.9 y 5.1 para las muestras de limón Meyer y naranjilla respectivamente, valores inferiores a seis y fuera del criterio de aceptación de calidad de MacroGen (RIN > 7), la calidad y número de *reads* generados en la secuenciación NGS fueron adecuados para usarlos en el análisis bioinformático de detección e identificación de virus de plantas.

El descubrimiento de nuevos virus fitopatógenos se ha incrementado rápidamente con la introducción de la secuenciación masiva paralela (Pecman et al., 2017). Sin embargo, la secuenciación masiva paralela produce gran cantidad de

datos que implican un desafío a nivel computacional relacionado con el procesamiento y almacenamiento de los mismos (Kaur & Kaur, 2015). Para una detección de virus precisa y confiable es necesario tener una cantidad adecuada de *reads*, que dependerán del tamaño del genoma del virus y la complejidad del viroma. Factores como el grado de infección, edad de la planta y respuesta de la planta a la infección pueden influenciar en la cantidad de *reads* virales recuperados de los datos crudos (Visser, Bester, Burger, & Maree, 2016).

Visser y colaboradores (2016) en su estudio mencionan que el número de lecturas de NGS necesarias para representar un genoma completo de virus a partir de una muestra extraída de material vegetal, específicamente para la detección de *Closterovirus*, es de 1'000,000 *reads* de alta calidad. El número de *reads* obtenidos en la secuenciación de la muestra de limón Meyer fue de 88'747,638, sin embargo, después de realizar la limpieza de las secuencias, eliminar aquellas de mala calidad, eliminar *reads* duplicados y filtrar el genoma de la planta se obtuvo un total de 38'964,192 *reads* de buena calidad, lo que concuerda con la cantidad estimada que permitieron detectar *Citrus tristeza virus* perteneciente al género *Closterovirus*.

Los secuenciadores Illumina normalmente producen lecturas con mala calidad al final de las lecturas, actualmente existen herramientas disponibles para recortar datos de lectura de extremo emparejado (Blawid, Silva, & Nagata, 2017). Blawid y colaboradores (2017) en su estudio realizado para descubrir y secuenciar nuevos genomas virales de plantas mediante secuenciación masiva paralela, usaron la



herramienta Trimmomatic para recortar lecturas de extremo emparejado, eliminar adaptadores y bases de N de baja calidad, para lo cual emplearon los siguientes parámetros: ILLUMINACLIP: TruSeq3 PE.fa: 2: 30: 10, LEADING: 3, TRAILING: 3, SLIDINGWINDOW: 4:20 y MINLEN: 36 y obtuvieron 10´323,599 *reads*. Trimmomatic presenta un alto valor predictivo positivo (*positive predictive value*, PPV) por lo que también se empleó esta herramienta para procesar los *reads* de secuenciación obtenidos en muestras de naranjilla y limón Meyer, los parámetros empleados fueron ILLUMINACLIP:TruSeq3 PE-2.fa:2:30:10, SLIDINGWINDOW:5:20 LEADING:5 TRAILING:5 y MINLEN:50 y se obtuvieron 69´555,658 *reads* (limón Meyer) y 76´156,490 *reads* (naranjilla). Al subir o bajar los valores de umbral de calidad de los parámetros (LEADING, TRAILING, SLIDINGWINDOW y MINLEN) se obtiene un menor o un mayor número de *reads* dependiendo de los valores que se asignen, mientras menos estrictos sean estos valores existe menos probabilidad de recortar erróneamente *reads* de buena calidad (Bolger, Lohse, & Usadel, 2014).

En un estudio realizado por Barrero y sus colaboradores (2017) se compararon cuatro ensambladores (SPAdes, Velvet, ABySS y SOAPdenovo) usados frecuentemente en la detección de virus, de los cuales SPAdes fue el que produjo mejores resultados con un porcentaje alto de genomas ensamblados, secuencias ensambladas más largas y un mayor número de bases ensambladas. Blawid y colaboradores (2017) utilizaron *k-mers* entre 21 y 99 para evaluar varios ensambladores de genomas virales entre ellos ABySS, SPAdes y Velvet y obtuvieron que Velvet generó *contigs* más diversos, mientras que SPAdes generó el *contig* más

