Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ciencias Biológicas
Unidad de Posgrado

Identificación de nuevos virus en variedades de *Solanum tuberosum* L. (Solanaceae) mediante secuenciación de alto rendimiento de ARN de interferencia y obtención de genomas virales en muestras provenientes de Cusco, Perú

TESIS

Para optar el Grado Académico de Magister en Biología Molecular

AUTOR

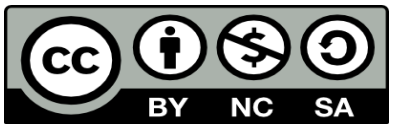
Melody AMAO PEÑA

ASESOR

Dr. Pablo Sergio RAMÍREZ ROCA

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Amao, M. (2026). *Identificación de nuevos virus en variedades de Solanum tuberosum L. (Solanaceae) mediante secuenciación de alto rendimiento de ARN de interferencia y obtención de genomas virales en muestras provenientes de Cusco, Perú*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Melody Amao Peña
Tipo de documento de identidad	DNI
Número de documento de identidad	45418516
URL de ORCID	**https://orcid.org/0000-0002-9756-6347**
Datos de asesor	
Nombres y apellidos	Pablo Sergio Ramírez Roca
Tipo de documento de identidad	DNI
Número de documento de identidad	06183797
URL de ORCID	https://orcid.org/0000-0001-9309-7021
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Talledo Rivera, Miguel Ángel Francisco
Tipo de documento	DNI
Número de documento de identidad	25625144
Miembro del jurado 1	
Nombres y apellidos	Mamani Zapana, Priscila Rosse
Tipo de documento	DNI
Número de documento de identidad	43058540
Miembro del jurado 2	
Nombres y apellidos	Vera Obando, Nora Yessenia
Tipo de documento	DNI
Número de documento de identidad	41851323
Miembro del jurado 3	
Nombres y apellidos	
Tipo de documento	

Número de documento de identidad	
Datos de investigación	
Línea de investigación	“_”
Grupo de investigación	“_”
Agencia de financiamiento	PNIA
Ubicación geográfica de la investigación	Perú-Cusco
Año o rango de años en que se realizó la investigación	2016-2025
URL de disciplinas OCDE	https://purl.org/pe-repo/ocde/ford#1.06.00

Expediente N.º 281-UPG-FCB-2025

ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR AL GRADO ACADÉMICO DE MAGÍSTER EN BIOLOGÍA MOLECULAR

Siendo las 14:00 horas del jueves 09 de abril de 2026, en el Auditorio de la Facultad de Ciencias Biológicas, el Jurado de Tesis conformado por:

Mg. Miguel Ángel Francisco Talledo Rivera	(Presidente)
Mg. Priscila Rosse Mamani Zapana	(Miembro)
Mg. Nora Yessenia Vera Obando	(Miembro)
Dr. Pablo Sergio Ramírez Roca	(Asesor)

Se reunió para la sustentación oral y pública de la Tesis para optar al Grado Académico de Magíster en Biología Molecular, que solicitara la señora Bachiller **MELODY AMAO PEÑA.**

Después de darse lectura al **Expediente N.º 281-UPG-FCB-2025,** en el que consta haberse cumplido con todas las disposiciones reglamentarias, los señores miembros del Jurado presenciaron la exposición de la Tesis Titulada:

"Identificación de nuevos virus en variedades de *Solanum tuberosum* L. (Solanaceae) mediante secuenciación de alto rendimiento de ARN de interferencia y obtención de genomas virales en muestras provenientes de Cusco, Perú", y formuladas las preguntas, éstas fueron absueltas por la graduando.

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América

FACULTAD DE CIENCIAS BIOLÓGICAS

UNIDAD DE POSGRADO

Acto seguido el Jurado procedió a la deliberación la que dio como resultado el calificativo de: Excelente con la nota de Diecinueve (19)

A continuación, el Presidente del Jurado de Tesis recomienda que la Facultad proponga que la Universidad le otorgue el grado académico de Magíster en Biología Molecular, a la señorita **MELODY AMAO PEÑA.**

Siendo las 15:30 se levantó la sesión, recibiendo la graduando las felicitaciones de los señores miembros del Jurado y público asistente.

Se extiende la presente Acta en Lima, el jueves 09 de abril del año 2026.

Mg. Miguel Ángel F. Talledo Rivera
Profesor Asociado a D.E.
PRESIDENTE
mtalledor@unmsm.edu.pe

Dr. Pablo S. Ramírez Roca
Profesor Principal a D.E.
ASESOR
pramirezr@unmsm.edu.pe
RENACYT P0001952

Mg. Priscila R. Mamani Zapana
Profesora Auxiliar a T. C.
MIEMBRO
pmamaniz@unmsm.edu.pe

Mg. Nora Y. Vera Obando
Profesora Asociado a D.E.
MIEMBRO
nveraob@unmsm.edu.pe

UNIVERSIDAD NACIONAL MAYOR DE
SAN MARCOS
Universidad del Perú, Decana de América

Anexo 1

CERTIFICADO DE SIMILITUD

Yo, Pablo Sergio Ramírez Roca en mi condición de asesor acreditado mediante Dictamen de la UPG N.º 000384-2025-UPG-FCB/UNMSM de la tesis, cuyo título es "Identificación de nuevos virus en variedades de *Solanum tuberosum* L. (Solanaceae) mediante secuenciación de alto rendimiento de ARN de interferencia y obtención de genomas virales en muestras provenientes de Cusco, Perú", presentado por la bachiller Melody Amao Peña para optar al grado académico de Magíster en Biología Molecular,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el **6% de similitud**, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado de magister en biología molecular.

Lima, 07 de mayo de 2026

UNMSM
Firmado digitalmente por RAMIREZ ROCA Pablo Sergio FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 07.05.2026 18:28:53 -05:00

Firma del Asesor

Nombres y Apellidos del asesor: **PABLO SERGIO RAMÍREZ ROCA**

DNI: **06183797**

* Según corresponda: bachiller, magíster, doctor; título profesional o título de segunda especialidad

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

Agradecimientos

Agradezco a Dios por ser mi guía y fortaleza, por las bendiciones que me ha otorgado y por todo lo que he conseguido. Su presencia y apoyo continúan iluminando mi camino hacia todo lo que aún está por alcanzar.

A mis padres, con infinita gratitud, por su amor incondicional, su apoyo constante y por ser mi mayor inspiración. Este logro es tanto mío como suyo.

A mi hijo Liam, mi mayor fuente de motivación, por recordarme cada día la importancia de perseverar y llenarme de amor y propósito. Este logro es también para ti, que das sentido a todos mis esfuerzos.

Agradezco profundamente al Dr. Jan Kreuze, Mg. Segundo Fuentes, y a la Blga. Ana Pérez, por su confianza, guía constante, paciencia y dedicación a lo largo de este proyecto. Su experiencia y apoyo fueron fundamentales para alcanzar este objetivo.

Mi especial agradecimiento al Dr. Pablo Ramírez por las valiosas directrices brindadas durante el desarrollo de esta tesis y apoyo en este proceso.

Finalmente, extiendo mi agradecimiento al personal profesional y técnico del Laboratorio de Patología de Plantas del CIP, por su apoyo incondicional y por compartir sus conocimientos y experiencia en las diversas metodologías de trabajo. Su colaboración fue esencial para concretar este estudio.

Dedicatoria

Para mis padres,

Su amor y sacrificio me han guiado siempre. Cada paso que doy está inspirado en el ejemplo que ustedes me han dado. Este logro es una forma de honrar todo lo que han hecho por mí.

ÍNDICE

Índice de figuras……VII
Índice de tablas……X
Resumen……XII
Abstract……XIII
1.INTRODUCCIÓN……1
2. ANTECEDENTES……2
2.1. Importancia de la papa……2
2.2. Técnicas de diagnóstico……3
2.3. ARN de interferencia pequeños……3
2.4. HTS en la detección de virus……5
2.5. *Badnavirus*……6
2.6. *Torradovirus*……8
3. HIPÓTESIS……9
4. OBJETIVOS……9
4.1 Objetivo general……9
4.2 Objetivos específicos……9
5. MATERIALES Y MÉTODOS……10
5.1 Colecta de muestras……10
5.2. Obtención de los ARN pequeños……12
5.2.1. Extracción de ARN total……12
5.2.2. Enriquecimiento de los ARN pequeños……13
5.2.3. Aislamiento de ARN de interferencia (siARN)……13
5.3. Preparación de librerías de siARN……15
5.3.1. Adenilación del adaptador 3'……15
5.3.2. Ligación del adaptador 3' adenilado……16
5.3.3. Ligación del adaptador 5'……17
5.3.4. Transcripción reversa……17
5.3.5 PCR test……18
5.3.6 Enriquecimiento de PCR……19

5.3.7. Cuantificación de ADN purificado……..……20
5.4. Secuenciación de alto rendimiento…….……..…20
5.5. Análisis bioinformáticos mediante VirusDetect y VirFind……..20
5.6. Identificación de nuevos virus……24
5.7. Diseño de *primers*……24
5.7.1. *Primers* del género *Badnavirus*……...25
5.7.2. *Primers* del género *Torradovirus*……25
5.8. PCR……26
5.9. Clonación de productos de PCR y secuenciación…….……..27
5.10. Análisis bioinformáticos……28
5.11. Árbol filogenético……...28
6. RESULTADOS……29
6.1. Análisis de secuencias……29
6.2. Análisis comparativo entre VirusDetect y Virfind…….....32
6.3. Obtención de los seudogenoma de los virus nuevos…… ...34
6.3.1. Género *Badnavirus*……34
6.3.2. Género *Torradovirus*……..38
6.4. Resultados de PCR…….....……41
6.4.1 Género *Badnavirus*……..41
6.4.2 Género *Torradovirus*……44
6.5. Obtención del genoma completo……..47
6.6.1 Género *Badnavirus*……..……47
6.6.2. Género *Torradovirus*……..50
6.6. Árbol filogenético……54
6.6.1. Género *Badnavirus*……54
6.6.2. Género *Torradovirus*……..56
7. DISCUSIÓN……..58
8. CONCLUSIONES……62
9. RECOMENDACIONES……...63
10. REFERENCIAS BIBLIOGRÁFICAS……...64

11. ANEXOS…….69

ÍNDICE DE FIGURAS

Figura 1: Mecanismo de formación de los siARN……..4
Figura 2: Organización del genoma de *Badnavirus*……..…...7
Figura 3: Organización genómica del *Tomato torradovirus*…….9
Figura 4: Mapa de ubicación de los puntos de muestreo……11
Figura 5: Corrida electroforética en aa que muestra bandas de tamaño aproximado de 21–24 pb, correspondientes a siARN ….…….....14
Figura 6: Librerías de siARN en 2% de gel agarosa…….…….20
Figura 7: Diagrama de flujo del software VirusDetect …….....22
Figura 8: Diagrama de flujo de VirFind……..……..23
Figura 9: Planta correspondiente a la muestra Czo-004 con leve moteado clorótico en las hojas ……..…….... 35
Figura 10: Planta correspondiente a la Czo-040 con leve moteado clorótico y enrollamiento en las hojas……..…..….... 35
Figura 11: *Contigs* formados de la muestra Czo-004 mediante VD……...37
Figura 12: *Contigs* formados de la muestra Czo-040 mediante VD……..….37
Figura 13: *Primers* diseñados para completar el genoma de las muestras Czo-004 y Czo-040…….……..…….....38
Figura 14: Planta correspondiente a la muestra Czo-001 con clorosis y leve enrollamiento en las hojas 39
Figura 15: *Contigs* formados de la muestra Czo-001 mediante VD (ARN 1- *Torradovirus*). ……..……..40
Figura 16: *Contigs* formados de la muestra Czo-001 mediante VD (ARN2-*Torradovirus*)…….…….40
Figura 17: *Primers* diseñados para completar el genoma de la muestra Czo-117……..……...…41
Figura 18: Resultados de la amplificación por PCR de las muestras Czo-004 y Czo-040 utilizando los primers Badna-4 y Badna-6…….42
Figura 19: Amplificación por PCR de las muestras Czo-004 y Czo-040 con primers Badna-5…….…….42
Figura 20: Determinación de la temperatura óptima de PCR con el primer Badna-7……..……..43

Figura 21: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR de la muestra Czo-004……..…43
Figura 22: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR de la muestra Czo-040……..……..……..43
Figura 23: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR con primers Badna-5……..…….……44
Figura 24: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR con primers Badna-7……..……..……..…44
Figura 25: Gel de agarosa que contiene los productos de PCRs utilizando los los pares de primers Torra- 1 (3077 pb) y Torra2.1 (1761 pb)……...…45
Figura 26: Gel de agarosa que contiene el producto de PCR de los *primers* Torra- 3 (2160 pb)……..……..……..……45
Figura 27: Gel de agarosa que contiene el producto de PCR de los *primers* Torra-4 (1579 pb)……..……..……45
Figura 28: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR con primer Torra2……..……...46
Figura 29: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR con los primers Torra2.1 y Torra1…….....…..46
Figura 30: Digestión del vector pGEM-T Easy con EcoRI y productos de PCR con los primers Torra2.1, Torra2 y Torra4……..……..……46
Figura 31: Unión de las secuencias de *contigs* y las obtenidas mediante la secuenciación Sanger de la muestra Czo-004……..……..…47
Figura 32: Unión de las secuencias de los *contigs* y las obtenidas mediante la secuenciación tipo Sanger de la muestra Czo-040……..……..…48
Figura 33: Organización del genoma completo del nuevo Badnavirus obtenido de las muestras Czo-004 y Czo-040……..…….49
Figura 34: Ensamblaje de *contigs* y secuencias Sanger del ARN1 del nuevo Torradovirus (Czo-117) ……..…..…….51
Figura 35: Ensamblaje de *contigs* y secuencias Sanger del ARN2 del nuevo Torradovirus (Czo-117)…….....51
Figura 36: Organización genómica del ARN1 del nuevo Torradovirus identificado en Cusco……..……...53
Figura 37: Organización genómica del ARN2 del nuevo Torradovirus identificado en Cusco ……..…….53

Figura 38: Árbol filogenético del nuevo Badnavirus basado en las proteínas retrotranscriptasa y RNasa H……55
Figura 39: Árbol filogenético del nuevo Torradovirus basado en las proteínas RdRp ……57

ÍNDICE DE TABLAS

Tabla 1: Lugares de colecta de muestras de papa en el departamento de Cusco, noviembre 2016……..11
Tabla 2: Reactivos utilizados en la adenilación del adaptador 3'……..……16
Tabla 3: Reactivos utilizados en la ligación del adaptador 3' a los siARN……....16
Tabla 4: Reactivos utilizados en la reacción de ligación 5' a los siARNs……..…17
Tabla 5: Reactivos utilizados en la Transcripción reversa……..…17
Tabla 6: Reactivos utilizados en el PCR test……18
Tabla 7: Reactivos utilizados en el enriquecimiento de PCR y el primer que contiene el código de barra que identifica a cada muestra……..……..…19
Tabla 8: *Primers* diseñados para completar el genoma del nuevo *Badnavirus*……25
Tabla 9: *Primers* diseñados para completar el genoma del nuevo *Torradovirus*…….……....26
Tabla 10: *Primers Walking* diseñados para completar la secuenciación Sanger…….……..28
Tabla 11: Virus comunes identificados mediante el análisis con el *software* VirusDetect……....29
Tabla 12: Virus nuevos identificados mediante el análisis con el *software* VirusDetect……31
Tabla 13: Porcentaje de cobertura obtenido de las secuencias siARN al genoma del virus referencial, que son considerados como nuevos virus……..32
Tabla 14: Resultados obtenidos de la muestra Czo-061 utilizando los *softwares* VirFind y VirusDetect……..33
Tabla 15: Resultados obtenidos de la muestra Czo-088 utilizando los *softwares* VirFind y VirusDetect…….……....33
Tabla 16: Resultados obtenidos de la muestra Czo-117 utilizando los *softwares* VirFind y VirusDetect…….……....34
Tabla 17: Comparación de resultados obtenidos de la muestra Czo-129 utilizando los softwares VirFind y VirusDetect……..……....34
Tabla 18: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-004……..……..…....36

Tabla 19: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-040……..…..……37
Tabla 20: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-001……..……..……....39
Tabla 21: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-117……..……..…...40
Tabla 22: Porcentaje de identidad del genoma del *Badnavirus* obtenido de las muestras Czo-040 y Czo-004..……..……..…49
Tabla 23: Porcentaje de identidad del alineamiento de nucleótidos correspondiente a las secuencias amplificadas con el *primer* Torra 2……50
Tabla 24: Porcentaje de identidad del alineamiento del ARN 1 entre secuencias de USDA y muestras de Cusco..…..……..…52
Tabla 25: Porcentaje de identidad del ARN 2 entre secuencias de USDA y Cusco……..…..52
Tabla 26: Códigos de accesión de las secuencias de Torradovirus depositadas en NCBI……..…54
Tabla 27: Porcentaje de identidad de la región RT–RNasa H del nuevo *Badnavirus* frente a *Canna yellow mottle virus*…….……55
Tabla 28: Porcentaje de identidad de la región RdRp entre el nuevo *Torradovirus* y *Squash chlorotic leaf spot virus*……..56
Tabla 29: Porcentaje de identidad de la proteína de la cápside entre el nuevo torradovirus y *Squash chlorotic leaf spot virus* ……. 56
Tabla 30: Datos obtenidos de la colecta de las muestras de la provincia de Cusco…….……69
Tabla 31: Master mix empleada para la amplificación de los *primers* Badna-4 y Badna-6……..……..…80
Tabla 32: Master mix empleada para la amplificación de los *primers* BADNA5, BADNA7, TORRA1, TORRA2, TORRA3 y TORRA……..……80
Tabla 33: Temperatura y tiempo de *annealing* utilizados para la amplificación de los *primers* BADNA5, BADNA7, TORRA1, TORRA2, TORRA2.1, TORRA3 y TORRA4..……..……....81
Tabla 34: Resultados obtenidos de la secuenciación de las 144 muestras colectadas en Cusco……....…….82
Tabla 35: Código de barras utilizado para la preparación de librerías……100

RESUMEN

Introducción: El cultivo de papa es esencial para garantizar la seguridad alimentaria mundial, ya que es consumido por millones de personas. Sin embargo, este cultivo enfrenta diversas amenazas, siendo los virus uno de los patógenos más significativos. La nueva metodología de diagnóstico de secuenciación de alto rendimiento (HTS, por sus siglas en inglés) permite identificar tanto a los virus comunes como a los nuevos, logrando un diagnóstico efectivo. **Materiales y métodos:** Se recolectaron hojas de 144 plantas de papa de 12 diferentes localidades de Cusco para la preparación de librerías de siARN. El análisis bioinformático se llevó a cabo utilizando los *softwares* VirusDetect y VirFind. Se realizaron PCRs para completar los *gaps* de la secuenciación, y se utilizó el programa Mega X para construir el árbol filogenético. **Resultados:** Los resultados de VirusDetect revelaron una mayor cantidad de información sobre las secuencias virales en comparación con Virfind. El análisis filogenético demostró la existencia de un nuevo *Torradovirus* filogenéticamente cercano a *Squash clorotic leaf spot virus*, con un 71.52% de identidad entre ellos. Además, se reporta un nuevo *Badnavirus* que se halla cercano evolutivamente a *Canna yellow motle virus* con un porcentaje de identidad de 73.3% entre ellos. **Conclusión**: La secuenciación de alto rendimiento de siARN permitió identificar virus nuevos en plantas de papa.

Palabras claves: Virus, HTS, *Torradovirus*, *Badnavirus*, papa.

ABSTRACT

Introduction: Potato cultivation is essential for ensuring global food security as it is consumed by millions of people. However, this crop faces various threats, with viruses being one of the most significant pathogens. The new high-throughput sequencing (HTS) diagnostic methodology allows for the identification of both common and novel viruses, achieving effective diagnosis. **Materials and Methods:** Leaves were collected from 144 potato plants from 12 different localities in Cusco for the preparation of siRNA libraries. Bioinformatic analysis was performed using VirusDetect and VirFind software. PCRs were conducted to fill sequencing gaps, and the Mega X program was used to construct the phylogenetic tree. **Results:** VirusDetect results revealed more information about viral sequences compared to Virfind. The phylogenetic analysis demonstrated the existence of a new *Torradovirus* that is phylogenetically close to *Squash chlorotic leaf spot virus*, with a 71.52% sequence identity between them. The new Badnavirus is evolutionarily close to *Canna yellow mottle virus*, with 73.3% identity. **Conclusion:** High-throughput siRNA sequencing enabled the identification of new viruses in potato plants.

Keywords: Virus, HTS, Torradovirus, Badnavirus, potato

1. INTRODUCCIÓN

La papa constituye en la actualidad uno de los cinco cultivos más importantes del mundo y es consumida habitualmente por miles de millones de personas. (Xia, 2022). Sin embargo, este cultivo está sujeto a varias enfermedades virales. En este sentido, la detección temprana y precisa de virus en cultivos de papa es fundamental para garantizar la seguridad alimentaria y la salud de los cultivos. Los virus pueden causar enfermedades devastadoras en las plantas de papa, lo que resulta en pérdidas significativas de rendimiento entre un hasta un 90% y calidad de los cultivos (Matthews, 2019). Por lo tanto, la identificación oportuna de los virus es esencial para implementar medidas de control eficaces y mitigar los impactos negativos en la producción agrícola.

Existen varias metodologías que se utilizan para detectar y caracterizar virus en cultivos de papa, cada una con sus ventajas y limitaciones. Una de ellas la secuenciación de alto rendimiento de ARN de interferencia pequeños (siARN), que se ha convertido en una herramienta invaluable en el diagnóstico de virus debido a su capacidad para detectar una amplia gama de virus.

Esta técnica no solo permite la detección de virus conocidos, sino que también facilita la identificación de nuevos virus y variantes virales emergentes (Smith et al., 2017). Además, el uso de siARN aumenta la especificidad de las secuencias virales identificadas, ya que refleja la actividad de respuesta antiviral de las plantas (Barrero et al., 2017).

Los géneros *Torradovirus* y *Badnavirus* se han identificado en diversas plantas de importancia económica, como el tomate y el camote, respectivamente. Estas infecciones han sido asociadas con sintomatologías como clorosis foliar, necrosis y enanismo (Iskra-Caruana et al., 2010), afectando negativamente el rendimiento del cultivo. Considerando que el tomate pertenece a la familia Solanaceae, al igual que *Solanum tuberosum*, y que los torradovirus han sido reportados en especies de esta familia, resulta relevante caracterizar su presencia en papa para evaluar su posible impacto fitosanitario y en la producción.

En este trabajo se buscó identificar nuevos virus en variedades de *Solanum tuberosum* L. (Solanaceae) mediante High-Throughput Sequencing (HTS) de librerías preparadas de ARN pequeños de interferencia con posterior ensamblaje de las secuencias obtenidas mediante programas bioinformáticos (VirusDetect v1.6, VirFind y Geneious 11.1.5), y complementado con la secuenciación Sanger para completar los genomas virales y realizar análisis filogenéticos.

2. ANTECEDENTES

2.1 Importancia de la papa

La papa constituye en la actualidad uno de los cinco cultivos más importantes del mundo, con una superficie ocupada de más de 20,7 millones de hectáreas en todos los continentes (incluidas las zonas de mayor altitud de los países tropicales), y es consumida habitualmente por miles de millones de personas (Xia, 2022).

En Perú, el cultivo de papa representa el 25% del PBI agropecuario (Ministerio de Desarrollo Agrario y Riego,2023), hecho que lo convierte en un cultivo importante para la base económica y alimentaria para la población peruana.

En el 2020, la producción de papa en el Perú creció un 12.2% (INEI, 2022) entre los departamentos que destacaron por su mayor producción figuran Apurímac (131,5%), Huancavelica (64,3%), Cusco (27,0%), Puno (20,5%) y Ayacucho (18,9%), los cuales participaron con el 70,9% de la producción nacional de papa. También, se reportaron resultados positivos en Ica (37,6%), Lima (7,5%), Cajamarca (3,6%) y Áncash (1,8%) (INEI, 2022).

Los cultivos de papa se ven afectados por diversos patógenos que comprometen su crecimiento y producción, entre ellos hongos, bacterias, nematodos y virus. Estos últimos constituyen uno de los grupos más importantes debido a su amplia distribución, facilidad de transmisión y significativo impacto en el rendimiento y la calidad del cultivo (Jones & Naidu, 2019).Los virus causan pérdidas significativas en el rendimiento y calidad en una amplia variedad de cultivos, por lo tanto, la detección e identificación de virus es una faceta crucial de la producción exitosa de cultivos y de gran

importancia en términos de seguridad alimentaria mundial (Jones et al., 2017).

2.2 Técnicas de diagnóstico

Existen diversas técnicas que permiten identificar la presencia de un virus en las plantas como las pruebas serológicas, los bioensayos, la detección visual y las técnicas moleculares, como la RT-PCR, PCR, los cuales han aumentado la velocidad y la precisión del diagnóstico viral, pero en su forma dirigida, estas técnicas solo enfocan principalmente en la detección de virus previamente conocidos (Mumford et al., 2006). Actualmente se han desarrollado métodos para la detección simultánea de múltiples virus mediante HTS.