largo. En esta investigación las muestras de limón Meyer y naranjilla fueron ensambladas con SPAdes, Velvet y ABySS con K21, K33, K55 Y K77, al compararlos se obtuvo que SPAdes generó el contig más largo y el mayor número total de bases obtenidos en los ensamblajes para las dos muestras, lo cual confirma lo obtenido en los estudios mencionados anteriormente. Spades al resultar el mejor ensamblador, fue empleado en la identificación de virus fitopatógenos de naranjilla y limón Meyer, debido a que va construyendo paulatinamente *contigs* (lecturas ensambladas) usando en la primera ronda el k-mer más pequeño y continua el ensamblaje de secuencias basándose en los resultados de la ronda anterior con k-mers de tamaños más grandes, además tiene una buena capacidad para ensamblar correctamente la longitud completa de genomas virales (Barrero et al., 2017; Berthet et al., 2016). Ensambladores como Velvet, ABySS y SOAPdenovo no tienen ese sistema de ensamblaje, usan k-mers individuales que generan *contigs* redundantes (Barrero et al., 2017).

El porcentaje de cobertura genómica obtenido a través del mapeo de lecturas ensambladas con genomas de referencia sirve para detectar virus, mientras mayor sea la cobertura más confiables serán los resultados de la identificación (Visser, Bester, et al., 2016). El rendimiento de las secuencias virales depende tanto de la organización del genoma viral como de la cantidad de lecturas virales en los datos.

Los métodos serológicos y moleculares para la detección de enfermedades

virales en los cultivos son muy sensibles, pero tienen la capacidad de detectar únicamente virus fitopatógenos conocidos. Cada prueba es específica para cada familia, género o especie. Estos métodos no pueden identificar patógenos virales desconocidos, mal caracterizados o altamente variables (Adams et al., 2011; Jones, Baizan-Edge, MacFarlane, & Torrance, 2017). A través del uso de primers genéricos y PCR convencional, se pudo detectar e identificar en muestras de naranjilla únicamente la presencia de *Potato yellowing virus* (PYV) del género *Ilarvirus* con un porcentaje de identidad del 97.63% con respecto a la secuencia (MH937419.1) reportada en el GenBank. A diferencia de las técnicas tradicionales, la secuenciación de próxima generación (NGS) no necesita un conocimiento previo del patógeno y puede proporcionar un resultado específico de la especie a la cual pertenece (Adams et al., 2009; Visser, Bester, et al., 2016). Con la tecnología NGS se pudo detectar e identificar la presencia de *Tobacco virus 2*, *Potato leafroll virus*, *Potato yellowing virus*, *Lily symptomless virus* y *Tomato torrado virus* en naranjilla y *Citrus tristeza virus*, *Citrus vein enation*, *Potato yellowing virus*, *Potato leafroll virus* y *Lily symptomless virus* en muestras de limón.

En Ecuador se han realizado estudios de identificación y detección de virus en varios cultivos, sin embargo, los reportes de virus fitopatógenos en limón Meyer y naranjilla son escasos. *Citrus tristeza virus* (CTV) es el patógeno con mayor prevalencia identificado y reportado en cítricos en varios países (Jooste, 2017). En 2017 Quito-Avila y sus colaboradores caracterizaron en cultivos ecuatorianos de naranjilla un nuevo *Tymovirus* causante de atrofia y mosaico clorótico denominado

tentativamente *Naranjilla chlorotic mosaic virus* (NarCMV). En Colombia se han identificado para este cultivo virus como *Cucumber mosaic virus* (CMV), *Potato yellow vein virus* (PYVV) y *Alstroemeria necrotic streak virus* (ANSV) (Gallo et al., 2018). Los virus que pudieron ser detectados e identificados en naranjilla a través de secuenciación masiva paralela son diferentes a los reportados y pertenecen a diferentes familias, géneros y especies. Silvestre, Untiveros & Cuellar (2011) detectaron en cultivos de papa en Ecuador *Potato yellowing virus* (PYV) y lo identificaron a nivel de género, para lo cual realizaron un análisis filogenético que agrupó las secuencias de PYV con las del género *Illavirus* y señaló que su pariente más cercano fue *Fragaria chiloensis latent virus* (FCILV) con el que comparte 77% de identidad a nivel de nucleótidos y 85% de identidad a nivel de aminoácidos. La información del genoma completo de PYV no se encuentra disponible en el GenBank, únicamente se encuentra disponible las secuencias separadas de cada segmento de ARN de este virus. En las muestras analizadas de naranjilla y limón Meyer se identificó *Fragaria chiloensis latent virus* con un 8% del total de *contigs* ensamblados como patógeno de estos cultivos; a pesar de ello, al importar los *contigs* generados a Geneious v11.1.5 y realizar una búsqueda Blastn se obtuvo que los *contigs* correspondientes a *Fragaria chiloensis latent virus* presentaron un porcentaje de identidad de 97.14% con la secuencia del segmento 3 de RNA de *Potato yellowing virus* reportado en el Genbank, lo cual nos indicaría la presencia de un falso positivo en la identificación de virus fitopatógenos obtenida.

La presencia de PYV en hojas sintomáticas de cultivos de pimiento en la provincia de Pichincha fue confirmada usando cebadores universales para la familia *Bromoviridae*, con los que se obtuvo un amplicón de 300 pb que se secuenció y reveló una identidad de 93.1-97% a nivel de nucleótidos y 78-90% a nivel de aminoácidos con otros aislados de PYV (Sivaprasad, Garrido, Mendez, Garrido, & Ramos, 2016). Para comprobar la presencia de PYV obtenido en este estudio mediante el análisis Blastn de los cotings ensamblados, se realizó una confirmación a través de primers genéricos para *Illavirus* y PCR convencional en donde se obtuvieron amplicones del mismo tamaño obtenido en el estudio mencionado anteriormente (300 pb) en muestras de naranjilla, los mismos que fueron secuenciados mediante método Sanger, dando como resultado el virus PYV (*Potato yellowing virus*) con un porcentaje de identidad superior a 97.63%; esto confirmaría la presencia de este virus en cultivos de naranjilla. A pesar de no haberse obtenidos amplicones de los tamaños esperados para limón Meyer posiblemente debido a una baja calidad del ADNc de las muestras obtenidas a partir de ARN total (Khalifa, Varsani, Ganley, & Pearson, 2016), el análisis de datos de secuenciación masiva paralela mostraría la presencia de PYV en este cultivo.

Los cultivos son susceptibles a la infección por virus y causan pérdidas significativas de rendimiento y calidad. La rápida detección e identificación de virus fitopatógenos es importante para la producción de cultivos libres de patógenos

virales (Jones et al., 2017). Este estudio permitió tener un primer reporte de virus fitopatógenos que afectan a cultivos de naranjilla y limón Meyer usando secuenciación de próxima generación (NGS).

## CAPÍTULO VI

### CONCLUSIONES

A través de retrotranscripción, PCR convencional con primers genéricos para *Illavirus* y secuenciación Sanger de amplicones se detectó en muestras de naranjilla *Potato yellowing virus* con un porcentaje de identidad del 97.63% y 97.65% (MH937419.1) a nivel de nucleótidos.

Mediante secuenciación masiva paralela se identificaron las especies *Tobacco virus 2*, *Potato leafroll virus*, *Lily symptomless virus* y *Tomato torrado virus* con porcentajes de identidad de 95.55% (KY038943.1), 96.72% (JQ420902.1), 98.80% (MK649770.1) y 96.17% (MH587229.1) respectivamente en plantas de naranjilla y *Citrus tristeza virus*, *Citrus vein enation*, *Potato leafroll virus* y *Lily symptomless virus* con porcentajes de identidad de 99.96% (KY110738.1), 98.48% (LC433634.1), 93.92% (KX073467.2) y 99.57% (MK649770.1) en plantas de limón Meyer. Se confirmó la presencia de *Potato yellowing virus* en ambos cultivos con un porcentaje de identidad de 97.14% (MH937420.1) en limón Meyer y 98.80% (MK649770.1) en naranjilla.

La secuenciación de próxima generación (NGS) en combinación con análisis bioinformáticos han demostrado ser una herramienta eficiente y de alto rendimiento

para la detección de virus, a partir de una cantidad óptima de datos que representen una parte significativa del genoma del virus fitopatógenos.

El análisis comparativo realizado entre los ensambladores (SPAdes, Velvet y ABySS) mostró que SPAdes con kmer 77 produjo los mejores resultados de ensamblaje al generar el *contig* más largo en naranjilla (19,984 bp) y limón Meyer (16,747 bp) y además presentó el mayor número total de bases con 14´637,417 bp (limón Meyer) y 65´673,696 bp (naranjilla).

Se obtuvo el genoma completo de *Citrus vein enation* (6,103 bp) y el genoma casi completo de *Citrus tristeza virus* (19,195 bp) en limón Meyer, mientras que en naranjilla se obtuvo la secuencia casi completa del segmento 1 de RNA de *Tomato torrado virus* (7,287 bp) mediante el ensamblaje realizado con SPAdes.