Con el desarrollo de diferentes protocolos de HTS ha sido posible detectar múltiples virus conocidos y desconocidos presentes en los diversos cultivos, lo cual ha ido de la mano en el desarrollo de métodos bioinformáticos rápidos y robustos para permitir la eliminación de las secuencias del huésped y la clasificación del virus (Jones et al., 2017). Además, la creación de base de datos de transcriptomas y genomas ha contribuido al descubrimiento de nuevos virus (Roossinck et al., 2015).

2.3 ARN de interferencia pequeños

Los ARN de interferencia pequeños (siARN) juegan un rol fundamental en las plantas, ya que regulan la expresión génica y las protegen contra las infecciones virales. Los siARN, pueden dirigirse específicamente a secuencias virales complementarias, desencadenando la degradación del ARN viral o reprimiendo su traducción (Ruiz-Ferrer & Voinnet, 2009).

El mecanismo mediante el cual actúa este silenciamiento se resume en la Figura 1, donde se muestra la biogénesis y acción de los siARN derivados de virus de ARN (Pooggin, 2018). La ARN polimerasa dependiente de ARN viral (vRdRP) sintetiza cadenas complementarias de ARN a partir de una plantilla viral, generando ARN de doble cadena (dsARN).

Este dsARN es reconocido por las enzimas DICER, principalmente DCL4 y DCL2, que lo procesan en pequeños fragmentos de 21–24 nt denominados siARN. Los siARN primarios resultantes se incorporan al complejo RISC

(complejo de silenciamiento inducido por ARN), formando un complejo activo que contiene una de las hebras del dúplex. Finalmente, este complejo se dirige de manera específica al ARN mensajero (ARNm) complementario, promoviendo su degradación o la represión de su traducción (Bologna & Voinnet, 2014).

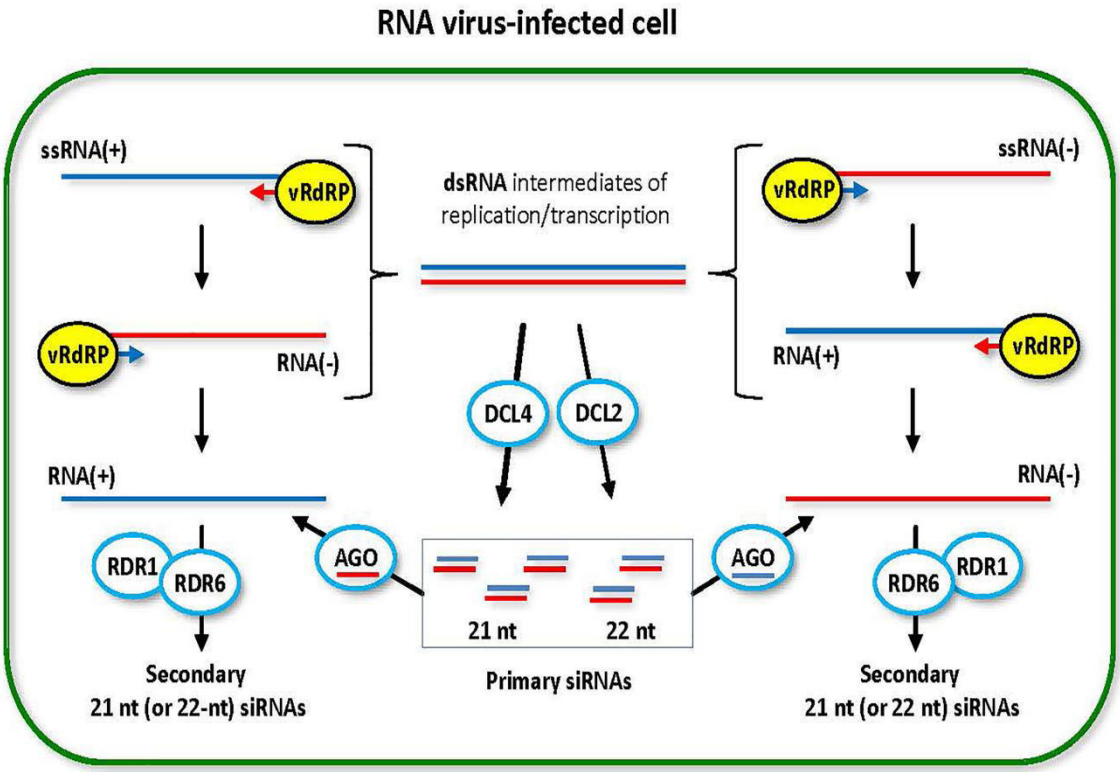


Figura 1: Mecanismo de formación de los siARN. (Fuente: Pooggin, 2018)

En 2009, dos estudios demostraron definitivamente que los siARN derivados de virus podrían usarse para detectar virus vegetales conocidos y previamente no caracterizados en plantas infectadas (Donaire et al., 2009; Kreuze et al., 2009). Desde entonces, este enfoque se ha utilizado ampliamente para detectar una variedad de virus que infectan las plantas, incluso ha sido posible reconstruir los genomas de virus de ADN como *Caulimovirus* y *Geminivirus* (Kreuze et al., 2009; Loconsole et al., 2012; Seguin et al., 2014). Se cree que la razón de esto se debe a que las regiones codificantes y no codificantes de ADN de doble cadena (dsDNA) viral circular se transcriben en ambas orientaciones en el núcleo para generar

precursores de ARN de doble cadena (dsRNA) de los siARN virales (Seguin et al., 2014).

Una de las principales ventajas del enfoque basado en siARN es su alta sensibilidad para detectar virus conocidos y potencialmente no caracterizados, dependiendo de la abundancia de siARN presentes y de la profundidad de secuenciación empleada (Kreuze et al., 2009). Esta metodología se basa en el enriquecimiento de secuencias virales derivadas de la respuesta antiviral de la planta, seguido de la secuenciación y el ensamblaje de ARN pequeños.

El análisis de siARN ha revolucionado el diagnóstico de virus en los cultivos agrícolas, ya que al detectar y analizar los siARN derivados de virus en las plantas, se puede identificar de manera precisa la presencia de patógenos virales. Esta técnica de diagnóstico basada en siARN ofrece una herramienta sensible y específica para monitorear y controlar las enfermedades virales en la agricultura, lo que contribuye significativamente a la seguridad alimentaria (Miozzi et.al., 2013).

2.4 HTS en la detección de virus

La aplicación de secuenciación de alto rendimiento (HTS), también conocida como secuenciación de próxima generación, ha demostrado ser muy exitosa para el descubrimiento de virus para resolver la etiología de la enfermedad en muchos cultivos agrícolas (Fox et al., 2018).

La secuenciación de siARNs ofrece nuevas oportunidades para identificar virus conocidos o desconocidos (novedosos) que infectan a las plantas, incluso en plantas asintomáticas con títulos virales extremadamente bajos (Aboul-Ata et al., 2011). HTS puede proporcionar miles o millones de secuencias de siARN a partir de materiales vegetales infectados con virus (Haididi et al., 2016). Cuando los siARN derivados de virus son lo suficientemente abundantes, se pueden ensamblar fragmentos del genoma del virus (Kreuze et al., 2009). Este enfoque requiere herramientas bioinformáticas relativamente sofisticadas.

Este proyecto se centrará en la obtención y caracterización del genoma de nuevos virus pertenecientes a los géneros *Badnavirus* y *Torradovirus*, mediante HTS en plantas provenientes de Cusco, Perú.

2.5 Badnavirus

Los *Badnavirus* infectan una amplia gama de plantas de cultivo económicamente importantes en todo el mundo (Lockhart & Olsweski , 1994). Por ejemplo, se ha reportado su presencia en cacao (*Theobroma cacao*), canna ornamental (*Canna indica*), ñame (*Dioscorea* spp.), higuera (*Ficus carica*) y grosellero (*Ribes* spp.), entre otras especies. Los síntomas causados por los *Badnavirus* son variables dependiendo del huésped, sus cultivares, especies de virus y condiciones ambientales (Alangar et al., 2023). En la mayoría de los casos, los síntomas son leves a moderados que incluyen el moteado clorótico o vetas necróticas, deformación de las hojas y una longitud reducida del entrenudo que conduce al retraso del crecimiento de las plantas (Bhat et al., 2016).

Los *Badnavirus* poseen un genoma de ADN circular de doble cadena que tiene un tamaño entre 7200–9200 pb (Bhat et al., 2016) en el cual se encuentran tres marcos de lectura abiertos (ORF), el ORF1 y ORF2 codifican para dos pequeñas proteínas de función desconocida de 20.8 kDa y 14.5 kDa respectivamente, y el ORF 3 codifica para una poliproteína de 208 kDa (ORF3) (Harper & Hull, 1998) que engloba a la proteína de movimiento célula-célula, la proteína de la cápside, la aspartil proteinasa y la replicasa con los dominios transcriptasa inversa y RNAsa H (Medberry et al., 1990) . La organización genómica del *Canna yellow mottle virus* (CaYMV), un miembro del género *Badnavirus*, se muestra en la Figura 2 (Zhang et al., 2017).

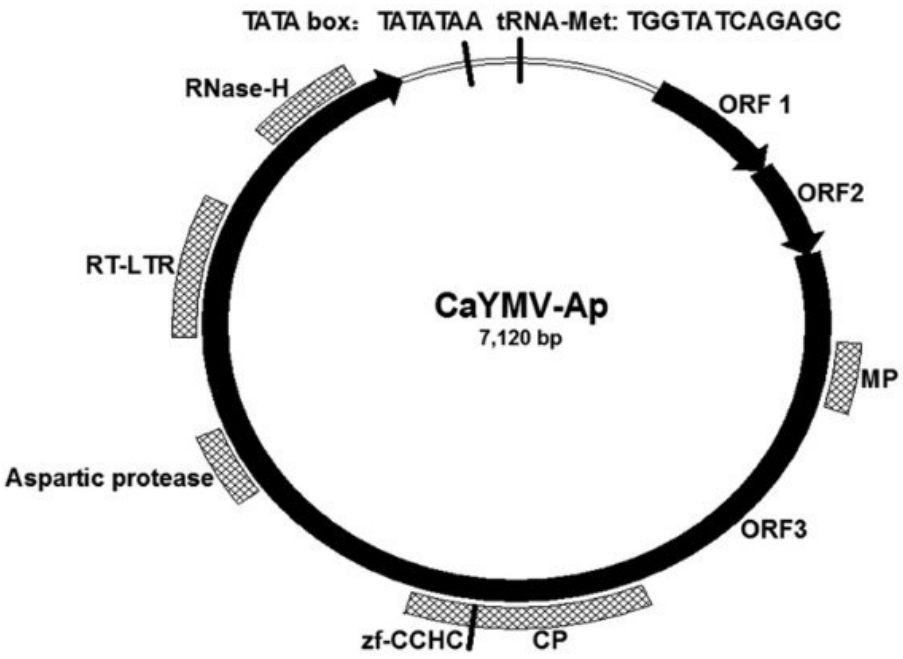


Figura 2: Organización del genoma del Badnavirus *Canna yellow mottle virus* (CaYMV) (Fuente: Zhang et al., 2017)

Se tiene conocimiento que los *Badnavirus* están presentes como secuencias integradas en algunos genomas de la planta huésped denominándose *Badnavirus* endógenos (Hohn et al., 2008).

La presencia de estas secuencias integradas representa un desafío en estudios basados en secuenciación y diagnóstico molecular, ya que puede dificultar la diferenciación entre formas virales activas (episomales) y fragmentos virales integrados en el genoma vegetal. Esta problemática ha sido ampliamente discutida en estudios sobre la familia *Caulimoviridae*, donde se señala la necesidad de estrategias moleculares y bioinformáticas para discriminar entre secuencias endógenas y virus infecciosos (Geering et al., 2014). En concordancia con estos enfoques, el presente estudio incorporó análisis bioinformáticos complementados con pruebas moleculares con el fin de diferenciar secuencias virales activas de posibles integraciones endógenas.

Los criterios de demarcación de las especies se definen según King et al., (2012):

- Diferencia en los hospederos

- Diferencia en más del 20% en la secuencia de nucleótidos que codifica al gen de la polimerasa (RT+ RNAsa H).
- Diferencias en las secuencias de productos génicos.
- Diferencias en las especificidades de los vectores.

2.6 Torradovirus

El género *Torradovirus* fue establecido en 2009 por Sanfaçon et al. (2009) tras la identificación de *Tomato torrado virus*, agente causal de la enfermedad del torrado en plantas de tomate (Verbeek et al., 2007). Esta enfermedad se caracteriza por necrosis foliar y malformación de frutos. Las hojas de las plantas infectadas muestran manchas necróticas y pueden tornarse de color bronce, mientras que los frutos pueden desarrollar manchas y deformaciones. Estos síntomas pueden provocar una disminución significativa en el rendimiento y la calidad de los cultivos de tomate (Massart et al., 2014).

El genoma segmentado y bipartito consta de dos moléculas monocatenarias de ARN (ssRNA) que contienen tres ORFs flanqueados por Regiones No Traducidas (UTRs) (Figura 3). Ambas UTR varían en tamaño entre las diferentes especies de *Torradovirus* con una UTR de 5' que varía de 106 a 140 nt y una UTR de 3' que varía de 628 a 1210 nt (Verbeek et al., 2008).

El ARN 1 tiene una longitud de 7.5–7.8 kb y contiene un único ORF con los dominios funcionales característicos del cofactor de proteasa (Pro-Co), helicasa, proteasa y RdRp en el extremo C-terminal. El ARN 2 tiene una longitud de 4–5.5 kb y contiene dos ORFs superpuestos. La proteína codificada hipotéticamente por ORF1 no tiene homología con otras proteínas virales ni con ningún dominio funcional conocido (Verbeek et al., 2007), mientras que ORF2 codifica un polipéptido que contiene los dominios de tres proteínas de cápside y una posible proteína de movimiento (Verbeek et al., 2007). Las proteínas de la cápside se denominan Vp35, Vp26 y Vp23 según sus pesos moleculares predichos. La posible proteína de movimiento del *Torradovirus* no tiene homología con las proteínas de movimiento de *Comoviridae*, aunque contiene un motivo LxxPxL conservado (Verbeek et al., 2007).

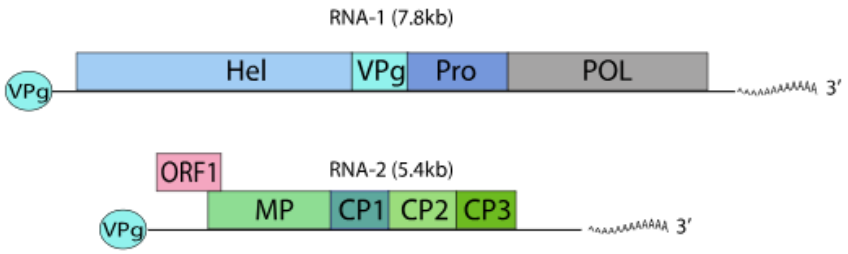


Figura 3: Organización genómica del *Tomato torradovirus* (Fuente: ExPASy, s. f.).

De acuerdo con los criterios de delimitación propuestos por Sanfancon et al., (2009), en promedio, deben presentar menos del 75% de identidad de secuencia de aminoácidos de la proteína de la cápside y menos del 80% de identidad de secuencia de aminoácidos en la región de proteinasa-polimerasa.

3. HIPOTESIS

H(+) En variedades de *Solanum tuberosum* L. provenientes de Cusco circulan virus no reportados previamente, los cuales pueden ser detectados mediante análisis de siARN por secuenciación de alto rendimiento (HTS) y confirmados mediante secuenciación dirigida.

H(-) No se detectan virus no reportados previamente en las muestras analizadas bajo el enfoque basado en siARN y secuenciación dirigida.

4. OBJETIVOS

4.1 Objetivo general

Identificar y caracterizar virus no reportados previamente en variedades de *Solanum tuberosum* L. provenientes de Cusco-Perú mediante secuenciación de alto rendimiento (HTS) de ARN de interferencia, así como obtener y analizar sus genomas completos para su caracterización molecular y filogenética.

4.2 Objetivos específicos

- Obtener librerías de siARN extraídas de variedades de papa provenientes de Cusco, Perú.

- Obtener lecturas de secuenciación HTS de siARN y realizar el control de calidad, ensamblaje y análisis bioinformático para la detección de secuencias virales utilizando VirusDetect v1.6.
- Identificar virus candidatos mediante el análisis de los resultados de la secuenciación de alto rendimiento utilizando los *softwares* VirusDetect v1.6 y VirFind.
- Completar el genoma de virus candidatos priorizados en función de su porcentaje de cobertura genómica y longitud de contigs, mediante secuenciación Sanger.
- Realizar el análisis filogenético con los genomas de los virus detectados.

5. MATERIALES Y MÉTODOS

5.1. Colecta de muestras

Se colectaron hojas de 144 plantas de papa provenientes de cinco provincias del departamento de Cusco (Anta, Paruro, Cusco, Paucartambo y Calca), mediante un muestreo no probabilístico dirigido.

En cada localidad se evaluaron hojas de 11 plantas; excepto en la localidad de Paucana (provincia de Paucartambo), donde se evaluaron hojas de 12 plantas. El muestreo priorizó plantas con sintomatología compatible con virosis, con el fin de incrementar la probabilidad de detección de agentes virales.

En total, se evaluaron 144 plantas distribuidas en 13 localidades pertenecientes a cinco provincias del departamento de Cusco. El muestreo incluyó nueve variedades de papa, cuya distribución por provincia, distrito y localidad se presenta en la Tabla 1, la Figura 4 y el Anexo 1.

Debido a que el muestreo fue no probabilístico y dirigido, el estudio no tuvo como objetivo estimar prevalencias poblacionales, sino detectar la presencia de virus en las muestras evaluadas.

Tabla1: Lugares de colecta y variedades de papa muestreadas en el departamento de Cusco, noviembre de 2016.

Provincia	Distrito	Localidad	Variedades de papa	Total muestras
Anta	Limatambo	Huerta	Yungay	11
Anta	Limatambo	Cartonpata	Cica	11
Anta	Limatambo	Ayaviri (Pampacona)	Ccompis	11
Anta	Ancahuasui	San Martin	Cica	11
Paruro	Huanoquite	Tiwicte	Cica	11
Paruro	Huanoquite	Tiwicte 2	Cica	11
Paruro	Yaurisque	Taray	Cica	11
Cusco	Santiago	Occopata-Cacchona	Ccompis/Huayro/Mactillo	11
Paucartambo	Challabamba	Chuza	Mactillo/Amarillis	11
Paucartambo	Challabamba	Pilco	Canchan	11
Paucartambo	Challabamba	Sunchubamba	Cica	11
Paucartambo	Colquepata	Paucana	Cica/Alianza	12
Calca	Pisac	Sacaca	Yungay	11
			Total	**144**

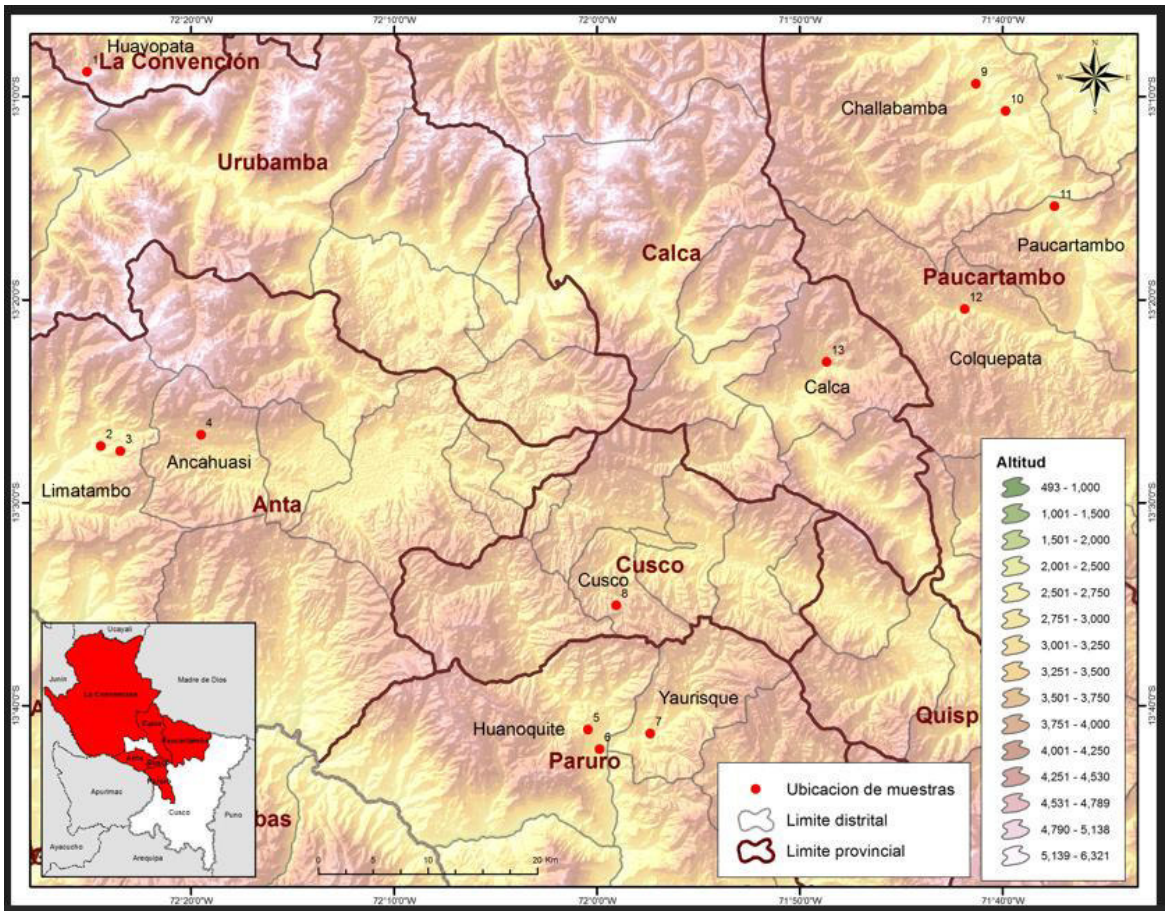


Figura 4: Mapa de ubicación de los puntos de muestreo. Los puntos rojos indican la ubicación de los sitios de muestreo.

De cada planta se recolectaron tres hojas (una de la parte apical, una de la parte media y una de la parte basal). Las muestras fueron colocadas en bolsas herméticas tipo Ziploc® que contenían papel filtro y sílica gel con indicador de humedad (gel anaranjado) para su desecación. El secado se inició inmediatamente después de la colecta y el sílica gel fue reemplazado hasta en tres oportunidades, según el cambio de color del indicador, con el fin de asegurar una deshidratación adecuada del tejido vegetal.

Una vez secas, las muestras fueron almacenadas a temperatura ambiente en recipientes herméticos (táperes) para su transporte al laboratorio, debidamente identificadas con el número de muestra y la localidad correspondiente. El tiempo transcurrido entre la colecta y el procesamiento en laboratorio fue de aproximadamente 3 días.

Se obtuvieron datos de colecta como la ubicación, el cultivar, edad de la planta, entre otras características, las cuales fueron registrados en el Anexo 1.

5.2 Obtención de los ARN pequeños

La obtención de los siARN se realizó utilizando el protocolo descrito por Gutiérrez y Kreuze (2014), el cual es utilizado en el laboratorio de Virología del Centro Internacional de la Papa (CIP).

5.2.1. Extracción de ARN total

El proceso implicó la maceración de 40 mg de hojas secas de papa en 1.5 ml de Trizol® en bolsas de plástico durante 10 minutos. Posteriormente, las muestras fueron molidas, centrifugadas a 12,000 rpm por 5 minutos y el sobrenadante se transfirió a un nuevo tubo Eppendorf®. Se agregaron 500 µL de cloroformo, se homogenizó y se centrifugó a 12,000 rpm por 10 minutos a 4 °C, separando dos fases. La fase acuosa se transfirió a un nuevo tubo y se le añadieron 500 µL de isopropanol, incubando y centrifugando nuevamente para obtener un *pellet*, este fue lavado con etanol al 75%, centrifugado y secado a temperatura ambiente antes de ser almacenado a -70 °C.

5.2.2. Enriquecimiento de los ARN pequeños

El pellet fue suspendido en 250 µL de agua libre de nucleasa, se dejó en movimiento por 2 horas a 80 rpm a 4 °C para su homogenización. Seguidamente se le adicionó 250 µL de LiCl 4M y se incubó toda la noche a 4 °C.

Posteriormente, esta mezcla se centrifugó a 12 000 rpm por 20 minutos a 4 °C, el sobrenadante el cual contenía a los ARN pequeños fue transferido a un nuevo tubo Eppendorf® de 1.5 ml. El *pellet* que contenía a los ARN largos se lavó dos veces con etanol al 75% y centrifugados a máxima velocidad por 10 minutos a 4°C, y secados a temperatura ambiente para luego ser almacenados a -70 °C. Los ARN largos obtenidos fueron utilizados en pruebas de confirmación y para completar los genomas de virus de ARN de interés, como los pertenecientes al género *Torradovirus*. En el caso de *Badnavirus*, debido a su naturaleza como virus de ADN, se trabajó con ADN total extraído mediante el método CTAB (véase sección 5.8).