El pipeline bioinformático establecido permitió extraer una cantidad óptima de secuencias a partir de datos de secuenciación masiva paralela, con los cuales se obtuvieron 25 y 38 *contigs* correspondientes a virus fitopatógenos para limón Meyer y naranjilla respectivamente a partir de 32,024 *contigs* (limón Meyer) y 134,506 *contigs* (naranjilla) ensamblados mediante SPAdes, con los que se pudieron identificar varias especies de virus presentes en estos cultivos.

VirFind identificó secuencias virales en las dos muestras, los virus fitopatógenos identificados mediante esta herramienta fueron los mismos



encontrados a través del pipeline establecido. En limón Meyer se identificó *Citrus tristeza virus* y *Citrus vein enation* con porcentajes de identidad de 99.65% (KU358530.1) y 98.46% (LC433634.1) respectivamente, mientras que en naranjilla se identificó *Potato leafroll virus*, *Potato yellowing virus*, *Lily symptomless virus*, *Tomato torrado virus* con porcentajes de identidad de 96.01% (JQ420902.1), 98.73% (MH937420.1), 96.80% (MK649770.1), 95.96% (MH587229.1) respectivamente.

## **CAPÍTULO VII**

### **RECOMENDACIONES**

Efectuar ensamblajes *reference-guided* a partir de los *contings* obtenidos para obtener secuencias genómicas completas de los virus fitopatógenos identificados en naranjilla y limón Meyer.

Emplear protocolos de extracción de ARN de doble cadena (dsRNA) que permitan obtener una muestra más estable, de mejor calidad y con menos degradación.

Realizar un estudio más amplio, que incluya a las provincias del país con mayor índice de producción tanto de limón Meyer como de naranjilla para poder determinar y corroborar la incidencia de los virus fitopatógenos identificados en esta investigación.

Verificar la presencia de los virus identificados en las dos muestras usando primers específicos y pruebas serológicas ELISA.

Ejecutar pruebas de patogenicidad/postulados de Koch que permitan verificar los agentes causales de los síntomas descritos en las muestras.

Realizar la detección de virus fitopatógenos de otras especies de cítricos,

variedades de naranjilla y cultivos de importancia económica en el país mediante secuenciación masiva paralela y análisis bioinformático para poder implementar un registro de virus en plantas en el Ecuador.

## CAPÍTULO VIII

### BIBLIOGRAFÍA

- Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., ... Boonham, N. (2009). *Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology*. *Molecular Plant Pathology*, 10(4), 537-545. <https://doi.org/10.1111/j.1364-3703.2009.00545.x>
- Adams, I. P., Glover, R. H., Monger, W. A., Thwaites, R., Mumford, R., Jackeviciene, E., ... Boonham, N. (2011). *Next-Generation Sequencing and Metagenomic Analysis: A Universal Diagnostic Tool in Plant Pathology*. En *Handbook of Molecular Microbial Ecology II* (pp. 63-72). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118010549.ch7>
- Alavandi, S., & Poornima, M. (2012). *Viral metagenomics: a tool for virus discovery and diversity in aquaculture*. *Indian journal of virology: an official organ of Indian Virological Society*, 23(2), 88-98. <https://doi.org/10.1007/s13337-012-0075-2>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). *Basic local alignment search tool*. *Journal of Molecular Biology*, 215(3), 403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Alvarez-Quinto, R. A., Cornejo-Franco, J. F., & Quito-Avila, D. F. (2017). *Characterization of a not so new potexvirus from babaco (Vasconcellea x heilbornii)*. *PLOS ONE*, 12(12). <https://doi.org/10.1371/journal.pone.0189519>

- Barba, M., Czosnek, H., & Hadidi, A. (2013). *Historical perspective, development and applications of next-generation sequencing in plant virology*. *Viruses*, 6(1), 106-136. <https://doi.org/10.3390/v6010106>
- Barrero, R. A., Napier, K. R., Cunnington, J., Liefing, L., Keenan, S., Frampton, R. A., ... Bellgard, M. I. (2017). *An internet-based bioinformatics toolkit for plant biosecurity diagnosis and surveillance of viruses and viroids*. *BMC Bioinformatics*, 18(1), 26. <https://doi.org/10.1186/s12859-016-1428-4>
- Bawden, F. C., & Kassanis, B. (1950). *Some effects of host-plant nutrition on the multiplication of viruses*. *Annals of Applied Biology*, 37(2), 215-228. <https://doi.org/10.1111/j.1744-7348.1950.tb01040.x>
- Berthet, N., Descorps-Declère, S., Nkili-Meyong, A. A., Nakouné, E., Gessain, A., Manuguerra, J.-C., & Kazanji, M. (2016). *Improved assembly procedure of viral RNA genomes amplified with Phi29 polymerase from new generation sequencing data*. *Biological Research*, 49(1), 39. <https://doi.org/10.1186/s40659-016-0099-y>
- Blawid, R., Silva, J. M. F., & Nagata, T. (2017). *Discovering and sequencing new plant viral genomes by next-generation sequencing: description of a practical pipeline*. *Annals of Applied Biology*, 170(3), 301-314. <https://doi.org/10.1111/aab.12345>
- Bolaños, V. (2007). *Determinación de las características físicas y químicas del Limón Meyer (Citrus limon)*. Universidad Técnica del Norte.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). *Trimmomatic: a flexible trimmer for Illumina sequence data*. *Bioinformatics*, 30(15), 2114-2120.

<https://doi.org/10.1093/bioinformatics/btu170>

Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J., & Mumford, R. (2014). *Methods in virus diagnostics: From ELISA to next generation sequencing*. *Virus Research*, 186, 20-31.

Bucheli, J. (2013). *Selección de genotipos promisorios de naranjilla (Solanum quitoense Lam.) provenientes de cruzamientos interespecíficos adaptados a condiciones subtropicales*. Quito, Ecuador.

Carter, M., Shieh, J. C., Farra, N., & Harris, G. (2015). *Guide to research techniques in neuroscience* (Segunda). Elsevier Inc. .

Castro, L., Timmer, L., Leguizamón, J., Müller, G., & Corrales. José Antonio. (2000). *Enfermedades de los cítricos en Colombia*. Bogotá, Colombia.

Cooper, J. (2007). *Virus and amp; Related Plant Pathogens*.

Dávila, M. F. (2016). *Diversidad de hongos asociados a lesiones foliares en naranjilla (Solanum quitoense Lam.) en varias localidades de la Amazonía Ecuatoriana, 2015*. Pontificia Universidad Católica Del Ecuador.

Dombrovsky, A., Glanz, E., Pearlsman, M., Lachman, O., & Antignus, Y. (2010). *Characterization of Pepper yellow leaf curl virus, a tentative new Polerovirus species causing a yellowing disease of pepper*. *Phytoparasitica*, 38(5), 477-486.  
<https://doi.org/10.1007/s12600-010-0120-x>

Dutilh, B. E., Reyes, A., Hall, R. J., & Whiteson, K. L. (2017). *Editorial: Virus Discovery by*

*Metagenomics: The (Im)possibilities. Frontiers in microbiology*, 8, 1710.  
<https://doi.org/10.3389/fmicb.2017.01710>

Fiallos, J. (2000). *Naranjilla INIAP-Palora: Híbrido interespecífico de alto rendimiento*. Palora, Ecuador.

Gallo, Y., Toro, L. F., Jaramillo, H., Gutiérrez, P. A., & Marín, M. (2018). *Identificación y caracterización molecular del genoma completo de tres virus en cultivos de lulo (Solanum quitoense) de Antioquia (Colombia)*. *Revista Colombiana de Ciencias Hortícolas*, 12(2), 281-292. <https://doi.org/10.17584/rcch.2018v12i2.7692>

Gergerich, R. C., & Dolja, V. V. (2006). *Introduction to Plant Viruses, the Invisible Foe. The Plant Health Instructor*. <https://doi.org/10.1094/PHI-I-2006-0414-01>

Green, K. J., Mollov, D., Tran, L. T., Alvarez-Quinto, R. A., Ochoa, J. B., Quito-Avila, D. F., & Karasev, A. V. (2018). *Characterization of a New Tymovirus Causing Stunting and Chlorotic Mosaic in Naranjilla (Solanum quitoense)*. *Plant Disease*, 102(5), 911-918. <https://doi.org/10.1094/PDIS-10-17-1534-RE>

Green, K. J., Mollov, D., Tran, L. T., Alvarez-Quinto, R. A., Ochoa, J. B., Quito-Avila, D. F., & Karasev, A. V. (2017). *Characterization of a New Tymovirus Causing Stunting and Chlorotic Mosaic in Naranjilla (Solanum quitoense)*. *Plant Disease*, 102(5), 911-918. <https://doi.org/10.1094/PDIS-10-17-1534-RE>

Guayasamín, M. (2015). *Evaluación ex ante del impacto socio-económico del manejo convencional y mejorado del cultivo de Naranjilla (Solanum quitoense) en el Ecuador*. Universidad Central del Ecuador.