El tubo que contenía los ARN pequeños (sobrenadante) se le adicionó 500 µL de Isopropanol e incubado a temperatura ambiente por 10 minutos, seguidamente fue centrifugado a máxima velocidad durante 20 minutos a 4 °C. Terminado este procedimiento, se eliminó el sobrenadante, para luego lavar el pellet utilizando 500 µL de etanol al 75% y centrifugarlo a 14000 rpm durante 10 minutos a 4 °C. Culminado este paso, se dejó secar a temperatura ambiente y se almacenó a -70 °C.

5.2.3. Aislamiento de ARN de interferencia (siARN)

El *pellet* que contiene los pequeños ARN fue disuelto en 30 µL de agua libre de nucleasa y fueron corridos en gel de agarosa al 3.5%

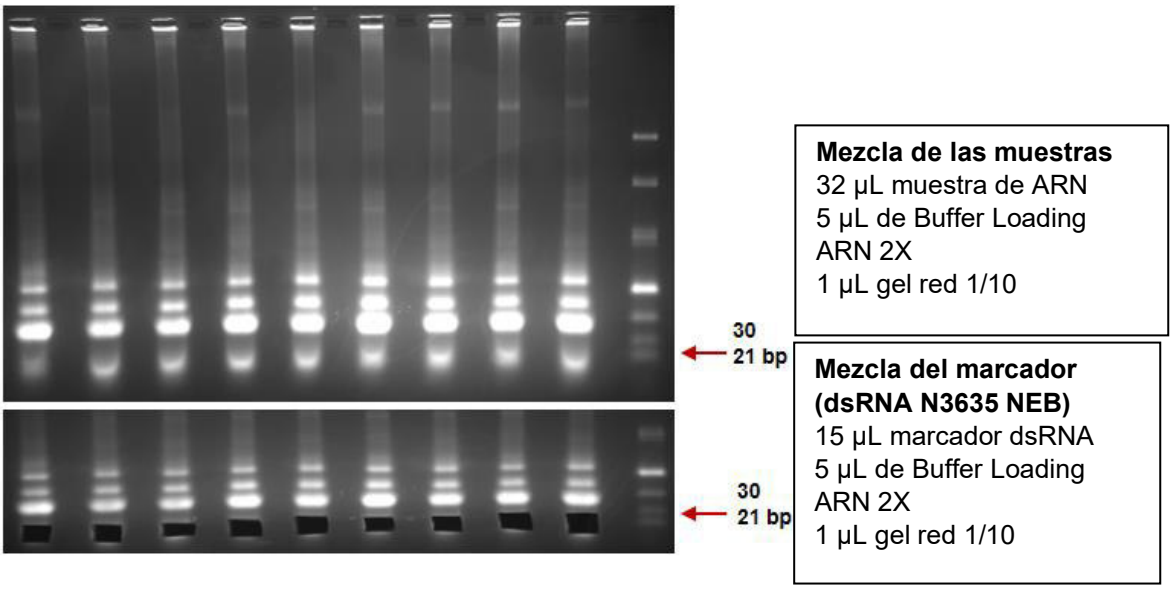


Figura 5: Corrida electroforética en gel de agarosa que muestra bandas de tamaño aproximado de 21–24 pb**,** correspondientes a siARN.

Los siARN tienen un tamaño entre 21-24 pb, en la figura 5 se observan bandas que se encuentran en ese rango, por lo cual estas bandas se aislaron utilizando un bisturí para cada muestra. El fragmento de gel de agarosa que contiene la banda (20 – 25 pb) se colocó en una columna de extracción (Freeze N Squeeze™ DNA gel Extraction Spin Clomuns (Bio RAD)) de gel, y puesto a congelar a -20 °C por 20 minutos, y centrifugado a 14 000 rpm por 5 minutos a temperatura ambiente.

El ARN fue colectado en un tubo colector y fue llevado hasta un volumen de 500 µL utilizando agua libre de nucleasa, para luego adicionar 250 µL de cloroformo y se mezcló vigorosamente. Se dejó incubar en hielo por 3 minutos, luego el tubo fue centrifugado a 12 000 rpm por 10 minutos a 4°C.

En el tubo se formaron dos fases, la fase acuosa (parte superior) fue transferida a un tubo Eppendorf® de 1.5 ml estéril y se le adicionó 1 µL de glicógeno (2 µg/µL), 1 ml de etanol al 100% y 40 µL de NaAc al 3 M. Se dejó incubar por una hora a -70 °C. Posteriormente el tubo fue centrifugado a 14000 rpm por 20 minutos a 4°C, se eliminó el sobrenadante y se le adicionó 600 µL de etanol al 75 %. Se volvió a

centrifugar a máxima velocidad por 10 minutos a 4°C, se eliminó el sobrenadante y el pellet se dejó secar a temperatura ambiente.

Posteriormente, el pellet fue resuspendido en 15 µL de agua libre de nucleasa.

5.3 Preparación de librerías de siARN

El protocolo que se utilizó para la preparación de la librería fue descrito por (Chen et al. 2012), al cual se le realizó algunas modificaciones descritas en el protocolo de laboratorio de Virología del CIP (Gutierrez y Kreuze, 2014).

En la preparación de las librerías no se incluyó un control negativo; no obstante, las muestras fueron procesadas simultáneamente con otras provenientes de distintas provincias del Perú que resultaron negativas para infección viral, lo que permitió descartar contaminación cruzada. En las etapas de RT-PCR y PCR se emplearon como controles negativos NFW (*nuclease-free water*) y ARN y ADN extraídos de tejido de planta sana.

La trazabilidad de las muestras se aseguró mediante la asignación de códigos de barra (*barcodes*) únicos para cada muestra, cuyo mapa de correspondencia muestra-barcode se presenta en el Anexo 6.

Los *primers* y adaptadores fueron sintetizados por la empresa Integrated DNA Technologies (IDT) con pureza HPLC, excepto los *primers* utilizados en el PCR test.

5.3.1 Adenilación del adaptador 3´

El adaptador 3' (5'-/5Phos/CAGATCGGAAGAGCACACGT/3AmMO/-3') fue diluido con Buffer de Elución (Buffer Tris-HCl pH 8, Roche®) hasta obtener una concentración final de 100 µM (stock).

Posteriormente se preparó una reacción de adenilado en un tubo de PCR (Tabla 2), los componentes se mezclaron por pipeteo y fueron incubados a 25 °C en el termociclador (Verity 96 well Thermal cycler).

Tabla 2: Reactivos utilizados en la adenilación del adaptador 3'.

Componentes	Reacción de adenilado Vol (µl)	Concentración final
Adaptador 3' (100 µM)	10	10 µM adaptador 3
10X Buffer ARN ligasa	10	1x Buffer ARN ligasa
10 µM ATP	10	1µM ATP
Agua libre de Nucleasa	20	------
50% PEG 8000	40	20 % PEG 8000
T4 ARN ligasa 1 (10U/µL)	10	1 U/µL T4 ARN ligasa 1

Se purificó el adaptador 3´ del producto de la reacción de adenilado mediante la adición de agua y cloroformo, seguido de centrifugación. La fase acuosa se transfirió a un tubo Eppendorf® y se trató con isopropanol y glicógeno, incubándola y centrifugándola para obtener un *pellet*. Tras lavarlo con etanol, se centrifugó nuevamente para completar la purificación.

Posteriormente el *pellet* se dejó secar a temperatura ambiente y se resuspendió con 30 µL de agua libre de nucleasa. Luego se midió la concentración utilizando el equipo Nanodrop 1000 y se procedió a ajustar la concentración hasta obtener 10 µM.

5.3.2 Ligación del adaptador 3´adenilado

Los siARNs y el adaptador 3' adenilado fueron desnaturalizados a 65 °C por 45 segundos, y colocados inmediatamente en hielo.

Posteriormente se preparó la reacción de ligación (Tabla 3).

Tabla 3: Reactivos utilizados en la ligación del adaptador 3' adenilado a los siARN.

Componentes	Reacción de ligación3' Vol. (µL)	Concentración final
siARN	10	
Adaptador 3' adenilado 10 µM	1	0.5 µM Adaptador 3´adenilado
10X T4 Buffer ARN ligasa (sin ATP)	2	1x T4 Buffer ARN ligasa (sin ATP)
50% PEG 8000	6	15 % PEG 8000
ARNasa out (40U/µL)	1	2 U/µL ARNasa out
T4 ARN ligasa 2 truncada (200U/µL)	1	10 U/µL ligasa 2

La mezcla se homogenizó mediante pipeteo y se incubó toda la noche (mínimo 16 horas) a 18 °C utilizando un termociclador, luego se adicionó 0.5 µL del *primer* de transcripción reversa 5'-/5AmMC6/ACGTGTGCTCTTCCGATCTG-3' (20 µM), y se inactivo la reacción a 65 °C por 15 min, para luego incubar a 25°C por 10 minutos.

5.3.3 Ligación del adaptador 5´

El adaptador 5´ (5-GUUCAGAGUUCUACAGUCCGACGAUC-3´) fue diluido hasta una concentración de 10 µM y desnaturalizado a 65 °C por 45 segundos. Posteriormente se preparó la reacción de ligación del adaptador 5´ (Tabla 4).

Tabla 4: Reactivos utilizados en la reacción de ligación del adaptador 5' a los siARNs.

Componentes	Reacción de ligación5 Vol. (µL)	Concentración final
Producto de ligación 3´	21.5	
Adaptador 5' (10uM)	1	0.33 µM 5' adaptador
10X T4 Buffer ARN ligasa (sin ATP)	1	1x T4 Buffer ARN ligasa (sin ATP)
ATP (stock 10 µM)	3	1 µM ATP
Agua libre de Nucleasa	1.5	----
ARNasa out (40 U/µL)	1	1.3 U/µL ARNasa out
T4 ARN ligasa 1 (10 U/µL)	2	0.65 U/µL T4 ARN ligasa 1

Se mezcló por pipeteo y se incubó a 25 °C por 4 horas en el termociclador (Veriti 96 well termal cycler).

5.3.4 Transcripción reversa

Se preparó la siguiente reacción de transcripción reversa (Tabla 5):

Tabla 5: Reactivos utilizados en la Transcripción reversa

Componentes	Reacción de transcripción Vol. (µL)	Concentración final
5' reacción de ligación	31	
dNTP (10 mM)	2	0.44 mM dNTP
DTT (0.1M)	3	6.35 mM DTT
SuperScript III First strand reaction buffer (5X)	9	1X First strand Buffer
ARNasa out (40 U/µL)	0.5	0.44 U/µLRNAsa out
SuperScript III rev-transcriptasa (200 U/µL)	1	4.35 U/µL Rev-transcriptasa

La reacción se mezcló por pipeteo y se incubó a 50 °C por una hora, la reacción se inactivo incubando a 75 °C por 15 min en el Termociclador (Veriti 96-Well Thermal Cycler).

5.3.5 PCR TEST

Se utilizó el siguiente par de *primers* a la concentración de 10 µM:

Forward sRNA PCR1:

AATGATACGGCGACCACCGAGATCTACACGACAGGTTCAGAGTTCT ACAGTCCGACGA*T*C

Reverse PCR2_1:

CAAGCAGAAGACGGCATACGAGA**TATCACG**GTGACTGGAGTTCAGA CGTGTGCTCTTCCGATC*T

Se preparó la siguiente reacción de PCR (Tabla 6).

Tabla 6: Reactivos utilizados en el PCR test.

Componentes	Reacción de PCR vol. (uL)	Concentracion final
Primer Forward sRNA PCR 1 (10 µM)	1	0.4 µM Primers
Primer Reverse PCR2_1 (10 µM)	1	0.4 µM Primers
dNTPs (10 mM)	1	0.4 mM dNTPs
5X Buffer HF	5	1X Buffer HF
Agua libre de nucleasa	14	---
$MgCl_2$ (50 mM)	0.25	0.5 mM $MgCl_2$
Phusion DNA polimearsa (2 U/µL)	0.25	0.02 U/µLDNA polimerasa
Producto de RT	2.5	---

El programa de amplificación comprendió una desnaturalización inicial a 95 °C por 2 min, seguida de 20 ciclos de desnaturalización a 98 °C por 11 s, alineamiento a 65 °C por 30 s y extensión a 72 °C por 30 s, finalizando con una extensión a 72 °C por 2 min.

La confirmación de la amplificación de librerías de siARN se realizó corriendo los productos amplificados de PCR test en un gel de agarosa al 2% en Buffer TAE utilizando 5 µL de SalB, con 1 µL de GelRed® diluido 1/10 y 25 µL de producto de PCR. Con lo cual se pudo escoger las muestras para realizar el siguiente paso.

5.3.6 Enriquecimiento de PCR

Se preparó la siguiente reacción de PCR (Tabla 7), para cada muestra se utilizó el mismo *forward* sRNA PCR 1 sintetizados utilizando el método de purificación mediante HPLC y distintos *primers reverse* (código de barra), identificados como PCR2_1 al PCR2_48 (Anexo 6). Los *primers reverse* se diferencian en los 6 nucleótidos centrales del *primer* (código de barra), con cual se puede reconocer la muestra cuando se obtienen los resultados de la secuenciación:

(CAAGCAGAAGACGGCATACGAGA**NNNNNN**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T)

Tabla 7: Reactivos utilizados en el enriquecimiento de PCR y el primer que contiene el código de barra que identifica a cada muestra.

Componentes	Reacción de PCR Vol. (µL)	Concentración final
Primer Forward sRNA PCR 1 (10 µM)	1	0.3 µM Primers
Primer reverse PCR2_* (10 µM)	1	0.3 µM Primers
dNTPs (10 µM)	1	0.33 µM dNTPs
5X Buffer HF	6	1X Buffer HF
Agua libre de nucleasa	10.2	
$MgCl_2$ (50 mM)	0.3	0.5 mM MgCl2
Phusion DNA polimerasa (2 U/µL)	0.5	0.04 U/µL DNA polimerasa
Producto de RT	10	

El programa de amplificación comprendió una desnaturalización inicial a 95 °C por 2 min, seguida de 20 ciclos de desnaturalización a 98 °C por 12 s, alineamiento a 65 °C por 30 s y extensión a 72 °C por 30 s, finalizando con una extensión a 72 °C por 2 min.

Los productos amplificados fueron corridos en geles de agarosa 2% en buffer TBE utilizando 4 µL de SalB, 1 µL de Gel red diluido 1/10 y 30 µL del producto de PCR (Figura 6), los cuales fueron purificados con High Pure PCR Product Purification Kit (Roche®), siguiendo las instrucciones del proveedor con algunas modificaciones (ver sección 11.2).

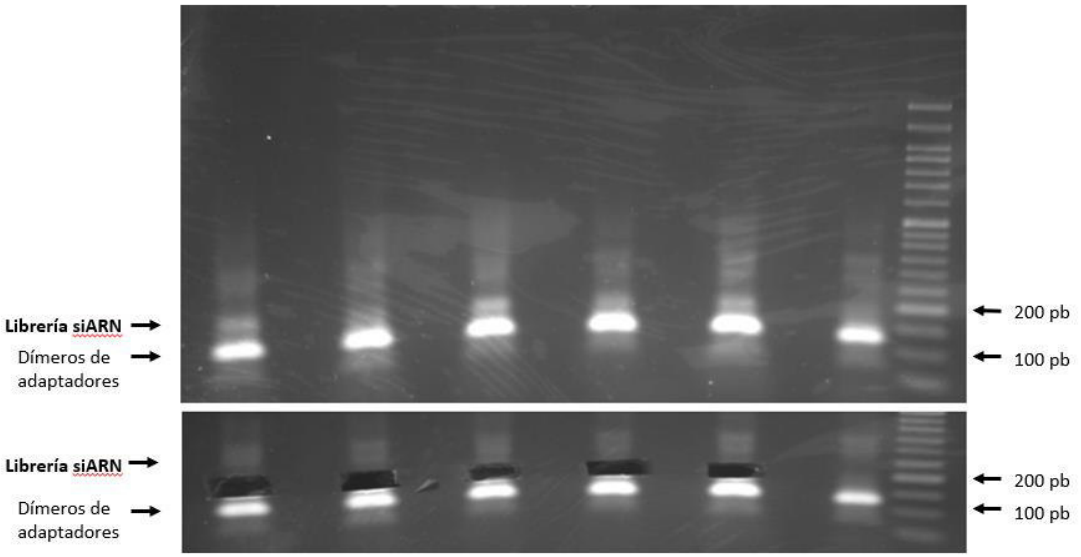


Figura 6: Librerías de siARN en 2% de gel agarosa, formando una banda en el gel, que fue cortado para su purificación. El marcador utilizado fue 50 pb DNA ladder (Invitrogen®).

5.3.7 Cuantificación de ADN purificado

Cada muestra fue cuantificada utilizando un fluorómetro Qubit 2.0 (Invitrogen®, USA) (ver sección 11.3), y luego se preparó la mezcla (*pool*) de librerías. Se prepararon en total 3 librerías constituidas de 48 muestras cada una. Cada mezcla de librería se preparó colocando la misma cantidad de muestra (12 ng) de cada una de ellas en un solo tubo Eppendorf®, el cual fue identificado con su respectiva etiqueta y enviado a secuenciar.

5.4 Secuenciación de alto rendimiento

Las librerías fueron sometidas a HTS, la cual fue realizado por la empresa Fasteris (Suiza), utilizando la plataforma IIlumina HiSeq 2500/4000 systems.

5.5 Análisis bioinformáticos mediante VirusDetect y VirFind

Las lecturas de secuenciación (reads) de siARN obtenidos mediante HTS se descargaron en el servidor del Centro Internacional de la Papa utilizando los *softwares* WinSCP 5.13 y PuTTY como interfases para la conexión a la nube (servidor Amazon).

El análisis con VirusDetect se realizó utilizando la versión 1.6 en entorno Linux (Zheng et al., 2017). Para la limpieza de las lecturas de siRNA se empleó el script *sRNA_clean.pl*, incluido en el paquete del software, el cual eliminó secuencias con bases ambiguas o de baja calidad, lecturas sin

adaptador, adaptadores vacíos y secuencias con longitud menor a 15 pb. La evaluación de la calidad de las lecturas filtradas se realizó mediante el programa FastQC versión 0.11.9.

https://www.bioinformatics.babraham.ac.uk/projects/download.html#fastqc.

El análisis con VirFind (plataforma web) se realizó utilizando los parámetros establecidos en el formulario de envío. El pipeline incluyó control de calidad y *trimming* de lecturas, mapeo a genoma de referencia, ensamblaje *de novo* y comparación contra bases de datos virales de NCBI mediante Blastn y Blastx, empleando un *e-value* por defecto de 0.01. Las secuencias sin coincidencias fueron evaluadas mediante búsqueda de dominios conservados en la base de datos CDD de NCBI, con un *e-value* de 0.05.
Los análisis comparativos de las secuencias se realizaron utilizando los *softwares* VirusDetect y Virfind (Xu, et al., 2017), utilizando el genoma de la papa para sustraer las secuencias provenientes de la planta y facilitar la identificación de los virus presentes en las muestras. Las figuras 7 y 8 muestran el flujo de trabajo de cada *software*.

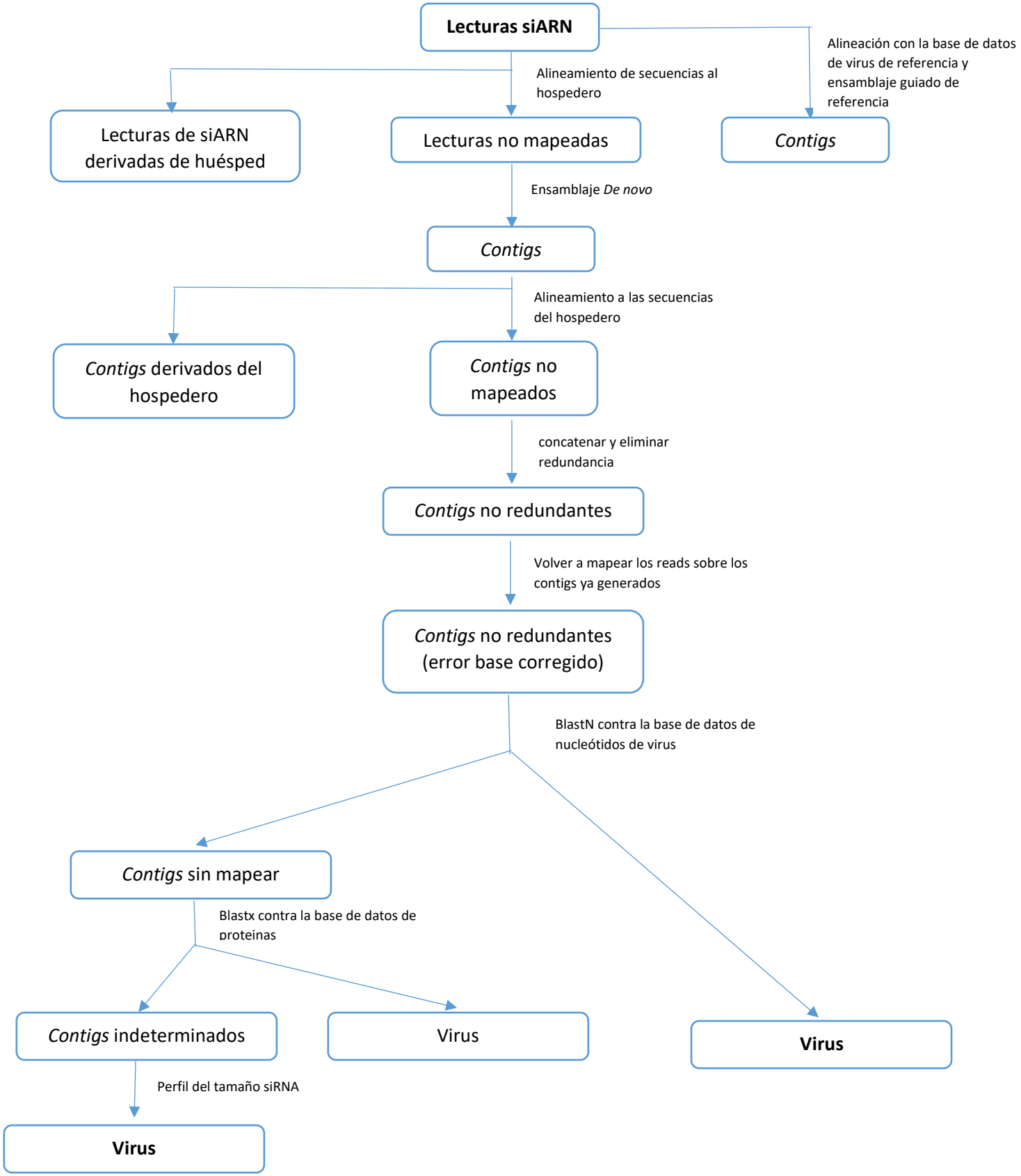


Figura 7: Diagrama de flujo del *software* VirusDetect para la identificación de virus utilizando datos de secuenciación de siARN.

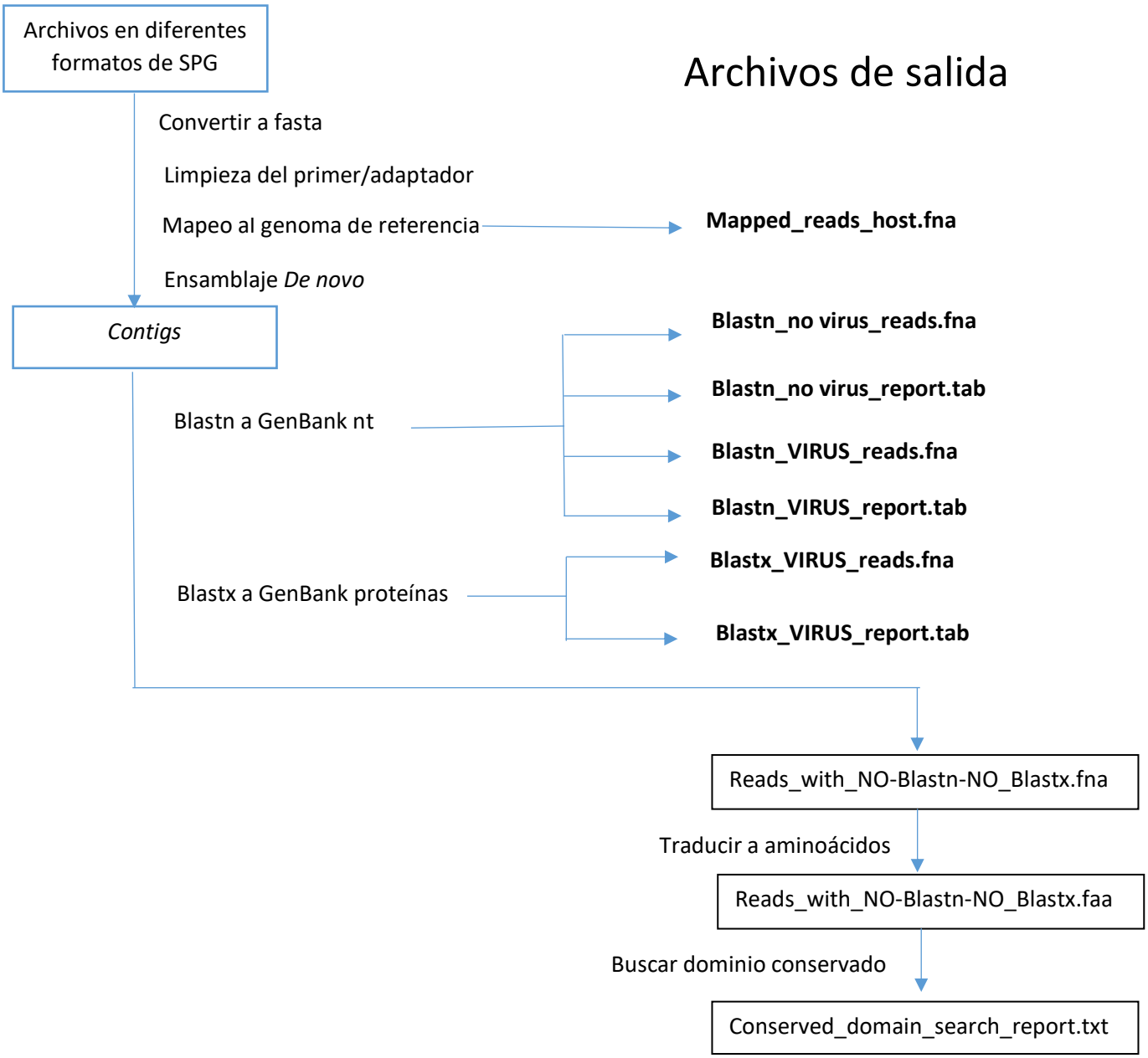


Figura 8: Diagrama de flujo de VirFind para la detección y descubrimiento de virus utilizando datos de secuenciación masiva (HTS).