Guevara, F. (2017). *Detección e identificación de virus de los géneros Potexvirus, Tombusvirus y Tobamovirus en aguas de riego de cultivos de Babaco (Vasconcellea x heilbornii. var.pentagona), en la provincia de Pichincha*. Universidad de las Fuerzas Armadas - ESPE.

Ho, T., & Tzanetakis, I. E. (2014a). *Development of a virus detection and discovery pipeline using next generation sequencing*. *Virology*, 471-473, 54-60. <https://doi.org/10.1016/J.VIROL.2014.09.019>

Ho, T., & Tzanetakis, I. E. (2014b). *Development of a virus detection and discovery pipeline using next generation sequencing*. *Virology*, 471-473, 54-60. <https://doi.org/10.1016/j.virol.2014.09.019>

Hull, R. (2009a). *Comparative plant virology* (2nd ed.). Elsevier/Academic Press.

Hull, R. (2009b). *Propagation of Major Plant-Virus Hosts*. En *Current Protocols in Microbiology* (pp. 16A.2.1-16A.2.4). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9780471729259.mc16a02s14>

Illumina Inc. (2017). *An introduction to Next-Generation Sequencing Technology*. U.S.

INIAP. (2018). *Naranjilla*. Recuperado de

<http://tecnologia.iniap.gob.ec/index.php/explore-2/mfruti/rnaranjilla>

Ismail, M. H. (1997). The use of enzyme — linked immunosorbent assay (ELISA) for quantitative detection of potato virus Y in potato and other test plants. *Microbiological Research*, 152(3), 307-313. [https://doi.org/10.1016/S0944-5013\(97\)80045-6](https://doi.org/10.1016/S0944-5013(97)80045-6)



Jeong, J.-J., Ju, H.-J., & Noh, J. (2014). *A Review of Detection Methods for the Plant Viruses*. *Research in Plant Disease*, 20(3), 173-181.

<https://doi.org/10.5423/RPD.2014.20.3.173>

Jones, S., Baizan-Edge, A., MacFarlane, S., & Torrance, L. (2017). *Viral Diagnostics in Plants Using Next Generation Sequencing: Computational Analysis in Practice*. *Frontiers in Plant Science*, 8, 1770. <https://doi.org/10.3389/fpls.2017.01770>

Jooste, T. (2017). *The detection of Citrus tristeza virus genetic variants using pathogen specific electronic probes*. Stellenbosch University.

Kaur, S., & Kaur, S. (2015). *Genomics with Cloud Computing*. *International Journal of Scientific & Technology Research*, 4(04).

Khalifa, M. E., Varsani, A., Ganley, A. R. D., & Pearson, M. N. (2016). *Comparison of Illumina de novo assembled and Sanger sequenced viral genomes: A case study for RNA viruses recovered from the plant pathogenic fungus Sclerotinia sclerotiorum*. *Virus Research*, 219, 51-57. <https://doi.org/10.1016/j.virusres.2015.11.001>

Koboldt, D., Larson, D., Chen, K., Ding, L., & Wilson, R. (2012). *Massively parallel sequencing approaches for characterization of structural variation*. *Methods in molecular biology (Clifton, N.J.)*, 838, 369-384. [https://doi.org/10.1007/978-1-61779-507-7\\_18](https://doi.org/10.1007/978-1-61779-507-7_18)

Kukurba, K. R., & Montgomery, S. B. (2015). *RNA Sequencing and Analysis*. *Cold Spring Harbor protocols*, 2015(11), 951-969. <https://doi.org/10.1101/pdb.top084970>

León, C. (2009). *Perfil de limones y limas*. Ecuador.

Loconsole, G., Saldarelli, P., Doddapaneni, H., Savino, V., Martelli, G. P., & Saponari, M. (2012). *Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae*. *Virology*, 432(1), 162-172. <https://doi.org/10.1016/J.VIROL.2012.06.005>

Luis, M., Llauger, R., & Collazo, C. (s. f.). *Principales enfermedades bacterianas de los cítricos*. La Habana, Cuba.

Macrogen. (2018a). *Raw Data Report*.

Macrogen. (2018b). *Secuenciación de próxima generación*. Recuperado de <https://dna.macrogen.com/eng/support/ngs/intro/wgs.jsp>

Maliogka, V. I., Minafra, A., Saldarelli, P., Ruiz-García, A. B., Glasa, M., Katis, N., & Olmos, A. (2018). Recent advances on detection and characterization of fruit tree viruses using high-throughput sequencing technologies. *Viruses*, 10(8), 1-23. <https://doi.org/10.3390/v10080436>

Marschner, H. (2006). *Mineral nutrition of higher plants*. (Elsevier, Ed.) (2.<sup>a</sup> ed.). Alemania.

Medina, E. (2007). *Determinación de las características físicas y químicas de Limón Meyer (Citrus limon)*. Universidad Técnica Del Norte.

Mesa, H. J. (2017). *Análisis del transcriptoma y viroma de Passiflora edulis f. edulis en cultivos de Antioquía utilizando métodos de Secuenciación de Nueva Generación*. Universidad Nacional de Colombia.

- Mutuku, J. M., Wamonje, F. O., Mukeshimana, G., Njuguna, J., Wamalwa, M., Choi, S.-K., ... Harvey, J. J. W. (2018). *Metagenomic Analysis of Plant Virus Occurrence in Common Bean (Phaseolus vulgaris) in Central Kenya. Frontiers in Microbiology, 9*, 2939. <https://doi.org/10.3389/fmicb.2018.02939>
- Naidu, R. A., & Hughes, A. (2018). *Methods for the detection of plant virus diseases*.
- Pantaleo, V., & Chiumenti, M. (2018). *Viral metagenomics: methods and protocols*. (Institute Sustainable Plant Protection of the CNR, Ed.). Bari, Italia: Springer Nature.
- Pecman, A., Kutnjak, D., Gutiérrez-Aguirre, I., Adams, I., Fox, A., Boonham, N., & Ravnkar, M. (2017). *Next Generation Sequencing for Detection and Discovery of Plant Viruses and Viroids: Comparison of Two Approaches. Frontiers in microbiology, 8*, 1998. <https://doi.org/10.3389/fmicb.2017.01998>
- Peña, I., Pérez, J., López, D., & Batista, L. (s. f.). *Principales enfermedades virales y afines de los cítricos*. La Habana, Cuba.
- Pennazio, S., & Roggero, P. (1997). *Mineral nutrition and systemic virus infections in plants. Phytopathologia Mediterranea, 36(1)*, 54-66.
- Prabha, K., Baranwal, V., & Jain, R. (2013). *Applications of next generation high throughput sequencing technologies in characterization, discovery and molecular interaction of plant viruses. Indian journal of virology: an official organ of Indian Virological Society, 24(2)*, 157-165. <https://doi.org/10.1007/s13337-013-0133-4>
- Quito-Avila, D. F., Jelkmann, W., Tzanetakis, I. E., Keller, K., & Martin, R. R. (2011). *Complete sequence and genetic characterization of Raspberry latent virus, a novel*

member of the family Reoviridae. *Virus Research*, 155(2), 397-405.  
<https://doi.org/10.1016/j.virusres.2010.11.008>

Revelo, J., Viteri, P., Vásquez, W., Valverde, F., León, J., & Gallegos, P. (2010). *Manual del Cultivo Ecológico de la Naranja*. Quito, Ecuador.

Roossinck, M. J., Martin, D. P., & Roumagnac, P. (2015). Plant Virus Metagenomics: *Advances in Virus Discovery*. <https://doi.org/10.1094/PHYTO-12-14-0356-RVW>

Rosales, M., Sepúlveda, P., Rojas, C., Medina, C., Sepúlveda, G., Brown, J., & Mora, R. (2011). *Virus transmitidos por insectos vectores en tomate en la Región de Arica y Parinacota: situación actual y manejo*. Chile.

Rowhani, A., & Falk, B. W. (1995). *Enzyme-Linked Immunosorbent Assay (ELISA) Methods to Certify Pathogen (Virus)-Free Plants*. En *Plant Cell, Tissue and Organ Culture* (pp. 267-280). Berlin, Heidelberg: Springer Berlin Heidelberg.  
[https://doi.org/10.1007/978-3-642-79048-5\\_21](https://doi.org/10.1007/978-3-642-79048-5_21)

Shih, S., & Doran, P. (2009). *In vitro propagation of plant virus using different forms of plant tissue culture and modes of culture operation*. *Journal of Biotechnology*, 143(3), 198-206. <https://doi.org/10.1016/j.jbiotec.2009.07.007>

Silva, W., Gómez, P., Viera, W., Viteri, P., & Ron, L. (2016). *Selección de líneas promisorias de Naranja para mejorar la calidad de la fruta*. *Revista Científica Ecuatoriana* (Vol. 3).

Sivaprasad, Y., Garrido, P., Mendez, K., Garrido, A., & Ramos, L. (2016). *First Report of Potato Yellowing Virus Infecting Pepper in Ecuador*. *Journal of Plant Pathology*,

97(4sup), 75. <https://doi.org/10.4454/JPP.V97I4SUP.023>

Spann, T., & Schumann, A. (2009). *The Role of Plant Nutrients in Disease Development with Emphasis on Citrus and Huanglongbing*. Florida.