5.6 Identificación de nuevos virus

Los resultados de VirusDetect fueron tabulados en un Excel (Anexo 5) donde se seleccionaron los posibles nuevos virus, basados en su porcentaje de identidad y cobertura con respecto al genoma de referencia.

De acuerdo con los criterios de demarcación de especies establecidos para los géneros Badnavirus y Torradovirus (King et al., 2012; Sanfaçon et al., 2009), se identificaron aquellos contigs que presentaron baja identidad respecto a secuencias de referencia. En el caso de Badnavirus, se evaluó la región correspondiente al ORF3 que codifica la poliproteína, incluyendo los dominios RT-RNaseH, aplicando como criterio una identidad nucleotídica inferior al 80 % respecto a especies reconocidas. Para Torradovirus (familia Secoviridae), la demarcación se basó en la identidad aminoacídica en regiones conservadas como la proteína de la cápside (CP) y la región Pro-Pol que incluye la ARN polimerasa dependiente de ARN (RdRp), considerando como umbral una identidad inferior al 80 % en dichas regiones para la delimitación de una posible nueva especie.

La ubicación y orientación de los contigs se determinó mediante alineamientos contra genomas de referencia obtenidos del NCBI. Para el género *Badnavirus* se utilizó como referencia el acceso *Canna yellow mottle-associated virus* (KX066020.1), y para *Torradovirus* los accesos Squash chlorotic leaf spot virus (KU052530.1) y *Squash chlorotic leaf spot virus* (KU052531.1).

El ensamblaje del pseudogenoma se realizó considerando identidad nucleotídica, orden, orientación y un solapamiento mínimo entre contigs. Los reads originales fueron remapeados sobre el ensamblaje para verificar continuidad y cobertura. Las regiones faltantes se completaron mediante diseño de primers específicos, amplificación por PCR y confirmación por secuenciación Sanger.

5.7. Diseño de *primers*

Se realizó un alineamiento múltiple de los pseudogenomas obtenidos de cada género, ubicando las secuencias consenso, donde se diseñaron los

primers utilizando el programa Oligoanalyzer (https://www.idtdna.com/pages/tools/oligoanalyzer).

5.7.1. Diseño de *primers* para el género *Badnavirus*

Según el alineamiento múltiple, se diseñaron *primers* para completar las secuencias faltantes en el pseudogenoma (Tabla 8).

Tabla 8: *Primers* diseñados para completar el genoma del nuevo *Badnavirus.*

Nombre	Secuencia	Inicio-Termino (*)	Amplicón (nt)
BADNA4_F	GACGTTGAGGGTTTAGTTCTGAAAATCAG	943-971	726
BADNA4_R	CATAAGGACTTCAAGGGTGTGCT	1674-1696	
BADNA5_F	ACAAGGCTGGGCTACGATG	2378-2396	2266
BADNA5_R	GACCAATTTGTAGAAGGTGATGCTC	4615-4639	
BADNA6_F	GAAGCATGTGACGCCCAAAATG	4903-4924	507
BADNA6_R	CCAGATGTTTCCTGTGCTGCT	5410-5430	
BADNA7_F	CACATCCAAGACGAGAAGGAAC	6335-6356	1649
BADNA7_R	CGATCAGTGATTCAGTCTGTCTTTC	991-1015	

(*) Posición de los *primers* según el genoma de referencia del CaYMAV-Ci (KX0660201)

5.7.2. *Diseño de primers* para el género *Torradovirus*

Las especies virales en el género *Torradovirus* posee un genoma bipartito lineal, por lo cual se procedió a utilizar dos adaptadores BanD one y BandD Two que se van a unir a los extremos 5’ y 3’, para poder obtener información de esos extremos. Ambos adaptadores se utilizaron para el ARN1 Y ARN2. Adicionalmente se diseñaron *primers* para completar las secuencias faltantes en los pseudogenomas de ambos ARNs (Tabla 9).

Tabla 9: *Primers* diseñados para completar el genoma del nuevo *Torradovirus.*

	Nombre	Secuencia	Región	Amplicon (nt)
RNA1	TORRA1_F	TGCCGCTAAATGTAACAATC	990	3077
	TORRA1_R	CCATTCCATCATAAAACAGCTC	3673	
	TORRA1.1_F	CCYAAGGTRTCAATMTGTG	394	3673
	TORRA1_R	CCATTCCATCATAAAACAGCTC	3673	
	TORRA2_F	GAGCCCTCAATTTTAACTTGTAG	3817	1182
	TORRA2_R	GGGTYGTTGTAGTATTGACGT	4999	
	TORRA2_F	GAGCCCTCAATTTTAACTTGTAG	3817	1761
	TORRA2.1_R	CTGAGAACACCTGAGAATCTG	5578	
RNA2	TORRA3_F	TAAGGTTTTCTGCGTTGTTG	251	2160
	TORRA3_R	CGCTGAGGTTGCATTATATTG	2411	
	TORRA4_F	TGGTGACCTTATCCTGATC	2085	1579
	TORRA4_R	CAGACAAAGTAGAGATTCTGC	3664	
	TORRA5_F	TGGACTCATAGTGTTACTCAAC	263	2120
	TORRA5_R	GCACMGTSACGTAWATCTG	2383	
	TORRA6_F	CTGTRTACGTGYTSAATCAG	2303	601
	TORRA6_R	TCCAAAACTGAGGTGAATATG	2904	
	TORRA6_F	CTGTRTACGTGYTSAATCAG	2303	1089
	TORRA6.1_R	GTTCTTCGGTGGATGATTAC	1089	

(*) Posición de los *primers* según el genoma de referencia del *Squash chlorotic leaf spot virus* ARN1 (KU052530.1) y ARN2 (KU052531.)

5.8. PCR

Para la detección de *Badnavirus*, el ADN de las muestras Czo-004 y Czo-040 fue extraído mediante el protocolo CTAB modificado descrito por Doyle y Doyle (1990). Para *Torradovirus*, se utilizaron ARN largos previamente obtenidos durante la preparación de librerías, los cuales fueron sometidos a retrotranscripción para la síntesis de ADNc.

Las reacciones de PCR se realizaron en un volumen final de 20 µL, empleando el Master Mix detallada en la Tabla 31: 10 µL de 2X Phusion™ Flash High-Fidelity PCR Master Mix (Thermo Scientific) (1X final), 0.5 µL de cada cebador (10 µM) para una concentración final de 0.25 µM, 2 µL de ADN o ADNc molde y 7 µL de agua libre de nucleasas (NFW).

Las condiciones generales de amplificación fueron: desnaturalización inicial a 98 °C por 10 s; seguida de 35 ciclos de 98 °C por 1 s, alineamiento a temperatura

específica según el par de cebadores (52–62 °C) durante 20–45 s, y extensión a 72 °C por 15 s; con una extensión final a 72 °C por 1 min y mantenimiento a 4 °C. Las temperaturas específicas de alineamiento para cada par de primers (Badna-5, Badna-7, Torra-1, Torra-2, Torra-2.1, Torra-3 y Torra-4) se presentan en la Tabla 32.

En cada ensayo se incluyó un control negativo (agua libre de nucleasas) y un control de planta sana para descartar amplificaciones inespecíficas o contaminación. Los productos de PCR fueron analizados mediante electroforesis en gel de agarosa al 1% en buffer TBE 1X y visualizados bajo luz UV. Los amplicones obtenidos fueron purificados utilizando el kit Roche High Pure PCR Product Purification Kit, siguiendo las instrucciones del fabricante, hasta alcanzar una concentración aproximada de 40 ng/µL

5.9. Clonación de productos de PCR y secuenciación

Los productos de PCR fueron clonados en el vector pGEM-T Easy Vector System I (Promega®), siguiendo las recomendaciones del fabricante. Las ligaciones fueron transformadas en células competentes de *Escherichia coli* DH5α mediante choque térmico. Los transformantes fueron sembrados en medio LB suplementado con ampicilina (100 µg/mL), IPTG y X-gal para selección azul/blanco. Las colonias blancas fueron seleccionadas como candidatas positivas.

Para cada amplicón se analizaron al menos tres clones independientes. La purificación del plásmido recombinante se realizó con el E.Z.N.A Plasmid Kit (Omega Bio-Tek), siguiendo las instrucciones del proveedor (https://omegabiotek.com/product/plasmid-dna-purification-e-z-n-a-plasmid-dna-mini-kit/?cn-reloaded=1)

La verificación del inserto del producto de PCR al vector se realizó digiriendo el plásmido recombinante con la enzima *EcoRI* (Promega®) (https://worldwide.promega.com/products/cloning-and-dna-markers/restriction-enzymes/ecori/?catNum=R6011#resources) y corriendo el producto digerido en gel de agarosa al 1%.

Los plásmidos recombinantes fueron enviados a secuenciar a la empresa Macrogen (Korea) mediante el método Sanger. Para aquellos insertos muy largos se tuvo que diseñar *primers walking* (Tabla 10).

Tabla 10: *Primers Walking* diseñados para completar la secuenciación Sanger

Primers	Sequencia
Badna-5-1F	GTAATGCCCCAAGACTTCGAC
Badna-5-1R	CAACCCCCTTATGATCTTCTGTTG
TORRA1_2F	AGAGCAATAGGATCTACTCGC
TORRA1_2R	GGTAGAAAACTTGATTAGCGGG
TORRA2_2F	AGGTACTTCTTGGATTTCTCTGC
TORRA2_2R	CAATGGCTGGGTAGTTGTAGT

5.10. Análisis bioinformáticos

La calidad de las secuencias de ADN obtenidas mediante la secuenciación Sanger fueron analizadas utilizando el programa Chromas, que nos permitió calcular el porcentaje de fidelidad de cada base nitrogenada que permitió así eliminar aquellas secuencias que presentaban una baja calidad.

La obtención del genoma completo de los virus, se realizó ensamblando las secuencias obtenidas mediante HTS y Sanger utilizando el programa Seqman 5.00 y Geneious 11.1.5.

5.11. Árbol filogenético

Los árboles filogenéticos fueron construidos utilizando las regiones genómicas que constituyen el criterio de demarcación de especies para cada género viral. Para *Badnavirus* se emplearon secuencias nucleotídicas correspondientes a la región que codifica la retrotranscriptasa (RT) y la RNasa H, mientras que para *Torradovirus* se utilizó la región correspondiente a la ARN polimerasa dependiente de ARN (RdRp).

Las secuencias obtenidas en este estudio y las secuencias de referencia descargadas del NCBI fueron alineadas utilizando el algoritmo ClustalW implementado en MEGA (versión 11). El análisis filogenético se realizó

mediante el método Neighbor-Joining utilizando el modelo de sustitución p-distance y 1000 réplicas de bootstrap. (Tamura et al., 2021).

6. RESULTADOS

6.1 Análisis de secuencias

Las 144 muestras fueron analizadas utilizando el software VirusDetect v.1.6, utilizando la base de datos vrl_Plants_217_U95 identificándose la presencia de secuencias correspondientes a virus comunes y virus nuevos (Anexo 5). Entre los virus comunes se identificaron secuencias de ARN que pertenecen a distintos géneros entre ellos tenemos a *Potexvirus* (77 muestras), *Potyvirus* (140 muestras), *Carlavirus* (13 muestras), *Ilarvirus* (4 muestras), Pomovirus (1 muestra), *Polerovirus* (1 muestra) y *Nepovirus* (21 muestras) (Tabla 11). Asimismo, se detectaron secuencias del género *Ipomovirus* en seis muestras. Sin embargo, este género está reportado principalmente en cultivos de yuca y no se encuentra comúnmente asociado al cultivo de papa. Por lo cual, su presencia podría estar relacionada con factores epidemiológicos en campo, como la proximidad de cultivos alternativos o el movimiento de personas, insectos y materiales entre parcelas, más que con una infección establecida en el cultivo de papa.

Tabla 11: Virus comunes identificados mediante el análisis con el software Virusdetect

N° Campo	Provincia	Distrito	Localidad	Variedades	Virus comunes	Cantidad de muestras
1	Anta	Limatambo	Huerta	Yungay	Carlavirus Potexvirus Potyvirus	1 4 11
2	Anta	Limatambo	Cartonpata	Cica	Nepovirus Pomovirus Potexvirus Potyvirus	1 1 4 11
3	Anta	Limatambo	Ayaviri (Pampacona)	Ccompis	Carlavirus Nepovirus Potexvirus Ipomovirus Potyvirus	1 2 6 1 11
4	Anta	Ancahuasui	San Martin	Cica	Ipomovirus Nepovirus Potexvirus Potyvirus	5 3 5 11

5	Paruro	Huanoquite	Tiwicte	Cica	Nepovirus Potexvirus Potyvirus	2 5 11
6	Paruro	Huanoquite	Tiwicte 2	Cica	Potexvirus Potyvirus	4 11
7	Paruro	Yaurisque	Taray	Cica	Potexvirus Potyvirus Polerovirus	6 11 1
8	Cuzco	Santiago	Occopata-Cacchona	Ccompis, Huayro, Mactillo	Carlavirus Nepovirus Potexvirus Potyvirus	2 2 7 11
9	Paucartambo	Challabamba	Chuza	Mactillo, Amarillis	Carlavirus Nepovirus Potexvirus Potyvirus	6 2 7 11
10	Paucartambo	Challabamba	Pilco	Canchan	Carlavirus Ilarvirus Nepovirus Potexvirus Potyvirus	1 3 3 5 11
11	Paucartambo	Challabamba	Sunchubamba	Cica	Carlavirus Potexvirus Potyvirus	1 8 11
12	Paucartambo	Colquepata	Paucana	Cica, Alianza	Nepovirus Potexvirus Potyvirus	3 11 12
13	Calca	Pisac	Sacaca	Yungay	Carlavirus Ilarvirus Nepovirus Potexvirus Potyvirus	1 4 4 7

Además, se detectaron secuencias con un porcentaje de identidad menor al <80% con respecto al genoma de referencia durante el análisis con Blastx. Estas hallazgos sugieren la presencia de secuencias virales divergentes, ya que presentan baja identidad con virus previamente descritos., entre los cuales se destacan: Allexivirus (1 muestra), *Badnavirus* (2 muestras), Botrexvirus (1 muestra), Citrivirus (2 muestras), Comovirus (6 muestras), Cytorhabdovirus (2 muestras), Enamovirus (6 muestras), Foveavirus (1 muestra), Ilarvirus (4 muestras), Nepovirus (2 muestras), Potyvirus, Prunevirus (2 muestras), *Torradovirus* (11 muestras), Trichovirus (1 muestra) y Varicosavirus (5 muestras) (Tabla 12).

Tabla 12: Virus nuevos identificados mediante el análisis con el software VirusDetect

N°Campo	Provincia	Distrito	Localidad	Variedades de papa	Virus nuevos	Cantidad de muestras
1	Anta	Limatambo	Huerta	Yungay	Badnavirus	1
					Comovirus	1
					Enamovirus	1
					Torradovirus	5
2	Anta	Limatambo	Cartonpata	Cica	Comovirus	1
					Enamovirus	2
					Polerovirus	1
4	Anta	Ancahuasui	San Martin	Cica	Badnavirus	1
					Comovirus	1
					Nepovirus	1
7	Paruro	Yaurisque	Taray	Cica	Enamovirus	1
8	Cuzco	Santiago	Occopata-Cacchona	Ccompis, Huayro, Mactillo	Allexivirus	1
					Botrexvirus	1
					Citrivirus	2
					Cytorhabdovirus	1
					Foveavirus	1
					Prunevirus	2
					Trichovirus	1
					Varicosavirus	1
9	Paucartambo	Challabamba	Chuza		Cytorhabdovirus	1
					Nepovirus	1
					Varicosavirus	1
10	Paucartambo	Challabamba	Pilco	Canchan	Comovirus	3
					Varicosavirus	3
11	Paucartambo	Challabamba	Sunchubamba	Cica	Enamovirus	1
12	Paucartambo	Colquepata	Paucana	Cica, Alianza	Enamovirus	1
					Torradovirus	7

En la Tabla 13 se presenta el porcentaje de cobertura e identidad obtenido al comparar los contigs virales con las secuencias de referencia. La selección de *Badnavirus* y *Torradovirus* para el ensamblaje y completado genómico se basó en la cobertura obtenida, la continuidad y coherencia de los contigs alineados, la presencia de regiones conservadas empleadas como criterio de demarcación taxonómica y la posibilidad de completar las secuencias mediante PCR y secuenciación.

Tabla 13: Porcentaje de cobertura obtenido de las secuencias siARN al genoma del virus referencial, que son considerados como nuevos virus

Virus nuevos	% de cobertura	% de identidad
Allexivirus	2	34.18
Badnavirus	21	41.96
Botrexvirus	1.8	50.00
Citrivirus	6.2	39.81
Comovirus	ARN1-12/ARN2-27	42.57
Cytorhabdovirus	1.2	27.67
Enamovirus	13	23.39
Foveavirus	4.7	65.79
Ilarvirus	12.1	33.82
Nepovirus	ARN1-21.5/ARN2-7.2	72.41
Polerovirus	5.7	51.22
Prunevirus	8.5	30.13
Torradovirus	ARN1-20.7/ARN2-14.4	31.45
Trichovirus	3.9	69.64
Varicosavirus	20.3	29.76

6.2. Análisis comparativo entre VirusDetect y VirFind

Para el análisis comparativo entre ambas metodologías, se seleccionaron muestras representativas clasificadas según el número de virus detectados mediante VirusDetect. La clasificación se realizó considerando el número total de virus identificados por muestra, definiéndose categorías de baja, media y alta en función de los siguientes criterios cuantitativos: baja presencia (1 virus detectado), presencia intermedia (3 virus) y alta presencia (>4 virus). Estas muestras fueron posteriormente analizadas con VirFind para la comparación de resultados entre ambas herramientas.

En la Tabla 14 se muestra los resultados de los análisis llevados a cabo con ambos softwares de la muestra Czo-061. Ambas metodologías identificaron la presencia de Potato virus Y (PVY) mediante Blastn, con una cobertura del 87.3% obtenido con el software VirFind y del 92.8% con VirusDetect. Además, con el software VirFind se obtuvo 24 *contigs* con un tamaño que varía entre 200 y 2516 nt, mientras que con VirusDetect se obtuvo 23 *contigs* con un rango de tamaño que oscila entre 58 y 2832 nt. En ambos softwares se identificaron varios genomas de referencia correspondientes a PVY.

Tabla 14: Identificación de virus en la muestra Czo-061 mediante los *softwares* VirFind y VirusDetect.

Muestras	Virus detectados	Cantidad de *contigs*	%Cobertura	Longitud de *contigs* (nt)
Czo-061 (Virfind)	PVY	24	87.3	200-2516
Czo-061 (VirusDetect)	PVY	23	92.8	58-2832

La Tabla 15 presenta los resultados obtenidos con ambos *softwares* para la muestra Czo-088. El *software* Virfind identificó la presencia de 3 virus (PVY, PVS, PVX), mientras que VirusDetect identificó 5 virus (PVY, PVS, PVX, PVA y *papaya ringspot virus*) mediante el análisis con Blastn.

Con respecto a la formación de *contigs*, se observó que los análisis realizados con VirusDetect produjo una mayor cantidad de *contigs* en comparación con VirFind, además, el tamaño de estos también fue superior en el análisis realizado con VirusDetect. Así mismo también se observó una mayor cobertura de los virus identificados con VirusDetect en comparación con VirFind.

Tabla 15: Identificación de virus en la muestra Czo-088 mediante los softwares VirFind y VirusDetect.

Muestras	Virus detectados	Cantidad de *contigs*	%Cobertura	Longitud de *contigs* (nt)
Czo-088(Virfind)	PVY	2	2.6	221-224
	PVS	3	3.5	204-825
	PVX	26	99.8	189-1359
Czo-088 (VirusDetect)	PVY	68	97	41-1328
	PVS	58	27.1	41-167
	PVX	16	100	41-6466
	PVA	70	48.5	41-167

En cuanto al análisis comparativo de la muestra Czo-117, Virfind identificó 2 virus (PVY y *Torradovirus*), mientras que VirusDetect detectó 3 virus (PVY, PVX y *Torradovirus*). La cobertura obtenida con el análisis de VirFind para el PVY y el *Torradovirus* fue del 7.5% y del 34.1%, respectivamente, mientras que, con VirusDetect la cobertura de estos virus fue del 57.4% y del 65%, respectivamente (ver Tabla 16). Asimismo, se observó una mayor cantidad de *contigs* con VirusDetect en comparación con VirFind.

El PVX fue identificado con el *software* VirusDetect, obteniendo una cobertura de 95.5% y un total de 48 *contigs,* los cuales presentan un tamaño que varía entre 41-1042 nt.

Tabla 16: Identificación de virus en la muestra Czo-117 mediante los softwares VirFind y VirusDetect**.**

Muestras	**Virus detectados**	**Cantidad de *contigs***	**%Cobertura**	**Longitud de *contigs* (nt)**
Czo-117(Virfind)	PVY	4	7.5	176-290
	Torradovirus	10	34.1(ARN1)-31.1(ARN2)	217(ARN1)-939(ARN2)
Czo-117 (VirusDetect)	PVX	48	95.5	41-1042
	PVY	98	99.8	42-4217
	Torradovirus	22	57.4(ARN1)-65(ARN2)	114(ARN1)-962(ARN2)

En la Tabla 17 muestra los resultados obtenidos con ambos *softwares* de la muestra Czo-129, el *software* VirFind no identificó ningún virus, mientras que el *software* VirusDetect identificó la presencia de secuencias pertenecientes al Potato virus X (PVX) con una cobertura de 90.6%, una cantidad de *contigs* de 35 cuyos tamaños tienen un rango entre 42 – 833 nt.

Tabla 17: Identificación de virus en la muestra Czo-129 mediante los softwares VirFind y VirusDetect**.**

Muestras	**Virus detectados**	**Cantidad de *contigs***	**%Cobertura**	**Longitud de *contigs* (nt)**
Czo-129 (Virfind)	0	0	0	0
Czo-129 (VirusDetect)	PVX	35	90.6	42-833

6.3. Obtención del pseudogenoma de los virus nuevos de los géneros *Badnavirus* y *Torradovirus*

6.3.1. Género *Badnavirus*

Los virus nuevos pertenecientes al género *Badnavirus* fueron identificados en dos muestras (Czo-004 y Czo-040), los datos de la colecta se pueden observar en el Anexo 1.

Las Figuras 9 y 10 corresponde a la planta Czo-004 y Czo-040 respectivamente, en la cual se puede observar la presencia de un

moteado clorótico y hojas arrugadas, estos síntomas pueden o no ser generados por los nuevos *Badnavirus*, ya que en esas plantas se encontraron otro virus como el *PVY, Nepovirus y Comovirus.*



Figura 9: Planta correspondiente a la muestra Czo-004, la cual presenta un moteado clorótico leve en las hojas.



Figura 10. Planta correspondiente a la muestra Czo-040, la cual presenta un moteado clorótico leve y enrollamiento en las hojas.

En las Tablas 18 y 19 se puede observar los resultados obtenidos con VirusDetect correspondiente al análisis Blastx, de las muestras Czo-004 y Czo-040 respectivamente, en ambas tablas se observa la identificación de un virus del género *Badnavirus* cuyos *contigs* formados mostraron un porcentaje de identidad mínimo de 41.8% y un máximo de 97% con el genoma de referencia (Figura 11 y 12). Para la reconstrucción del pseudogenoma se alinearon todos los contigs asociados a este virus con el genoma de referencia y se colocó la letra “N” en los espacios vacíos.