Stevens, W. A. (1983). *Transmission of Plant Viruses*. En *Virology of Flowering Plants* (pp. 41-68). Boston, MA: Springer US. [https://doi.org/10.1007/978-1-4757-1251-3\\_3](https://doi.org/10.1007/978-1-4757-1251-3_3)

Stobbe, A. H., & Roossinck, M. J. (2014). *Plant virus metagenomics: what we know and why we need to know more*. *Frontiers in plant science*, 5, 150. <https://doi.org/10.3389/fpls.2014.00150>

Torres, A., Vargas, J., García, Y., Arteaga, Y., & Navarrete, H. (2017). *Solanum quitoense Lam, un rubro económico de alto valor en la Amazonía Ecuatoriana*. En *Gente, Bosque y Biodiversidad: El rol del bosque sobre la biodiversidad y las poblaciones* (p. 279). Puyo, Ecuador.

Untiveros, M., Perez-Egusquiza, Z., & Clover, G. (2010). *PCR assays for the detection of members of the genus Ilarvirus and family Bromoviridae*. *Journal of Virological Methods*, 165(1), 97-104. <https://doi.org/10.1016/J.JVIROMET.2010.01.011>

Valarezo Concha, A., Valarezo Cely, O., Mendoza, A., & Alvarez, H. (2014). *Guía técnica sobre el manejo de los cítricos en el Litoral ecuatoriano* (Vol. 101). Portoviejo, EC: INIAP, Estación Experimental Portoviejo, Programa Nacional de Fruticultura.

Valverde, R. A., & De La Torre-Almaraz, R. (2017). *Extraction and purification of large dsRNAs from virus-infected plants and fungi; applications in virus detection and identification*. *Revista mexicana de fitopatología*. scielomx .

Velasco, V. (2000). *Role of Mineral Nutrition on Plant Disease Tolerance*. *TERRA*, 17(3).

Visser, M., Bester, R., Burger, J. T., & Maree, H. J. (2016). *Next-generation sequencing for virus detection: Covering all the bases*. *Virology Journal*, 13(1).  
<https://doi.org/10.1186/s12985-016-0539-x>

Visser, M., Burger, J. T., & Maree, H. J. (2016). *Targeted virus detection in next-generation sequencing data using an automated e-probe based approach*. *Virology*, 495, 122-128. <https://doi.org/10.1016/J.VIROL.2016.05.008>

Viteri Diaz, P., Vásquez, W., León, J., Viera, W., Posso, M., Hinojosa, M., ... Ochoa, J. (2009). *Naranjilla de jugo (Solanum quitoense Latti.) injerta en patrones de solanáceas silvestres resistentes a Fusarium oxysporum y a Meloidogyne incógnita*. Quito, Ecuador.

Wamaita, M. J., Nigam, D., Maina, S., Stomeo, F., Wangai, A., Njuguna, J. N., ... Garcia-Ruiz, H. (2018). *Metagenomic analysis of viruses associated with maize lethal necrosis in Kenya*. *Virology Journal*, 15(1), 90. <https://doi.org/10.1186/s12985-018-0999-2>

Whitfield, A. E., & Rotenberg, D. (2015). *Disruption of insect transmission of plant viruses*. *Current Opinion in Insect Science*, 8, 79-87.  
<https://doi.org/10.1016/j.cois.2015.01.009>

Williams, S., Boehm, J., & Qu, F. (2017). *Viral Diseases of Plants*. Ohio.

Zhao, G., Krishnamurthy, S., Cai, Z., Popov, V. L., Travassos da Rosa, A. P., Guzman,

H., ... Wang, D. (2013). *Identification of Novel Viruses Using VirusHunter -- an Automated Data Analysis Pipeline*. *PLOS ONE*, 8(10), e78470.  
<https://doi.org/10.1371/journal.pone.0078470>

Zhao, G., Wu, G., Lim, E. S., Droit, L., Krishnamurthy, S., Barouch, D. H., ... Wang, D. (2017). *VirusSeeker, a computational pipeline for virus discovery and virome composition analysis*. *Virology*, 503, 21-30.

<https://doi.org/10.1016/j.virol.2017.01.005>

Zoni, A., & Martínez, J. (2015). *Efficacy of the next generation sequencing in the diagnosis and follow up oncohematologic diseases*. Madrid.