Tabla 18: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-004.

Reference	Length	Coverage (%)	#contig	Depth	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
ACF60613.1	1916	292 (15.2)	5	244.0	51.3	80.14	97.56	68.57	badnavirus	Cycad leaf necrosis virus, complete genome.
ADN88180.1	175	164 (93.7)	2	474.8	99.7	50.61	76.74	41.32	badnavirus	Pineapple bacilliform comosus virus, complete genome.
ANN87818.1	1684	1442 (85.6)	9	350.8	73.7	71.89	94.00	56.36	badnavirus	Canna yellow mottle virus isolate 0028, complete genome.
ANN87816.1	133	133 (100)	1	286.1	60.1	42.54	42.54	42.54	badnavirus	Canna yellow mottle virus isolate 0028, complete genome.
ACS44708.1	1792	1038 (57.9)	8	301.9	63.4	72.26	96.00	41.90	badnavirus	Sugarcane bacilliform virus isolate BataviaD, complete genome.
AAA42421.1	997	780 (78.2)	16	71.0	14.9	74.91	88.89	55.17	comovirus	Andean potato mottle virus polyprotein encoding small and large coat proteins and potential movement protein mRNA, complete cds.
ACX53644.1	1857	551 (29.7)	8	103.3	21.7	40.07	57.50	25.66	comovirus	Turnip ringspot virus isolate M12 segment RNA1, complete sequence.
ADI60054.1	1863	524 (28.1)	7	93.8	19.7	39.74	53.93	27.78	comovirus	Broad bean true mosaic virus isolate EV-11 segment RNA1, complete sequence.
AAA42422.1	729	245 (33.6)	5	48.2	10.1	94.00	95.83	84.21	comovirus	Andean potato mottle virus RNA polymerase mRNA, 3' end.
AIG59705.1	186	146 (78.5)	5	30.7	6.4	71.71	83.33	61.76	comovirus	Andean potato mottle virus strain PV-0057 coat protein gene, partial cds.

Tabla 19: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-040.

Reference	Length	Coverage (%)	#contig	Depth	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
ANN87817.1	173	96 (55.5)	2	247.6	77.9	54.90	61.67	45.24	badnavirus	Canna yellow mottle virus isolate 0028, complete genome.
AEC49888.1	176	78 (44.3)	2	270.5	85.1	62.03	68.00	59.26	badnavirus	Banana streak IM virus, complete genome.
ACS44708.1	1792	1518 (84.7)	13	336.8	106.0	69.96	97.96	40.00	badnavirus	Sugarcane bacilliform virus isolate BataviaD, complete genome.
ANN87818.1	1684	1466 (87.1)	15	309.7	97.5	71.97	95.92	41.18	badnavirus	Canna yellow mottle virus isolate 0028, complete genome.
ANN87816.1	133	111 (83.5)	1	178.4	56.2	41.96	41.96	41.96	badnavirus	Canna yellow mottle virus isolate 0028, complete genome.
AAA42421.1	997	954 (95.7)	2	182.0	57.3	72.16	72.62	67.44	comovirus	Andean potato mottle virus polyprotein encoding small and large coat proteins and potential movement protein mRNA, complete cds.
AKL83120.1	1856	1201 (64.7)	3	212.9	67.0	39.93	45.16	34.45	comovirus	Bean rugose mosaic virus isolate Parana segment RNA1 polyprotein gene, complete cds.
CBL43396.1	1499	782 (52.2)	10	86.9	27.4	45.44	78.57	31.30	nepovirus	Grapevine Bulgarian latent virus segment 2, complete sequence, isolate Serb1.
AAD09329.2	1626	456 (28)	7	64.5	20.3	57.61	82.14	38.21	nepovirus	Blackcurrant reversion virus RNA2 polyprotein mRNA, complete cds.
CBL43395.1	2095	1606 (76.7)	17	89.6	28.2	68.67	100	44.19	nepovirus	Grapevine Bulgarian latent virus segment 1, complete sequence, isolate Serb1.

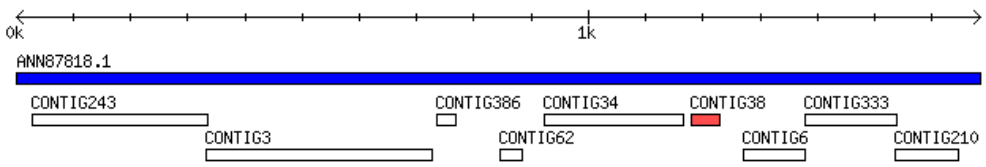


Figura 11: Representación gráfica de los *contigs* obtenidos mediante el análisis con el software VirusDetect de la muestra Czo-004. El rectángulo azul representa el genoma de referencia (CaYMAV-Ci). Los rectángulos blancos representan los *contigs* con baja identidad respecto al genoma de referencia, mientras que los de color rojo representan aquellos con alta identidad.

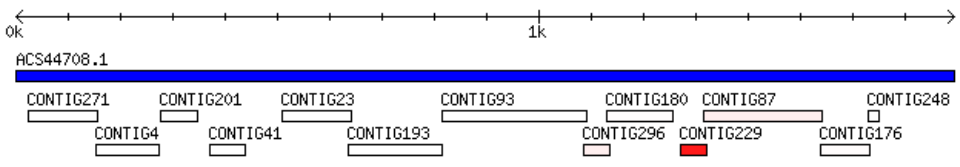


Figura 12: Representación gráfica de los *contigs* obtenidos mediante el análisis con el software VirusDetect de la muestra Czo-040. El rectángulo azul representa el genoma de referencia (CaYMAV-Ci). Los rectángulos blancos representan los *contigs* con baja identidad respecto al genoma de referencia, mientras que los de color rojo representan aquellos con alta identidad

El alineamiento se realizó utilizando el programa ClustalW, donde se alinearon cada uno de los *contigs* con un genoma de referencia representativo del género (KX066020.1 *Canna yellow mottle-associated virus* isolate CaYMAV-Ci, complete genome) que se encuentra disponible en el NCBI, posteriormente se juntaron todos los alineamientos utilizando el programa MegAlign, en la cual se pudo observar la ubicación de todos los *contigs* en el genoma de referencia. Seguidamente se utilizó un editor de secuencia (Editseq) que ayudó a obtener el pseudogenoma en forma lineal.

Se utilizó el programa Geneious para poder realizar las gráficas como se observa en la Figura 13 en la cual se observa el pseudogenoma obtenido de ambas muestras. Para poder completar el genoma del virus nuevo se procedió a diseñar *primers* (Tabla 8) en las regiones conservadas de ambas muestras.



Figura 13: Posición de los *primers* diseñados en los pseudogenomas obtenidos de las muestras Czo-004 y Czo-040, después de alinear los *contigs* al genoma de referencia CaYMAV-Ci. Las regiones blancas representan las secuencias faltantes del genoma, mientras que las regiones de color negro representan las secuencias obtenidas mediante el análisis con el software VirusDetect.

6.3.2. Género *Torradovirus*

Los virus nuevos pertenecientes al género *Torradovirus* fueron identificados en once muestras (Czo-001, Czo-002, Czo-007, Czo-008, Czo-010, Czo-112, Czo-113, Czo-115, Czo-117, Czo-119 y Czo-120), los datos de la colecta se pueden observar en el Anexo 1.

En la Figura 14 se muestra la planta Czo-001, en la cual se puede observar la presencia de síntomas como moteado clorótico y hojas arrugadas.



Figura 14: Planta correspondiente a la muestra Czo-001 con clorosis y leve enrollamiento en las hojas.

En las Tablas 20 y 21 se puede observar los resultados obtenidos del VirusDetect correspondiente al análisis Blastx, de las muestras Czo-001 y Czo-117, en las tablas se observa la identificación de un virus del género *Torradovirus* cuyos *contigs* formados obtuvieron un porcentaje de identidad mínimo de 49.3% y un máximo de 62.5% con el genoma de referencia (Figura 15 y 16). Para la reconstrucción del pseudogenoma se alinearon todos los contigs asociados a este virus con el genoma de referencia y se colocó la letra "N" en los espacios vacíos.

Tabla 20: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-001.

Reference	Length	Coverage (%)	#contig	Depth	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
ACT79984.1	1197	318 (26.6)	5	79.1	15.5	44.38	55.56	40.43	NA	Tomato chocolate spot virus segment RNA2, complete sequence.
ABV44417.1	190	40 (21.1)	1	123.7	24.3	52.50	52.50	52.50	torradovirus	Tomato marchitez virus isolate PRI-TMarV0601 segment RNA2 hypothetical protein and polyprotein genes, complete cds.
AMN91910.1	2207	386 (17.5)	9	41.1	8.1	58.40	85.71	42.55	torradovirus	Squash chlorotic leaf spot virus isolate Su12-10 segment RNA1, complete sequence.
AHA85557.1	1167	274 (23.5)	5	75.2	14.8	46.91	76.47	41.03	torradovirus	Carrot torradovirus 1 isolate CTV-1_RNA2_H6 ORF1 and polyprotein genes, complete cds.
AHA91818.1	973	213 (21.9)	6	55.5	10.9	49.30	58.82	41.86	torradovirus	Cassava Torrado-like virus isolate CM5460-10 segment RNA2 polyprotein gene, partial cds.

Tabla 21: Resultados del análisis obtenido mediante VirusDetect utilizando la herramienta Blastx de la muestra Czo-117.

Reference	Length	Coverage (%)	#contig	Depth	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
AMN91912.1	1054	685 (65)	8	142.7	34.4	41.07	62.50	30.00	torradovirus	Squash chlorotic leaf spot virus isolate Su12-10 segment RNA2, complete sequence.
ACB47565.1	1198	192 (16)	3	150.7	36.4	50.00	56.14	42.86	torradovirus	Tomato torrado virus isolate Wal'03 segment RNA2, complete sequence.
ACA49852.1	1074	176 (16.4)	3	120.6	29.1	49.43	56.14	42.86	torradovirus	Tomato torrado virus isolate ToTV-CE segment RNA2 polyprotein gene, partial cds.
AMM63581.1	190	157 (82.6)	1	193.2	46.6	31.45	31.45	31.45	torradovirus	Tomato marchitez virus isolate M clone pJL89-M-R2 segment RNA2, complete sequence.
AMN91910.1	2207	1266 (57.4)	12	86.1	20.8	52.52	87.18	30.00	torradovirus	Squash chlorotic leaf spot virus isolate Su12-10 segment RNA1, complete sequence.

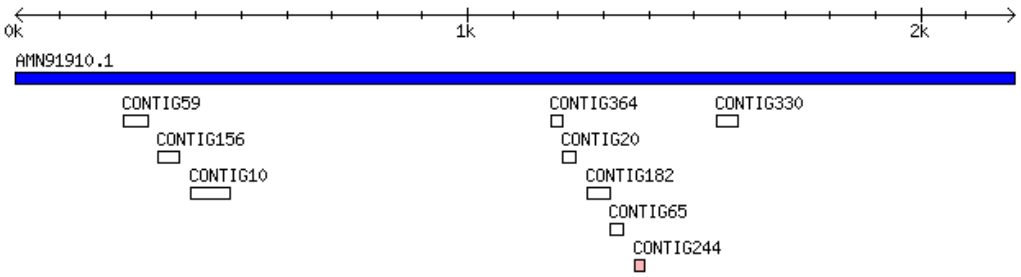


Figura 15. Representación gráfica de los *contigs* obtenidos mediante el análisis con el software VirusDetect de la muestra Czo-001. El rectángulo azul representa el genoma de referencia *Squash chlorotic leaf spot virus* isolate Su12-10 segment RNA1, los rectángulos blancos representan los *contigs* formados que presentan baja identidad con el genoma de referencia y el color rosado representa los *contigs* que presentan una alta identidad con el genoma de referencia.

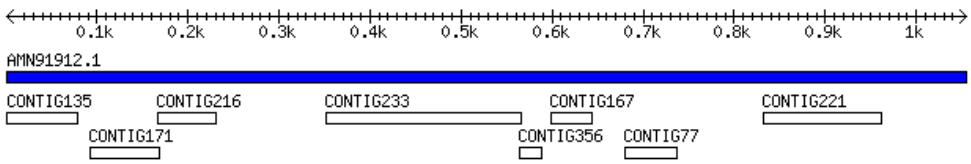


Figura 16. Representación gráfica de los *contigs* obtenidos mediante el análisis con el software VirusDetect de la muestra Czo-001. El rectángulo azul representa el genoma de referencia *Squash chlorotic leaf spot virus* isolate Su12-10 segment RNA2, los rectángulos blancos representan los *contigs* formados que presentan baja identidad con el genoma de referencia.

El alineamiento se realizó utilizando el programa ClustalW, donde se alinearon cada uno de los *contigs* con un genoma de referencia representativo del género (KU052530.1 *Squash chlorotic leaf spot virus*

isolate Su12-10 segment RNA1, complete sequence y KU052531.1 *Squash chlorotic leaf spot virus* isolate Su12-10 segment RNA2, complete sequence) que se encuentra disponible en el NCBI. Posteriormente, para la obtención del pseudogenoma del nuevo *Torradovirus* se utilizó la misma metodología descrita para *Badnavirus*.

Con el programa Geneious se realizó las gráficas como se observa en la Figura 17 en la cual se observa el pseudogenoma obtenido de la muestra Czo-117. Para poder completar el genoma del virus nuevo se procedió a diseñar *primers* (Tabla 9) en las regiones conservadas.

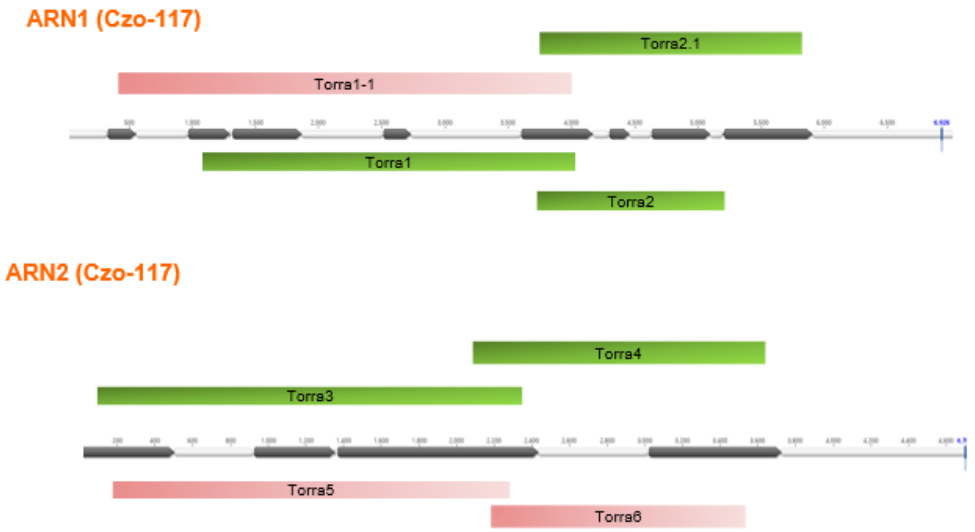


Figura 17: Posición de los *primers* diseñados en el pseudogenoma obtenido de la muestra Czo-117, después de alinear los *contigs* al genoma de referencia *Squash chlorotic leaf spot virus* isolate Su12-10. Las regiones blancas representan las secuencias faltantes del genoma, mientras que las regiones de color negro representan las secuencias obtenidas mediante el análisis con el software VirusDetect.

6.4. Resultados de PCR

6.4.1. Género *Badnavirus*

Se realizaron diversos protocolos de PCR (Anexo 3) utilizando los *primers* diseñados (Tabla 8).

En la Figura 18 se muestra los resultados obtenidos de la realización del PCR utilizando los *primers* Badna-4 y Badna-6 de las muestras Czo-004y Czo-040, siendo el producto de ellos 726 pb y 507 pb respectivamente.

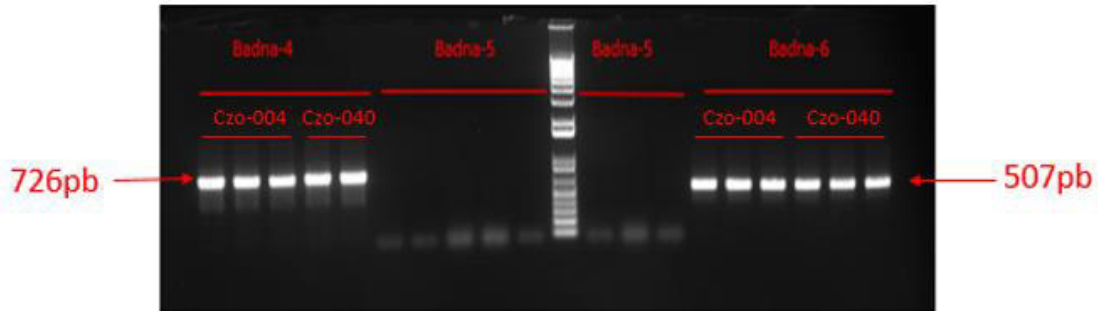


Figura 18: Resultados de la amplificación por PCR de las muestras Czo-004 y Czo-040 utilizando los *primers* Badna-4 y Badna-6

En la Figura 19 se muestra los resultados del PCR obtenida utilizando el *primers* Badna-5, cuyo producto de amplificación es de 2266 pb aprox.

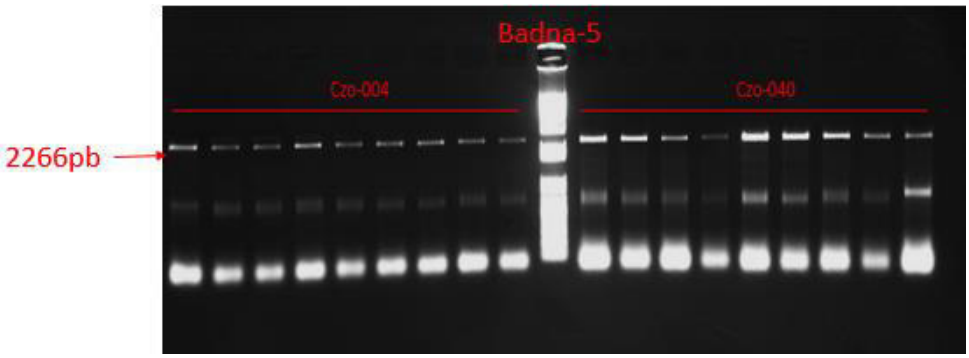


Figura 19: Resultados de la amplificación por PCR de las muestras Czo-004 y Czo-040 utilizando el par de *primers* Badna-5

En la Figura 20 se muestra los resultados de la gradiente de temperatura realizada para identificar la temperatura idónea para la amplificación del *Primers* Badna-7, observándose la presencia de una banda del tamaño esperado (1400 pb aprox) en todas las temperaturas.

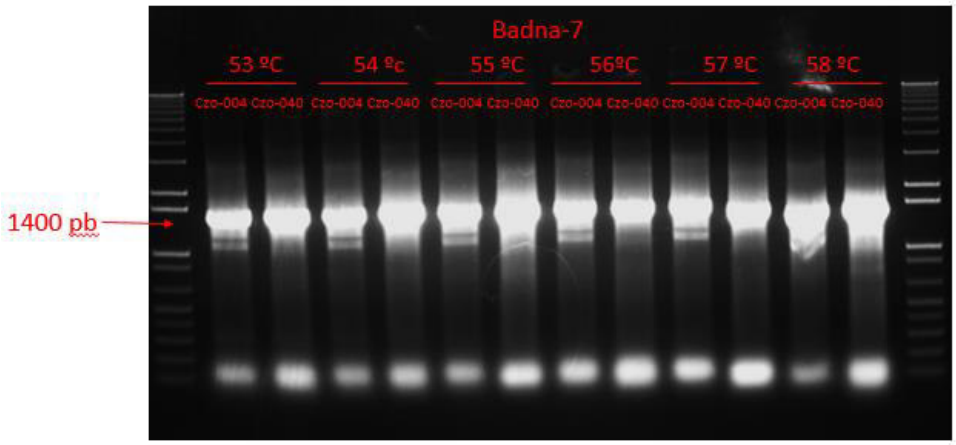


Figura 20: Determinación de la temperatura óptima de amplificación por PCR utilizando el par de *primers* Badna-7.

Todas las bandas esperadas fueron clonadas y enviadas a secuenciar mediante la metodología Sanger. Los resultados de la digestión con la enzima *EcoRI* se observan en las Figuras 21 al 24.

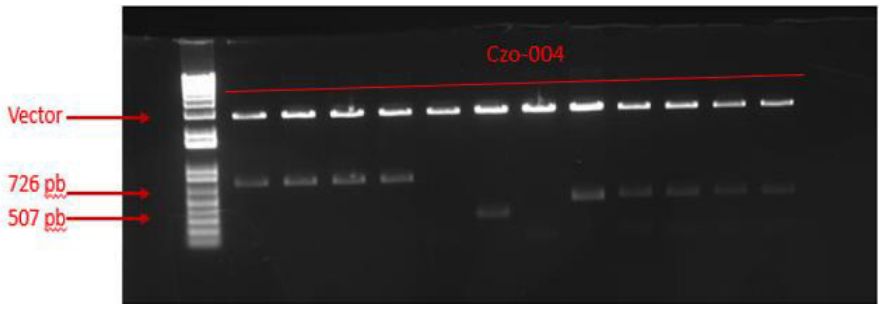


Figura 21: Resultados de la digestión del vector pGEMT-easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos de los *Primers* Badna-4 (726 pb) y Badna-6 (507 pb) pertenecientes a la muestra Czo-004**.**

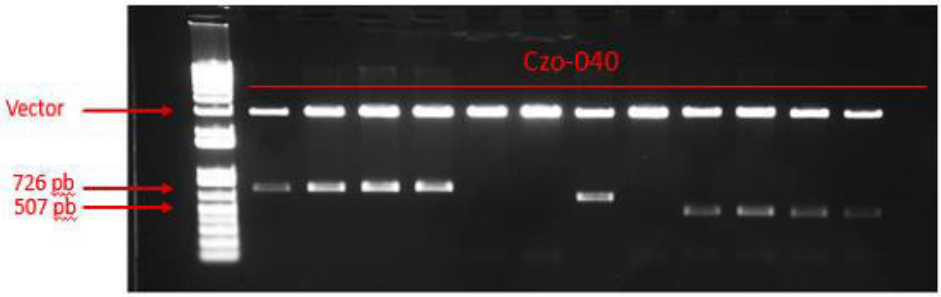


Figura 22: Resultados de la digestión del vector pGEM-T - easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos de los *Primers* Badna-4 (726 pb) y Badna-6 (507 pb) pertenecientes a la muestra Czo-040.

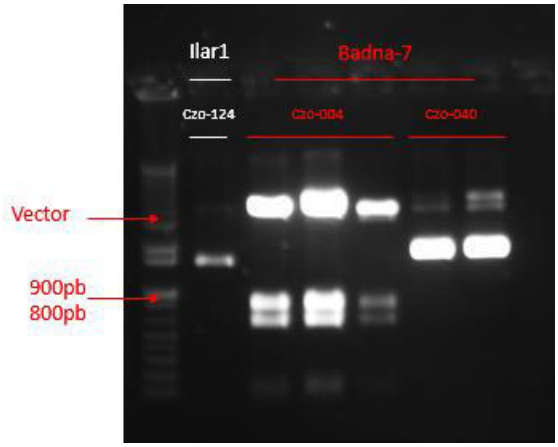


Figura 23: Resultados de la digestión del vector pGEM-T-easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos utilizando el *Primers* Badna-5.

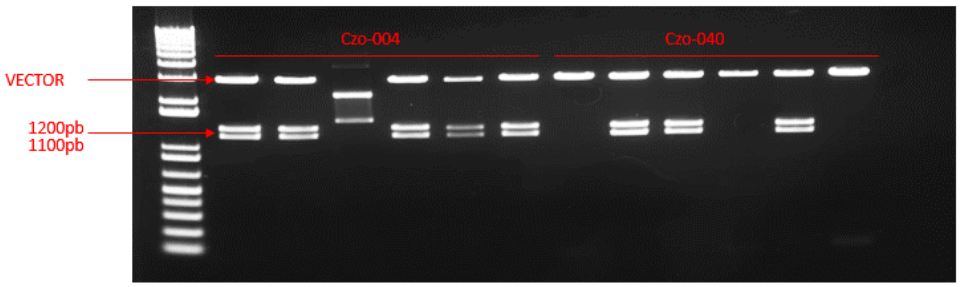


Figura 24: Resultados de la digestión del vector pGEM-T-easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos utilizando el par de *primers* Badna-7.

6.4.2. Género *Torradovirus*

En la Figura 25 se observa la amplificación de las muestras utilizando los *primers* Torra 1 (3077pb) y Torra 2.1 (1761 pb); en la Figura 26 lo correspondiente a los *primers* Torra 3 (2160 pb); y en la Figura 27 a los correspondientes a Torra4 (1579 pb).

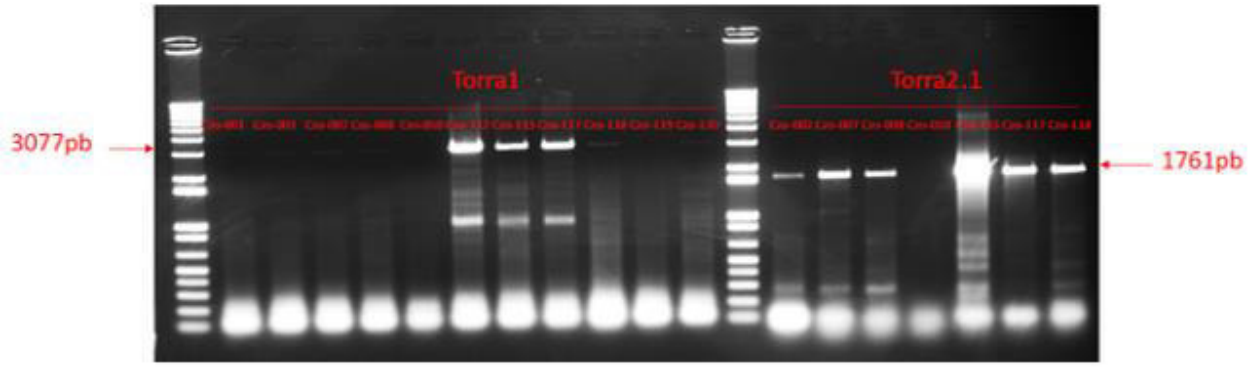


Figura 25. Gel de agarosa que contiene los productos de PCRs utilizando los *primers* Torra- 1 (3077 pb) y Torra2.1 (1761 pb).

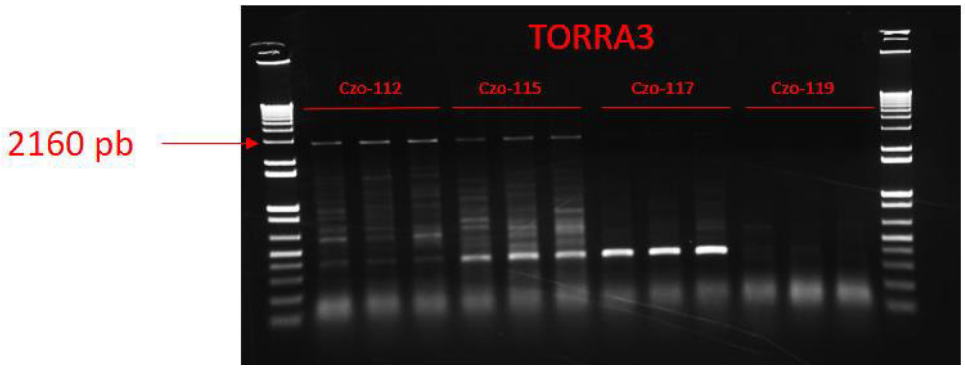


Figura 26: Gel de agarosa que contiene el producto de PCR de los *primers* Torra- 3 (2160 pb).

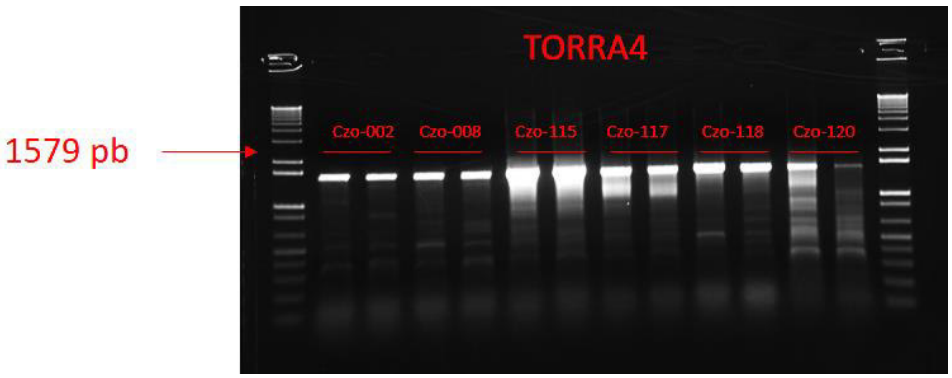


Figura 27: Gel de agarosa que contiene el producto de PCR de los *primers* Torra- 4 (1579 pb).

Todas las bandas esperadas fueron clonadas y enviadas a secuenciar mediante la metodología Sanger. Los resultados de la digestión con la enzima *EcoRI* se observan en las Figuras 28 al 30.

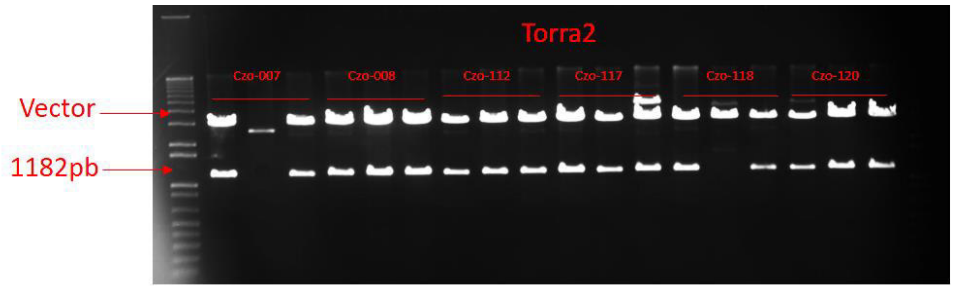


Figura 28. Resultados de la digestión del vector pGEM-T-easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos utilizando el *primers* Torra2 (1182pb).

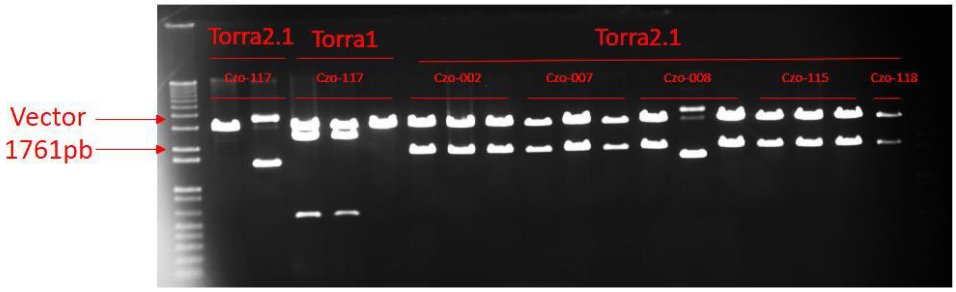


Figura 29: Resultados de la digestión del vector pGEM-T-easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos utilizando el *primers* Torra2.1 (1182pb) y Torra1.

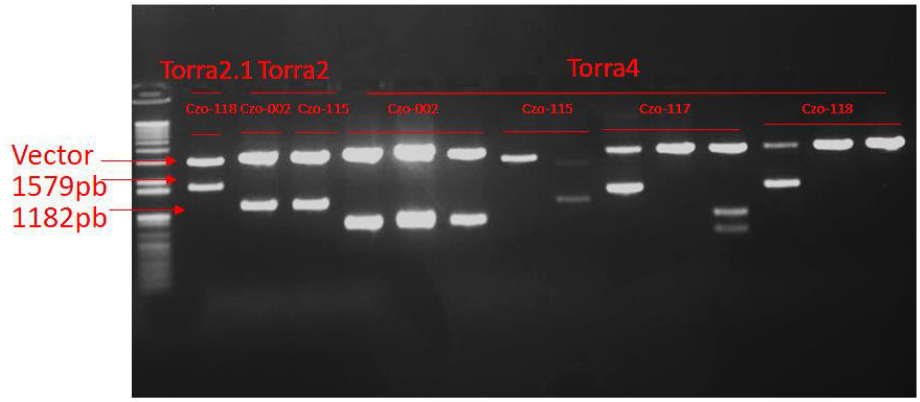


Figura 30: Resultados de la digestión del vector pGEM-T-easy utilizando la enzima *EcoRI*, se puede observar los productos de PCR obtenidos utilizando de los *primers* Torra2.1 (1761 pb), Torra2 (1182 pb) y Torra 4 (1579 pb).

Los productos de PCR de los *primers* Torra-3 fueron enviados a secuenciar como productos de PCR.

6.5. Obtención del genoma completo

El análisis de las secuencias mediante la metodología Sanger fue analizado utilizando el programa Chromas. Este software nos permitió identificar bases nitrogenadas de baja calidad al observar varios indicadores: picos mal definidos, valores de calidad menores de 40 y la presencia de artefactos. Así, pudimos eliminar las secuencias que presentaron una baja calidad, garantizando la fiabilidad de los datos obtenidos.

6.5.1. Género *Badnavirus*

El ensamblaje de los *contigs* y de las secuencias obtenidas mediante la secuenciación Sanger de las muestras Czo-004 y Czo-040 se realizó utilizando el programa SeqMan (Figuras 31 y 32).

Con respecto al *primer* badna-5 tiene un producto de PCR de 2266pb aproximadamente, se tuvo que realizar un diseño de *primers walking* para obtener información completa de esta región (tabla 10)

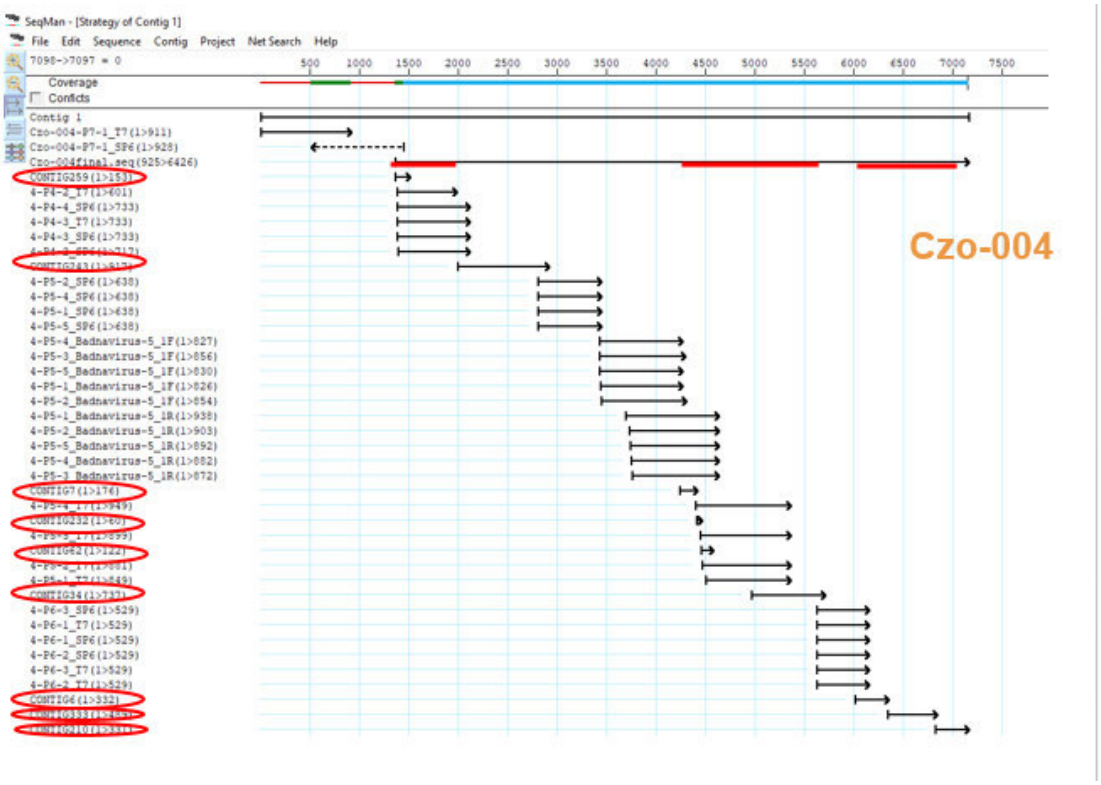


Figura 31: Ensamblaje de *contigs* con el programa Seqman. Se muestra la unión de las secuencias de los *contigs* formados con el VirusDetect (círculos rojos) y las secuencias obtenidas mediante la secuenciación Sanger de la muestra Czo-004. Las regiones de color rojo son las secuencias faltantes las cuales fueron completadas con los PCRs realizados.

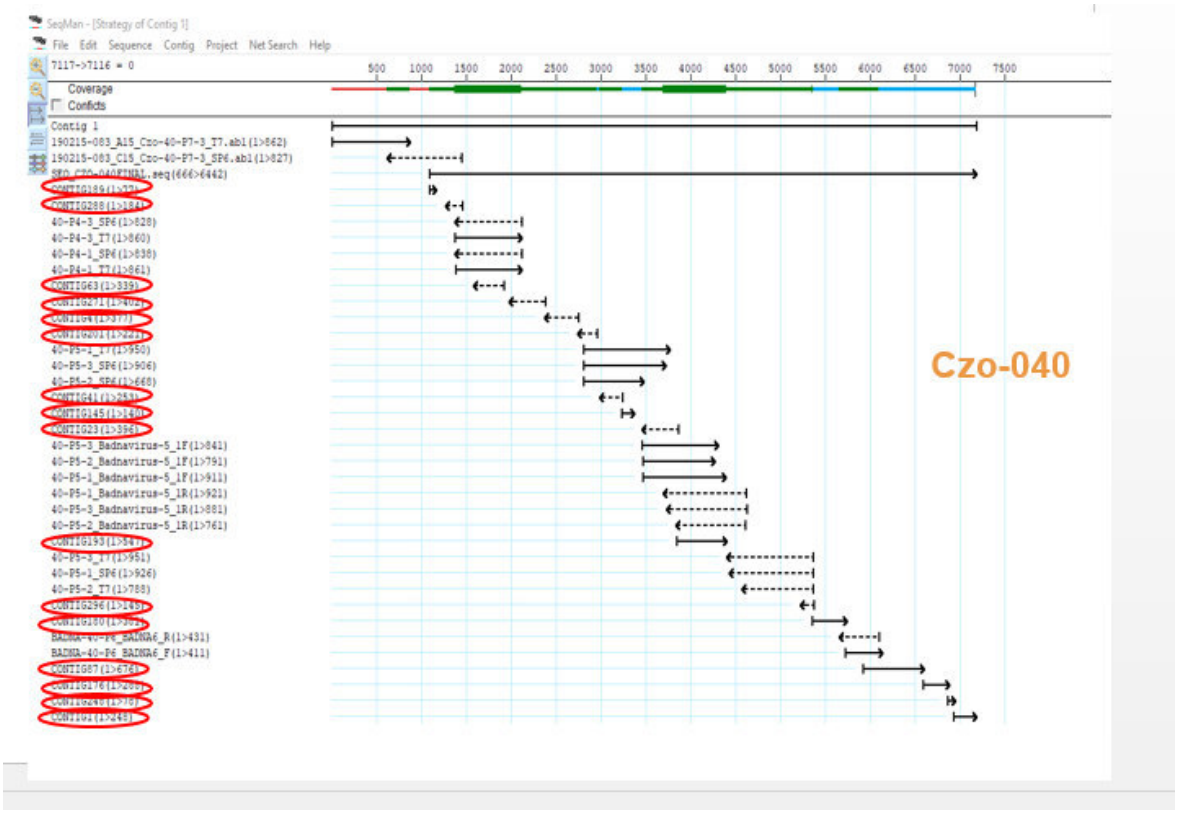


Figura 32: Ensamblaje de *contigs* con el programa Seqman. Se muestra la unión de las secuencias de los *contigs* formados con el VirusDetect (círculos rojos) y las secuencias obtenidas mediante la secuenciación Sanger de la muestra Czo-040. Las regiones de color rojo son las secuencias faltantes las cuales fueron completadas con las PCRs realizadas.

De esta forma se consiguió obtener el genoma completo del nuevo *Badnavirus*, se realizó un alineamiento con la secuencia completa de ambas muestras obteniendo una similaridad de 99.1% (Tabla 22). En la Figura 33 se observar la organización genómica del *Badnavirus* nuevo obtenido de la muestra Czo-004 y Czo-040. El genoma de este nuevo *Badnavirus* tiene un tamaño de 7096 y 7094 pb, respectivamente. Presenta 3 ORFs, el ORF 1 tiene un tamaño de 527 pb y el ORF2 tiene un tamaño de 392 pb, ambos codifican para proteínas de función desconocida; mientras que, el ORF3 posee un tamaño de 5426 pb y codifica una poliproteína de dedos de zinc que se ubica en la región 3482-3530pb. El gen que codifica una enzima Aspartato proteasa se encuentra en la región 4322-4649 pb, la enzima RT-LTR se ubica en la región 4988-5564pb y la enzima RNasa-H se encuentra en la región 5861-6242 pb. Asimismo, se identificó una caja TATA en la posición 6909–6916 pb y la tRNA unión de metionina en la posición 6-19pb.

Tabla 22: Porcentaje de identidad del genoma del *Badnavirus* obtenido de las muestras Czo-040 y Czo-004.

	Czo-004	**Czo-040**
Czo-004		99.1%
Czo-040	99.1%	

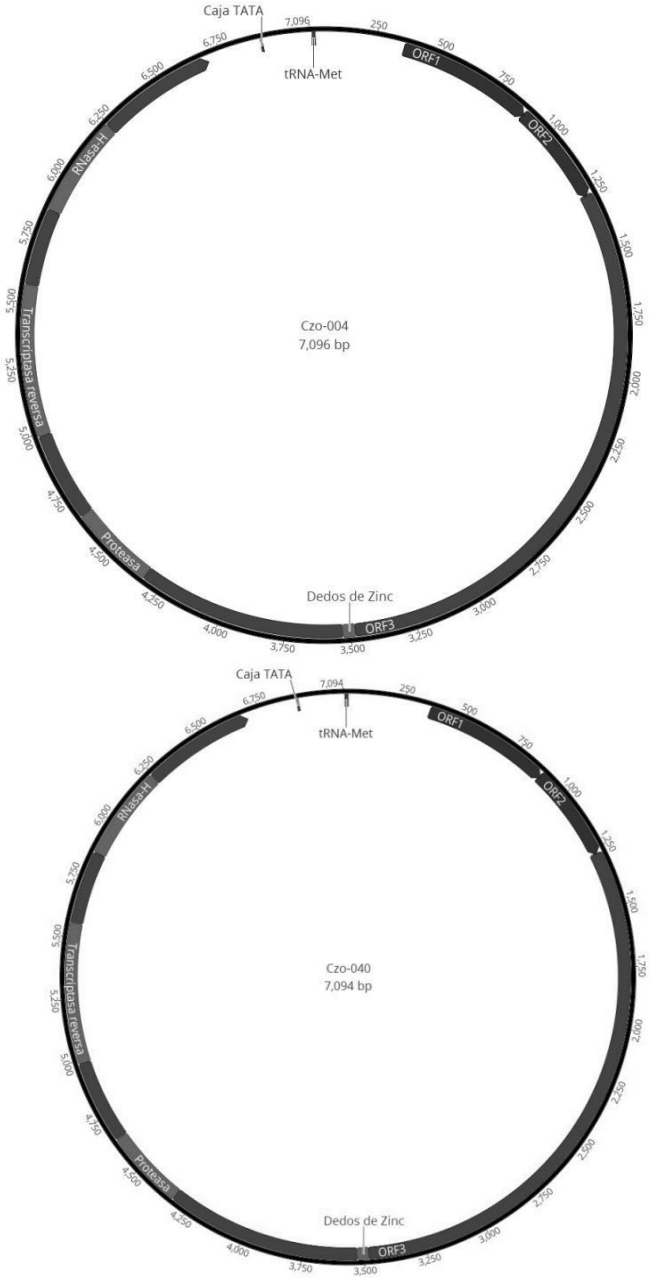


Figura 33: Organización del genoma completo del nuevo *Badnavirus* obtenido de las muestras Czo-004 y Czo-040.

6.5.2. Género *Torradovirus*

Como se observa en la Tabla 20 se identificó la presencia de secuencias del nuevo *Torradovirus* en 11 muestras, por lo cual se procedió a realizar la amplificación de los *primers* Torrado 2 y Torrado 2.1 los cuales amplifican la región que codifica para la enzima ARN polimerasa dependiente de ARN (RdRp), la cual es utilizada para realizar estudios filogenéticos.

En los resultados del PCR, se lograron obtener productos de PCR, para 8 muestras, excepto en las correspondientes a Czo-112, Czo-113 y Czo-119, lo cual probablemente se deba al bajo título del virus presente en esas muestras. Tras obtener los resultados de la secuenciación Sanger, de los primers Torra2 que amplifican la región de la RdRp se procedió a realizar un alineamiento múltiple, obteniendo como resultado que las muestras presentan una identidad entre ellas de 98.1% y el 100% (Tabla 23), lo cual nos indica la existencia de una sola especie viral entre las muestras. Por consiguiente, se seleccionaron dos muestras representativas para completar sus genomas, siendo estas la Czo-117 y la Czo-120, ya que fueron las únicas que amplificaron para todos los *primers*.

Tabla 23: Porcentaje de identidad del alineamiento de nucleótidos correspondiente a las secuencias amplificadas con el *primer* Torra 2.

	Czo-117	**Czo-115**	**Czo-008**	**Czo-010**	**Czo-001**	**Czo-007**	**Czo-120**	**Czo-002**
Czo-117		100%	100%	100%	100%	99.67%	98.74%	99.37%
Czo-115	100%		100%	100%	100%	99.67%	98.74%	99.37%
Czo-008	100%	100%		100%	100%	99.67%	98.74%	99.37%
Czo-010	100%	100%	100%		100%	99.67%	98.74%	99.37%
Czo-001	100%	100%	100%	100%		99.67%	98.74%	99.37%
Czo-007	99.67%	99.67%	99.67%	99.67%	99.67%		98.74%	99.37%
Czo-120	98.74%	98.74%	98.74%	98.74%	98.74%	99%		99.37%
Czo-002	99.37%	99.37%	99.37%	99.37%	99.37%	99.04%	98.11%	

El ensamble de los *contigs* y de las secuencias obtenidas mediante la secuenciación Sanger fue realizado con el programa Seqman (Figura 34 y 35).

Con respecto a los primers Torra1 y Torra3, que generan productos de PCR de aproximadamente 3077 y 2160 pb respectivamente, se tuvo que diseñar *primers walking* para obtener información completa de esta región.

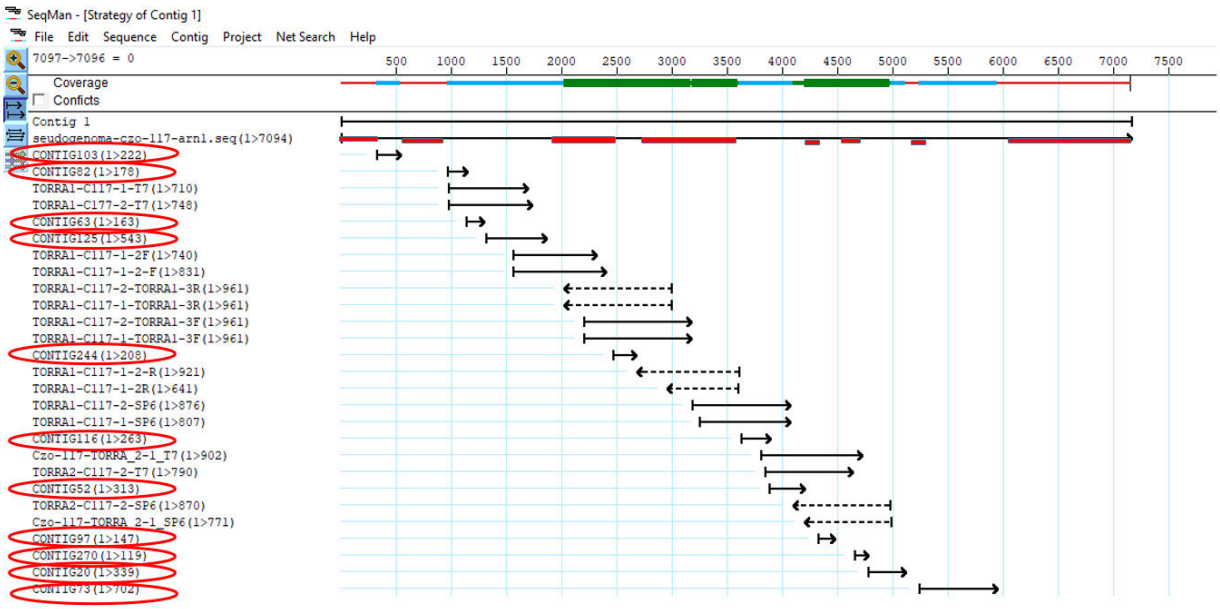


Figura 34: Ensamblaje de los *contigs* con el programa Seqman**.** Se muestra la unión de las secuencias de los *contigs* formados con el VirusDetect (círculos rojos) y las secuencias obtenidas mediante la secuenciación Sanger de la muestra Czo-117 pertenecientes al ARN1 del nuevo *Torradovirus*. Las regiones de color rojo son las secuencias faltantes las cuales fueron completadas con los PCRs realizados.

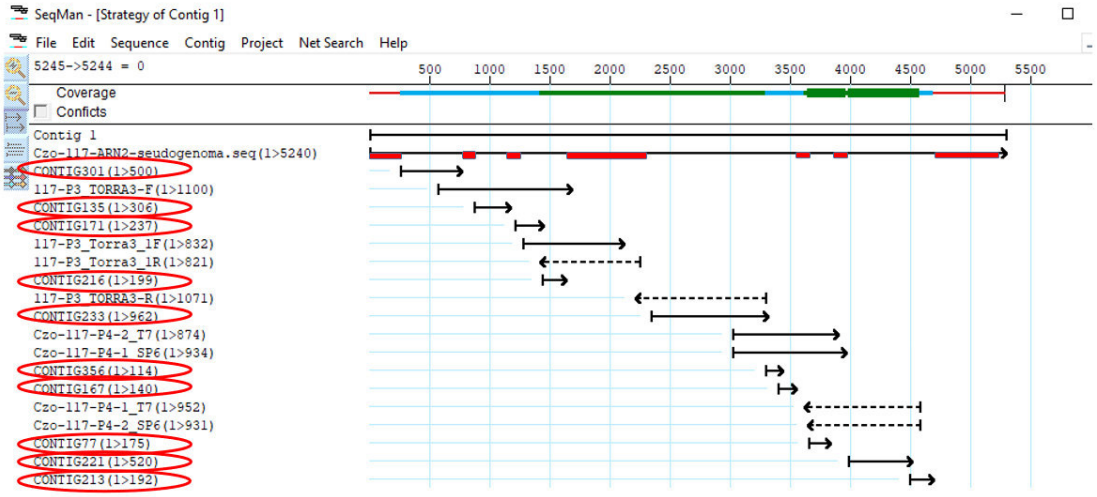


Figura 35: Ensamblaje de los *contigs* con el programa Seqman**.** Se muestra la unión de las secuencias de los *contigs* formados con el VirusDetect (círculos rojos) y las secuencias obtenidas mediante la secuenciación Sanger de la muestra Czo-117 pertenecientes al ARN2 del nuevo *Torradovirus*. Las regiones de color rojo son las secuencias faltantes las cuales fueron completadas con los PCRs realizados.

Para obtener las secuencias de los extremos 5' y 3' se intentó utilizar la metodología de PCR RACE, pero no se obtuvo buenos resultados. En colaboración con el departamento de Agricultura del Estados Unidos (USDA) nos compartió unas secuencias de *Torradovirus* que identificaron en muestras interceptadas provenientes de Perú. Al realizar el alineamiento de las secuencias correspondientes a ARN1 y ARN2 de las muestras Czo-117 y Czo-120 obtenidas en este estudio con aquellas compartidas por el USDA, se obtuvo un porcentaje de identidad de 98,38 % y 98,349 %, respectivamente, para el ARN1 (Tabla 24), y de 94,465 % y 94,56 % para el ARN2 (Tabla 25). Por lo cual se procedió a realizar un mapeo de las secuencias de los siARN (*reads*) utilizando el *software* Geneious, teniendo como referencia la secuencia obtenida de USDA. con lo cual se pudo obtener el genoma completo del virus.

Tabla 24: Porcentaje de identidad del alineamiento de nucleótidos del ARN 1 entre las secuencias compartidas por USDA y las secuencias obtenidas de las muestras procedentes de Cusco.

	Czo-117-ARN1	**Czo-120-ARN1**	**PQ-13-ARN1**
Czo-117-ARN1		98.44%	98.38%
Czo-120-ARN1	98.44%		98.34%
PQ-13-ARN1	98.38%	98.34%	

Tabla 25: Porcentaje de identidad del alineamiento de nucleótidos del ARN 2 entre las secuencias compartidas por USDA y las secuencias obtenidas de las muestras procedentes de Cusco.

	Czo-117-ARN2	**Czo-120-ARN2**	**PQ-13-ARN2**
Czo-117-ARN2		99.523%	94.465%
Czo-120-ARN2	99.523%		94.56%
PQ-13-ARN2	94.465%	94.56%	

De esta forma se logró obtener el genoma del nuevo *Torradovirus*, el cual posee un genoma bipartito (Figura 36 y 37), el ARN1 tiene un tamaño aproximado de 7100 pb y presenta un ORF en la región de 221-6862 pb la cual codifica para una poliproteína que contiene la secuencia de la proteína Helicasa que se encuentra en la región 1382-1703pb y la enzima RdRp que se ubica en la región 4154-5042 pb. El ARN2 tiene un tamaño de 5240 pb y

presenta dos ORFs, el ORF 1 tiene una longitud de 645pb el cual codifica para proteína de función desconocida, y el ORF2 tiene una longitud de 3825 pb y codifica para poliproteína, en la que se han identificado dos proteínas funcionales: una proteína de movimiento, ubicada entre las posiciones 1001 y 1574 pb, y una proteína de la cápside, aproximadamente entre las posiciones 3392 y 3719 pb.

Todas las secuencias obtenidas en este estudio fueron depositadas en la base de datos NCBI para su acceso público; los códigos de accesión correspondientes se indican en la Tabla 26.

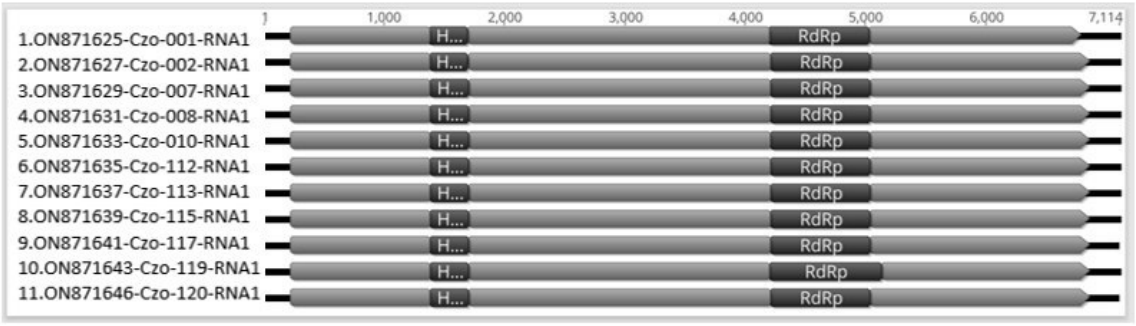


Figura 36: Organización genómica del ARN1 del nuevo *Torradovirus* identificados en Cusco.

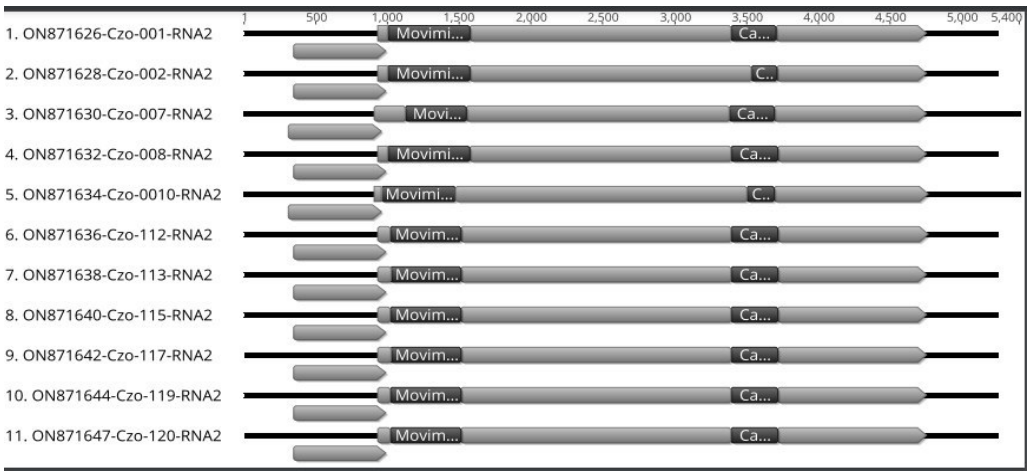


Figura 37: Organización genómica del ARN2 del nuevo *Torradovirus* identificados en Cusco.

Todas las secuencias obtenidas correspondientes al nuevo Torradovirus fueron depositadas en la base de datos NCBI para su acceso público; los códigos de accesión correspondientes se indican en la tabla 26.

Tabla 26: Códigos de accesión de las secuencias de *Torradovirus* depositadas en NCBI

Muestra	Segmento	Código de accesión	Longitud(pb)
Czo-001	ARN1	ON871625	7108
Czo-001	ARN2	ON871626	5247
Czo-002	ARN1	ON871627	7106
Czo-002	ARN2	ON871628	5239
Czo-007	ARN1	ON871629	7110
Czo-007	ARN2	ON871630	5338
Czo-008	ARN1	ON871631	7113
Czo-008	ARN2	ON871632	5247
Czo-010	ARN1	ON871633	7113
Czo-010	ARN2	ON871634	5380
Czo-112	ARN1	ON871635	7112
Czo-112	ARN2	ON871636	5245
Czo-113	ARN1	ON871637	7114
Czo-113	ARN2	ON871638	5247
Czo-115	ARN1	ON871639	7113
Czo-115	ARN2	ON871640	5246
Czo-117	ARN1	ON871641	7097
Czo-117	ARN2	ON871642	5239
Czo-119	ARN1	ON871643	7113
Czo-119	ARN2	ON871644	5246
Czo-120	ARN1	ON871646	7097
Czo-120	ARN2	ON871647	5239

6.6. Árbol filogenético

Se realizaron arboles filogenéticos con los genes que son determinantes para especiación en cada género, para lo cual se utilizó el programa Mega 11 utilizando el método *Neighbor joining* con *boostrap* de 1000.

6.6.1. Género *Badnavirus*

En la figura 38 muestra el árbol filogenético realizado con las secuencias de nucleótidos que codifican para las proteínas Retrotranscriptasa (RT) y RNAasa-H de distintas especies del género *Badnavirus* y *Caulimovirus* los cuales forman parte de la familia *Caulimoviridae*; como se observa las secuencias obtenidas de las muestras Czo-004 y Czo-040 se encuentran en un mismo clado, el virus más cercano evolutivamente es *Canna yellow motle virus* el cual presenta un porcentaje de identidad que varía entre 75,19 % y 75,35 % (Tabla 27).

Tabla 27: Porcentaje de identidad del alineamiento de nucleótidos de la región RT–RNasa H del nuevo *Badnavirus* frente a la secuencia de *Canna yellow mottle virus*.

	Czo-004	**Czo-040**	**MF074075.1**
Czo-004		99.08%	75.35%
Czo-040	99.08%		75.19%
MF074075.1	75.35%	75.19%	

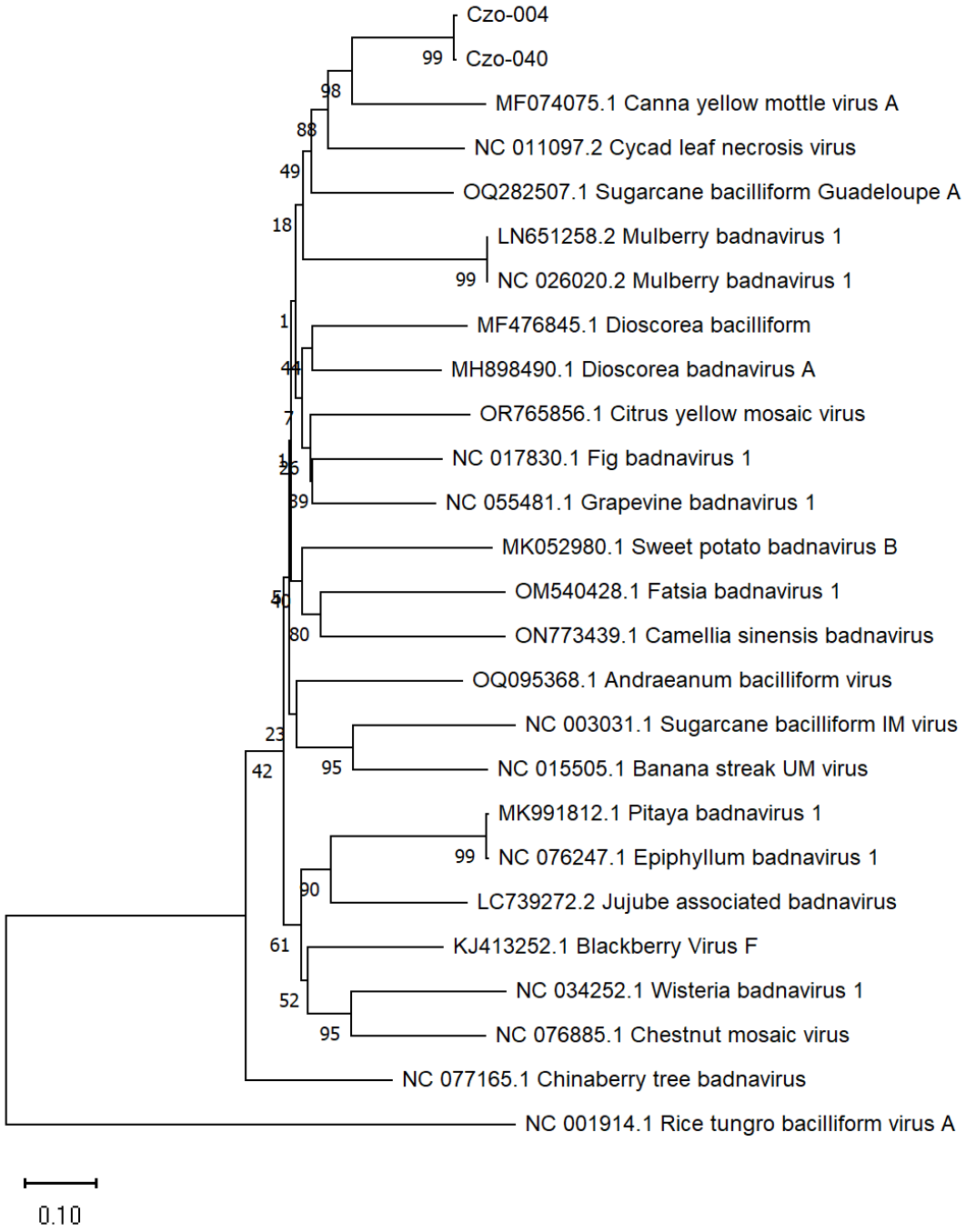


Figura 38: Árbol filogenético del nuevo *Badnavirus*, utilizando la secuencia de nucleótidos de la región que codifican para las proteínas Retrotransciptasa (RT) + RNAasa H.

6.6.2. Género *Torradovirus*

En la Figura 39 muestra el árbol filogenético realizado con las secuencias de aminoácidos que codifican para la proteína ARN polimerasa dependiente de ARN (RdRp) de distintas especies pertenecientes a la familia secoviridae; como se observa las secuencias obtenidas de las muestras de Cusco se encuentran en un mismo clado, el virus más cercano evolutivamente es *Squash chlorotic leaf spot virus* el cual presentan un porcentaje de identidad de 71.52% (Tabla 28), Asimismo, se realizó el alineamiento de las secuencias de aminoácidos del gen que codifica la proteína de la cubierta proteica, el cual indicó que el nuevo *Torradovirus* presenta una identidad del 32 % con *Squash chlorotic leaf spot virus* (Tabla 29).

Tabla 28: Porcentaje de identidad del alineamiento de aminoácidos de la región RdRp entre el nuevo *Torradovirus* y *Squash chlorotic leaf spot virus*.

	Czo-117	**Czo-120**	**YP_009389546**
Czo-117		100%	75.445%
Czo-120	100%		75.445%
YP_009389546	75.445%	75.445%	

Tabla 29: Porcentaje de identidad del alineamiento de aminoácidos de la región de la proteína de la cápside entre el nuevo *Torradovirus* y *Squash chlorotic leaf spot virus*.

	Czo-117	**Czo-120**	**UZN89715.1**
Czo-117		99.69%	32.104%
Czo-120	99.69%		32.025%
UZN89715.1	32.104%	32.025%	

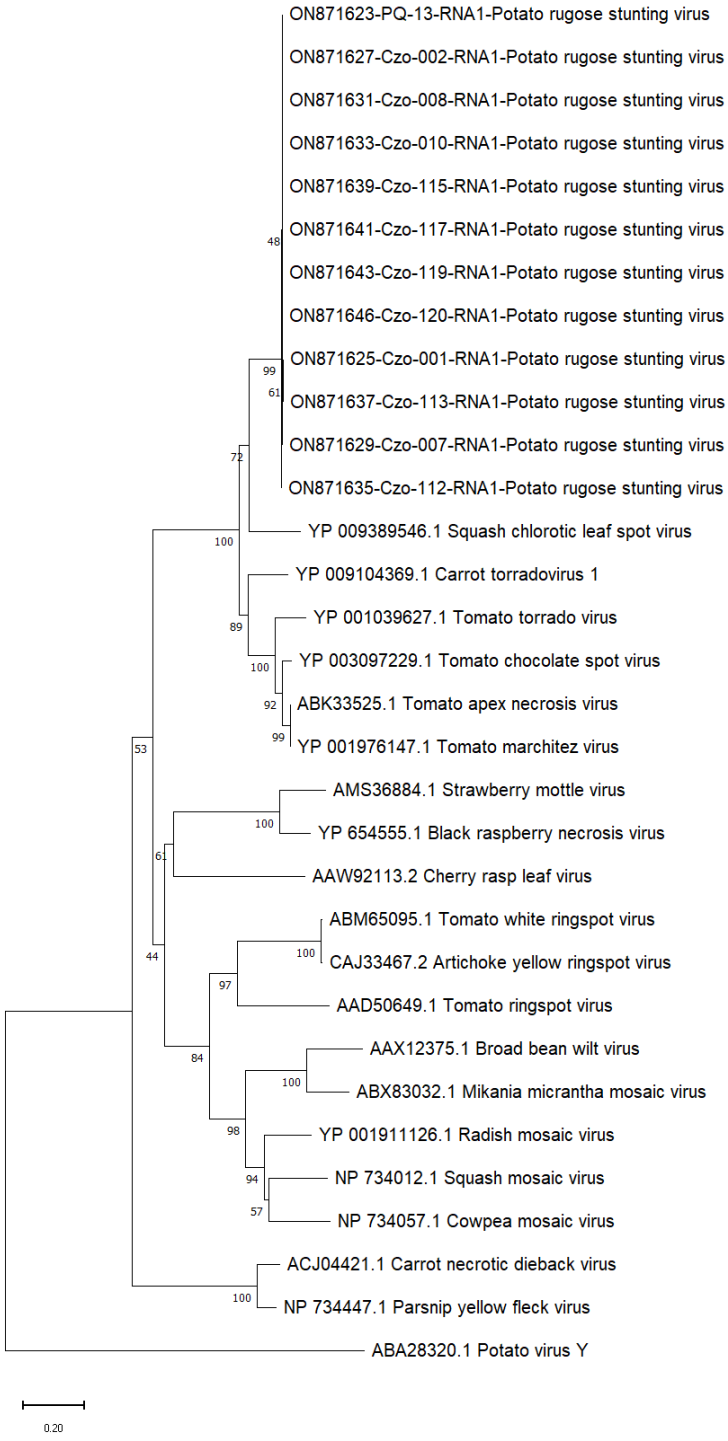


Figura 39: Árbol filogenético del nuevo *Torradovirus*, utilizando la secuencia de aminoácidos de la región que codifica para la proteína ARN polimerasa dependiente de ARN.

7. DISCUSIÓN

En la actualidad, existe un creciente interés en aplicar enfoques basados en ARN pequeños (sRNAs) para la detección de virus, así como para la inspección sanitaria oficial de plantas.

En este trabajo se utilizó esta metodología y se logró identificar la presencia de virus comunes (Tabla 11) y secuencias divergentes candidatas (Tabla 12) en plantas de papa, y se obtuvo el genoma de dos virus candidatos pertenecientes a los géneros *Badnavirus* (con genoma ADN) y *Torradovirus* (con genoma ARN). Estos resultados concuerdan con lo obtenido por Kreuze et al., (2019), quienes utilizaron esta metodología y lograron identificar virus de ARN y ADN presentes en plantas de camote.

En este conjunto de muestras, VirusDetect detectó 9 virus con una cobertura promedio de 68,1 %, mientras que VirFind identificó 6 virus con una cobertura promedio de 39,1 % (Tablas 14–17). Estos resultados indican que, bajo el criterio de detección empleado (análisis Blastn/Blastx y porcentaje de cobertura), VirusDetect mostró un mejor desempeño en la identificación y ensamblaje de secuencias virales en las muestras evaluadas. Estos resultados son similares a los obtenidos por Zheng et al. (2017), quienes compararon el *software* VirusDetect con el *software* Vigilancia y Diagnóstico de Virus y Viroides (VSD) utilizando muestras de melocotón. Los resultados mostraron que VirusDetect identificó 6 virus adicionales en comparación con VSD, los cuales no fueron reportados en ensayos preliminares.

El resultado del análisis de Blastx proporcionó información sobre los nuevos virus presentes en las muestras, esta herramienta ha sido utilizada por otros investigadores los cuales lograron identificar nuevos virus (Zhang et al., 2017, Li et al., 2013, Roy et al., 2013). La información obtenida mediante este análisis permitió reconstruir el pseudogenoma de los virus candidatos a nuevas especies (un *Badnavirus* y un *Torradovirus*), conforme a los criterios de demarcación establecidos para cada género. En el caso de *Badnavirus*, se consideró como umbral una identidad nucleotídica inferior al 80 % en la región ORF3 que incluye los dominios RT-RNaseH; mientras que para *Torradovirus* se aplicó un umbral inferior al 80 % de identidad aminoacídica

en regiones conservadas como la proteína de la cápside y la región Pro-Pol que contiene la RdRp. Los *contigs* generados por VirusDetect fueron suficientes para permitir el diseño de *primers* y así completar los genomas de dichos virus mediante secuenciación Sanger.

El nuevo *Badnavirus* se identificó en dos muestras, Czo-004 y Czo-040, colectadas en las localidades de Huerta y San Martín en Cusco, correspondientes a las variedades de papa Yungay y Cica, respectivamente. La sintomatología observada en las plantas infectadas con el *Badnavirus* fueron un moteado clorótico y hojas enrolladas. Sin embargo, debido a que en estas plantas también se identificaron virus comunes de papa como PVY y PVX, no se puede asociar esta sintomatología con el nuevo virus.

El árbol filogenético realizado con las secuencias de nucleótidos que codifican las proteínas de RT + RNAsa H, evidencia que los badnavirus presentes en las muestras Czo-004 y Czo-040 se agrupan en un mismo clado cercano a *Canna yellow mottle virus,* pero distinto al mismo con un porcentaje de identidad de 75% (Tabla 26) entre ellos. Estos resultados nos indicarían que ambos virus corresponden a una nueva especie viral según el criterio de demarcación de especies dentro de los *Badnavirus* (King et.al., 2012).

El nuevo *Torradovirus* se identificó en 11 muestras provenientes de dos localidades de Cusco (Huerta y Paucana), correspondientes a las variedades de papa Yungay y Cica. En las plantas infectadas se observaron sintomatologías como clorosis y leve enrollamiento foliar, al igual que en *Badnavirus* en estas plantas también se identificaron otros virus comunes de papa como PVX, PVY, PVA y PVB. Por la presencia de coinfecciones no es posible atribuir la sintomatología observada exclusivamente al Torradovirus candidato.

Los genomas obtenidos tienen un tamaño de ARN1: 7100 pb aprox y de ARN2: 5400 pb aproximado. El ARN1 presenta un ORF que codifica para una poliproteína donde se encuentran la Helicasa, proteasa y la RdRp, la secuencia de aminoácidos que codifican a la proteína RdRp se utilizó para realizar el árbol filogenético dando como resultados que las 11 muestras se encuentran en el mismo clado y cercano evolutivamente a *Squash clorotic*

leaf spot virus. Al realizar el alineamiento de las secuencias se observó que el *Torradovirus* identificado en las muestras de Cusco y el *Squash clorotic leaf* spot virus presentan una identidad de 70% entre ellas (Tabla 27). Por otro lado, el alineamiento de las secuencias de aminoácidos que codifican la proteína de la cápside evidenció una identidad del 32 % entre el *Torradovirus* identificado en las muestras de Cusco y *Squash chlorotic leaf spot virus* (Tabla 28), lo que indica una marcada divergencia en esta región estructural. Estos resultados indican que estas muestras deben ser consideradas como nuevo virus según el criterio de demarcación de especies descrito por Sanfancon et. al.,2009.

Estos nuevos virus, *Badnavirus y Torradovirus*, han sido identificados en cultivos de papa en varias provincias del Perú. *Badnavirus* fue detectado en cultivos provenientes de Puno y Huánuco (resultados no reportados), mientras que *Torradovirus* se encontró en cultivos provenientes de las provincias de Huánuco, Ica, Junín y Lima, según Álvarez et. al., 2023.

El nuevo *Badnavirus* y *Torradovirus* fueron detectados en las variedades de papa Yungay y Cica**,** las cuales presentan una resistencia limitada o tolerancia parcial frente a infecciones virales. Yungay ha sido descrita como moderadamente susceptible a diversos virus de la papa, lo que favorece la acumulación y diseminación de patógenos cuando se utiliza semilla no certificada (Salazar, 1996), mientras que Cica puede albergar infecciones latentes sin impedir la replicación viral (Valkonen, 2007). La detección de estos nuevos virus en ambas variedades sugiere que la resistencia genética disponible es insuficiente frente a virus emergentes, resaltando la importancia del monitoreo fitosanitario y del uso de semilla certificada (Valkonen, 2007).

En este estudio se sugiere la denominación de *Potato rugose stunting virus* (género *Torradovirus*), de acuerdo con lo propuesto por Álvarez et al. (2023), así como la denominación de *Nisqa virus* (género *Badnavirus*), para las nuevas especies virales detectadas. Álvarez et al. sugirieron el nombre *Potato rugose stunting virus* para el nuevo *Torradovirus,* dado que mediante el uso de la técnica de HTS se identificó a este virus como el único agente viral detectado de en plantas colectadas en diferentes provincias del Perú

(Huánuco, Ica, Junín y Lima), las cuales presentaron una sintomatología caracterizada por mosaico leve y enanismo con rugosidad.

No obstante, actualmente se desconoce el efecto que estos nuevos virus puedan ocasionar en la producción de papa, así como sus propiedades biológicas, por lo que aún no es posible establecer su importancia epidemiológica ni su impacto real en este cultivo.

El adecuado manejo de cultivos, basado en diagnósticos certeros, es fundamental para minimizar pérdidas económicas y maximizar la productividad agrícola. En este contexto, las nuevas perspectivas tecnológicas, como la secuenciación de alto rendimiento, prometen revolucionar el diagnóstico y manejo de virus en la agricultura, promoviendo un enfoque más proactivo y eficiente para garantizar la seguridad alimentaria a nivel global. Sin embargo, es importante destacar que este análisis se basó en muestras recopiladas en 2016, y debido a la alta variabilidad que presentan los virus, es necesario continuar realizando investigaciones relacionadas con la diversidad viral presente en los cultivos de importancia económica en el Perú.

Dada la importancia de identificar nuevos virus en cultivos como la papa, es recomendable combinar técnicas de biología molecular como HTS y PCRs con herramientas bioinformáticas avanzadas para facilitar la identificación de virus. Además, la vigilancia continua y la evaluación de los riesgos asociados con los virus son fundamentales para detectar y prevenir la introducción y propagación de nuevas enfermedades virales en los cultivos de papa.

8. CONCLUSIONES

- La secuenciación de alto rendimiento de siARN permitió identificar dos nuevas especies virales en papa, un *Badnavirus* y un *Torradovirus*, con el apoyo en conjunto de PCR y secuenciación Sanger permitió completar el genoma del *Badnavirus* y obtener parcialmente el genoma del *Torradovirus*, confirmando su presencia en las muestras analizadas y permitiendo su caracterización genómica.
- El software *VirusDetect* se mostró como una herramienta robusta para la identificación de virus conocidos y desconocidos. En comparación con *VirFind*, *VirusDetect* proporcionó mayor cobertura y mejor caracterización de los genomas virales, alcanzando, por ejemplo, un 99,8% del genoma de PVY, frente al 7,5**%** obtenido con *VirFind,* evidenciando la mayor eficacia de *VirusDetect* para la identificación y análisis genómico de virus en papa.
- El análisis filogenético reveló que los nuevos *Badnavirus* y *Torradovirus* se agrupan en clados bien definidos, próximos a virus conocidos pero lo suficientemente divergentes para sugerir nuevas especies.

9. RECOMENDACIONES

- Realizar estudios biológicos sobre los nuevos virus identificados y analizar exhaustivamente el efecto que estos causan en el rendimiento y la salud de las plantas de papa.
- Determinar los mecanismos de propagación de estos nuevos virus, con la identificación de insectos vectores responsables de su transmisión.
- Analizar las posibles sinergias entre los nuevos virus y otros virus conocidos que afectan a los cultivares de papa en el Perú.
- Desarrollar y evaluar estrategias de manejo integrado de plagas para mitigar el impacto de los nuevos virus en los cultivos de papa.

10. REFERENCIAS BIBLIOGRÁFICAS

Aboul-Ata, E., Aboul-Ata, M., Mazyad, H., El-Attar, A., Mohamed Soliman, A., Anfoka, G., & Czosnek, H. (2011). Diagnosis and control of cereal viruses in the Middle East. *Advances in Virus Research*, 81, 33–61.

Alangar, S., Ramasamy, S., & Velusamy, B. (2023). Emerging and re-emerging diseases caused by badnaviruses. *Pathogens*, *12*(2), 245. https://doi.org/10.3390/pathogens12020245

Alvarez-Quinto, R., Amao, M., Muller, G., Fuentes, S., Grinstead, S., Fuentes-Bueno, I., Roenhorst, A., Westenberg, M., Botermans, M., Kreuze, J., & Mollov, D. (2023). Evidence that an unnamed isometric virus associated with potato rugose disease in Peru is a new species of genus *Torradovirus*. *Phytopathology*, *113*(9), 1716–1728. https://doi.org/10.1094/PHYTO-11-22-0449-V

Barrero, R., Napier, K., Cunnington, J., Liefting, L., Keenan, S., Frampton, R., Szabo, T., Hunter, A., Ward, L., Whattam, M., & Bellgard, M. (2017). An internet-based bioinformatics toolkit for plant biosecurity diagnosis and surveillance of viruses and viroids. *BMC Bioinformatics*, *18*, 26. https://doi.org/10.1186/s12859-016-1428-4

Bhat, A., Hohn, T., & Selvarajan, R. (2016). Badnaviruses: The current global scenario. *Viruses*, *8*(6), 177. https://doi.org/10.3390/v8060177

Bologna, N., & Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in *Arabidopsis*. *Annual Review of Plant Biology*, *65*, 473–503. https://doi.org/10.1146/annurev-arplant-050213-040418

Chen, Y.-R., Zheng, Y., Liu, B., Zhong, S., Giovannoni, J., & Fei, Z. (2012). A cost-effective method for Illumina small RNA-seq library preparation using T4 RNA ligase 1 adenylated adapters. *Plant Methods*, *8*, 41. https://doi.org/10.1186/1746-4811-8-41.

Donaire, L., Wang, Y., González-Ibeas, D., Mayer, K. F. X., Aranda, M. A., & Llave, C. (2009). Deep sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology*, *392*(2), 203–214. https://doi.org/10.1016/j.virol.2009.07.005

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*, 13–15.

Expasy. (n.d.). ViralZone: Badnavirus. https://viralzone.expasy.org/729

Hadidi, A., Flores, R., Candresse, T., & Barba, M. (2016). Next-generation sequencing and genome editing in plant virology. *Frontiers in Microbiology, 7*, 1325. https://doi.org/10.3389/fmicb.2016.01325

Harper, G., & Hull, R. (1998). Cloning and sequence analysis of banana streak virus DNA. *Virus Genes, 17*(3), 271–278. https://doi.org/10.1023/A:1008063023887

Hohn, T., Richert-Pöggeler, K., Harper, G., Schwarzacher, T., Teo, C. H., Teycheney, P.-Y., & Hull, R. (2008). Evolution of integrated plant viruses. En *Plant virus evolution* (pp. 54–76).

Instituto Nacional de Estadística e Informática (INEI). (2022, abril). *La producción de papa creció 12,2 % en abril de 2020*. https://m.inei.gob.pe/prensa/noticias/produccion-de-papa-crecio-122-en-abril-del-2020-12263/

Fox, A., Al Rwahnih, M., Boonham, N., & Candresse, T. (2018). Application of HTS for routine plant virus diagnostics: State of the art and challenges. *Frontiers in Plant Science, 9*, 1082. https://doi.org/10.3389/fpls.2018.01082

Geering, A. D. W., Maumus, F., Copetti, D., Choisne, N., Zwickl, D. J., Zytnicki, M., McTaggart, A. R., Scalabrin, S., Vezzulli, S., Wing, R. A., Quesneville, H., & Teycheney, P. Y. (2014). Endogenous florendoviruses are major components of plant genomes and hallmarks of virus evolution. *Nature Communications, 5*, 5269. https://doi.org/10.1038/ncomms6269

Gutiérrez, D., & Kreuze, J. F. (2014). *Protocols for siRNA library preparation*. International Potato Center (CIP)

Iskra-Caruana, M. L., Chabannes, M., & Baurens, F. C. (2010). A four-partner plant–virus interaction: Enemies can also come from within. *Molecular Plant-Microbe Interactions, 23*(11), 1394–1402. https://doi.org/10.1094/MPMI-23-11-1394

Jones, S., Biazan-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

Jones, R. A. C., & Naidu, R. A. (2019). Global dimensions of plant virus diseases: Current status and future perspectives. *Annual Review of Virology, 6*, 387–409. https://doi.org/10.1146/annurev-virology-092818-015606

King, A. M. Q., Adams, M. J., Carstens, E. B., & Lefkowitz, E. J. (Eds.). (2012). *Virus taxonomy: Classification and nomenclature of viruses* (Ninth Report of the International Committee on Taxonomy of Viruses). Academic Press.

Kreuze, J. F., Pérez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., & Simon, R. (2009). Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology, 388*(1), 1–7. https://doi.org/10.1016/j.virol.2009.03.024

Li, Y.-Y., Zhang, R.-N., Xiang, H.-Y., Abouelnasr, H., Li, D.-W., & Yu, J.-L. (2013). Discovery and characterization of a novel carlavirus infecting potatoes in China. *PLOS ONE, 8*(6), e69255. https://doi.org/10.1371/journal.pone.0069255

Loconsole, G., Onelge, N., Potere, O., Giampetruzzi, A., Bozan, O., Satar, S., De Stradis, A., Savino, V., Yokomi, R. K., & Saponari, M. (2012). Identification and characterization of citrus yellow vein clearing virus, a putative new member of the genus *Mandarivirus*. *Phytopathology, 102*(12), 1168–1175. https://doi.org/10.1094/PHYTO-05-12-0103-R

Lockhart, B. E. L., & Olszewski, N. E. (1994). Plant virus badnavirus group. En *Encyclopedia of virology* (pp. 139–143). Academic Press.

Massart, S., Olmos, A., Jijakli, H., & Candresse, T. (2014). Current impact and future directions of high-throughput sequencing in plant virus diagnostics. *Virus Research, 188*, 90–96. https://doi.org/10.1016/j.virusres.2014.03.012

Medberry, S. L., Lockhart, B. E. L., & Olszewski, N. E. (1990). Properties of *Commelina yellow mottle virus* complete DNA sequence, genomic discontinuities, and transcript suggest that it is a pararetrovirus. *Nucleic Acids Research, 18*(18), 5505–5513. https://doi.org/10.1093/nar/18.18.5505

Miozzi, L. (2013). Analysis of small RNAs derived from *Tomato yellow leaf curl Sardinia virus* reveals a cross-reaction between the major viral hotspot and the plant host genome. *Virus Research, 178*(2), 287–296. https://doi.org/10.1016/j.virusres.2013.10.006

Mumford, R., Boonham, N., Tomlinson, J., & Barker, I. (2006). Advances in molecular phytodiagnostics: New solutions for old problems. *Plant Pathology, 55*(1), 1–9. https://doi.org/10.1111/j.1365-3059.2005.01363.x

Pooggin, M. M. (2018). Small RNA-omics for plant virus identification, virome reconstruction, and antiviral defense characterization. *Frontiers in Microbiology, 9*, 277. https://doi.org/10.3389/fmicb.2018.00277

Promega Corporation. (n.d.). *pGEM®-T Easy vector systems*. https://worldwide.promega.com/products/pcr/pcr-cloning/pgem-t-easy-vector-systems/?catNum=A1360

Matthews, R. E. F. (2019). *Plant virology* (5th ed.). Academic Press

Ministerio de Desarrollo Agrario y Riego (MIDAGRI). (2023). *Papa*. https://www.midagri.gob.pe/portal/23-sector-agrario/cultivos-de-importancia-nacional/183-papa?start=1

Roossinck, M. J., Martin, D. P., & Roumagnac, P. (2015**).** Plant virus metagenomics: Advances in virus discovery. *Phytopathology, 105*(6), 716–727. https://doi.org/10.1094/PHYTO-12-14-0356-RVW

Roy, A., Shao, J., Hartung, J. S., Schneider, W. L., & Brlansky, R. H. (2013). A case study on discovery of novel *Citrus leprosis virus* cytoplasmic type 2 utilizing small RNA libraries by next-generation sequencing and bioinformatics analyses. *Journal of Data Mining in Genomics & Proteomics, 4*(2), 1–8. https://doi.org/10.4172/2153-0602.1000129

Ruiz-Ferrer, V., & Voinnet, O. (2009). Roles of plant small RNAs in biotic stress responses. *Annual Review of Plant Biology, 60*, 485–510. https://doi.org/10.1146/annurev.arplant.59.032607.092557

Salazar, L. F. (1996). *Potato viruses and their control*. International Potato Center (CIP).

Sanfançon, H., Wellink, J., Le Gall, O., Karasev, A., van der Vlugt, R., & Wetzel, T. (2009). Secoviridae: A proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus *Torradovirus*. *Archives of Virology, 154*(6), 899–907. https://doi.org/10.1007/s00705-009-0367-z

Seguin, J., Rajeswaran, R., Malpica-López, N., Martin, R. R., Kasschau, K., & Dolja, V. V. (2014). De novo reconstruction of consensus master genomes of plant RNA and DNA viruses from siRNAs. *PLoS ONE, 9*(2), e88513. https://doi.org/10.1371/journal.pone.0088513

Smith, N. A., Eamens, A. L., & Wang, M.-B. (2011). Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathogens, 7*(5), e1002022. https://doi.org/10.1371/journal.ppat.1002022

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution, 38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Valkonen, J. P. T. (2007). Viruses: Economical losses and biotechnological potential. In D. Vreugdenhil et al. (Eds.), *Potato biology and biotechnology: Advances and perspectives* (pp. 619–641). Elsevier.

Verbeek, M., Dullemans, A. M., van den Heuvel, J. F. J. M., Maris, P. C., & van der Vlugt, R. A. A. (2007). Identification and characterisation of Tomato torrado virus, a new plant picorna-like virus from tomato. *Archives of Virology, 152*(5), 881–890. https://doi.org/10.1007/s00705-006-0909-5

Verbeek, M., Dullemans, A. M., van den Heuvel, J. F. J. M., Maris, P. C., & van der Vlugt, R. A. A. (2008). Tomato marchitez virus, a new plant picorna-like virus from tomato related to tomato torrado virus. *Archives of Virology, 153*(1), 127–134. https://doi.org/10.1007/s00705-007-1083-5

Xia, J. (2022). Propuesta para un Día Internacional de la Papa. *Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO)*. https://www.fao.org

Xu, P., Chen, F., Mannas, J. P., Feldman, T., Sumner, L. W., & Roossinck, M. J. (2017). VirusFinder: Software for efficient and accurate detection of viruses and their integration sites in host genomes through next-generation sequencing data. *PLoS ONE, 12*(6), e0180759. https://doi.org/10.1371/journal.pone.0180759

. Zhang, J., Dey, K., Lin, B., Wayne, B., Melzer, M., Sun, Y., Wang, S., Pu, X., Sun, D., & Hu, J. (2017). Characterization of *Canna yellow mottle virus* in a new host, *Alpinia purpurata*, in Hawaii. *Phytopathology, 107*(7), 791–799. https://doi.org/10.1094/PHYTO-01-17-0005-R

Zheng, Y., Gao, S., Padmanabhan, C., & Li, R. (2017). VirusDetect: An automated pipeline for efficient virus discovery using deep sequencing of small RNAs. *PeerJ, 5*, e3198. https://doi.org/10.7717/peerj.3198

11. ANEXOS

Anexo1: Datos de colecta

Tabla 30: Datos obtenidos de la colecta de las muestras de la provincia de Cusco.

Muestra No.	Fecha dd/mm/aaaa	Provincia	Distrito	Localidad	Campo No.	Altitud (m)	Edad cultivo	Estado fenológico	Inter-cultivo	Cultivar/ especie	Origen de semilla	Sistema de manejo	Uso de pesticidas	Notas
Czo-001	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-002	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-003	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-004	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-005	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprx. 95% de plantas mostraron síntomas virales
Czo-006	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Symptomless
Czo-007	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales

Czo-008	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-009	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-010	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-010B (extra)	28/11/2016	Anta	Limatambo	Huerta	1	2,871	4	Floración	No	Yungay	Andahuaylas	Consumo, parte para semilla	Comedores de hoja	Aprox. 95% de plantas mostraron síntomas virales
Czo-011	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-012	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-013	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-014	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-015	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-016	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-017	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-018	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-019	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-020	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			
Czo-020B	28/11/2016	Anta	Limatambo	Cartonpata	2	3,152	4	Floración	No	Cica	Chincheros			

Czo-021	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-022	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-023	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-024	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-025	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-026	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-027	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-028	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-029	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-030	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-031	28/11/2016	Anta	Limatambo	Ayaviri (Pampacona)	3	3,536	4?	Floración		Ccompis	Chincheros	Consumo		
Czo-031	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-032	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		

Czo-033	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-034	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-035	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-036	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-037	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-038	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-039	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-040	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-040B (extra)	28/11/2016	Anta	Ancahuasui	San Martin	4	3,592	4	Floración		Cica	Chacan-Anta	Consumo		
Czo-041	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-042	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-043	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-044	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-045	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-046	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-047	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-048	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-049	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			

Czo-050	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-050B (extra)	29/11/2016	Paruro	Huanoquite	Tiwicte	5	3,432	4			Cica	Chincheros			
Czo-051	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-052	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-053	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-054	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-055	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-056	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-057	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-058	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-059	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-060	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-060B (extra)	29/11/2016	Paruro	Huanoquite	Tiwicte 2	6	3,140				Cica				
Czo-061	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-062	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-063	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-064	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-065	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-066	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-067	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				

Czo-068	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-069	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-070	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-070B (extra)	29/11/2016	Paruro	Yaurisque	Taray	7	3,106	4	Floración		Cica				
Czo-071	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-072	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-073	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-074	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-075	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-076	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-077	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				**Sin síntoma**
Czo-078	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Ccompis				
Czo-079	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Huayro				
Czo-080	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Huayrp				

Czo-080B (extra)	29/11/2016	Cuzco	Santiago	Occopata - Cacchona	8	3,844	1	En crecimiento		Macctillo				
Czo-081	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Macctillo	Cuzco (mercado)			Amarillis (80%, Macctillo (20%). Cambia de semilla cada año
Czo-082	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Macctillo	Cuzco (mercado)			
Czo-083	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Macctillo	Cuzco (mercado)			
Czo-084	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-085	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-086	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-087	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-088	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-089	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-090	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-090B (extra)	30/11/2016	Paucartambo	Challabamba	Chuza	9	2,956	3.5	Floración		Amarillis	Cuzco (mercado)			
Czo-091	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-092	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			

Czo-093	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-094	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-095	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-096	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-097	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-098	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-099	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-100	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-100B (extra)	30/111/2016	Paucartambo	Challabamba	Pilco	10	2,770	2	Vegetativo		Canchan	Cuzco			
Czo-101	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			Cada año cambian de semilla por sequía para evitar problema de polilla, brotamiento rápido. También siembran en época seca para evitar Phytopthora (rancha).
Czo-102	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-103	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-104	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-105	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-106	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-107	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-108	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-109	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			

Czo-110	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-110B (extra)	30/11/2016	Paucartambo	Challabamba	Sunchubamba	11	2,851	4	Floración		Cica	Mercado Cuzco			
Czo-111	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-112	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-113	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-114	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-115	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-116	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-117	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-118	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-119	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-120	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-120B (extra)	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Cica	Mercado de Cuzco			
Czo-120C (extra)	30/11/2016	Paucartambo	Colquepata	Paucana	12	3,381	4	Floración		Alianza	Mercado de Cuzco			Variedad Alianza
Czo-121	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-122	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-123	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-124	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			

Czo-125	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-126	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-127	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-128	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-129	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-130	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			
Czo-130B (extra)	30/11/2016	Calca	Pisac	Sacaca	13	3,518	4	Floración		Yungay	Mercado de Cuzco			

Anexo 2: Protocolo de purificación de agarosa

Se utilizó el High Pure PCR Product Purification Kit (ROCHE), el protocolo utilizado para la purificación fue el siguiente:

- Cortar el gel agarosa que contiene la banda esperada (aprox 200mg) y colocarlo en un tubo de 2 mL estéril.
- Adicionar 600uL de Binding Buffer e incubar por 10 min a 65°C realizando un vortex cada 2 minutos.
- Luego adicionar 300 uL de isopropanol y realizar un vortex vigorosamente.
- Colocar esta mezcla dentro del High Pure Filter Tube y centrifugar a 12000 rpm por 1 minuto.
- Descartar el líquido precipitado y adicionar al filtro 600 uL de Wash Buffer y centrifugar a 12000 rpm por 1 minuto.
- Decantar el líquido precipitado.
- Adicionar al filtro 200 uL de Wash Buffer y centrifugar a 12000 rpm por 1 minuto descartar el precipitado.
- Centrifugar el filtro en seco a 12000 rpm por 2 minutos.
- Colocar el filtro en un tubo 1.5 ml estéril y adicionar 30 uL de Elution Buffer e incubar a temperatura ambiente por 5 minutos.
- Luego, centrifugar por 1 minuto a 12000 rpm.

Anexo 3: Protocolo de cuantificación con fluorómetro Qubit

Preparar la siguiente solución en tubos de 0.5ml estériles

- 0.5ul del fluorescente
- 197.5 uL de Elution Buffer
- 2 uL de muestra

Anexo 4: Protocolos de PCR:

- *Primers* Badna-4 y Badna-6

Tabla 31: Master mix empleada para la amplificación de los *primers* Badna-4 y Badna-6

Reactivos	Vol. (µL)
Buffer 5x	5
MgCl2	2.5
dNTPs (10uM)	0.5
Primer FORWARD(*)	0.5
Primer REVERSE(*)	0.5
NFW	13.8
Taq pol(200 U/uL)	0.2

(*) los *primers* Badna-4 y Badna-6 se utilizaron en una concentración de 10uM.

Las condiciones de amplificación fueron:

94 °C	2 min	
94 °C	30 seg	35 ciclos
60 °C	30 seg	
72 °C	1 min	
72 °C	10 min	
4 °C	∞	

- *Primers* Badna-5, Badna-7, TORRA1, TORRA2, TORRA2.1, TORRA3 y TORRA4

  La amplificación de estos *primers* fue realizada utilizando Phusion Flash High-Fidelity PCR Master Mix.

Tabla 32: Master mix empleada para la amplificación de los *primers* Badna-5, Badna-7, Torra-1, Torra-2, Torra-3 y Torra-4

Componentes	Vol. Ul	Concentracion final
2X Phusion ™ Flash PCR Master Mix	10uL	1x
cebadores_F (10 uM)(*)	0.5uL	0.25 uM
cebadores_R(10 uM)(*)	0.5uL	0.25 uM
NFW	7uL	
ADN	2uL	

(*) las concentraciones y el volumen de los componentes de la PCR fueron los mismos para cada par de cebadores.

Las condiciones de amplificación fueron:

98 °C 10 seg

98 °C 1seg

(*) (**) seg — 30 ciclos

72 °C 15seg

72 °C 1min

4 °C ∞

(*) y (**) la temperatura y el tiempo de *annealing* varía según el par de cebadores siendo:

Tabla 33: Temperatura y tiempo de *annealing* utilizados para la amplificación de los *primers* Badna-5, Badna-7, Torra-1, Torra-2, Torra-2.1 Torra-3 y Torra-4

Primers	**T° *Alineamiento***	**Tiempo de alineamiento**
Badna-5	62°C	45 seg
Badna-7	58°C	30 seg
Torra-1	52°C	45 seg
Torra-2	55°C	20 seg
Torra-2.1	56°C	30 seg
Torra-3	52°C	30 seg
Torra-4	55°C	30 seg

Anexo 5: Resultados obtenidos usando VirusDetect

Tabla 34: Resultados obtenidos de la secuenciación de las 144 muestras colectadas en cusco.

Muestra	Reference	Length	Coverage (%)	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
Cus-001	JQ971975	9714	8523 (87.7)	1.8	95.66	98.39	90.2	potyvirus	Potato virus Y
	ACT79984.1	1197	318 (26.6)	15.5	44.38	55.56	40.43	NA	Tomato chocolate spot virus
	ABV44417.1	190	40 (21.1)	24.3	52.5	52.5	52.5	torradovirus	Tomato marchitez virus
	AMN91910.1	2207	386 (17.5)	8.1	58.4	85.71	42.55	torradovirus	Squash chlorotic leaf spot virus
	AHA85557.1	1167	274 (23.5)	14.8	46.91	76.47	41.03	torradovirus	Carrot torradovirus
	AHA91818.1	973	213 (21.9)	10.9	49.3	58.82	41.86	torradovirus	Cassava Torrado-like virus
Cus-002	Q971975	9714	8351 (86)	2.4	95.28	100	89.89	potyvirus	Potato virus Y
	AMN91912.1	1054	116 (11)	7.4	50.86	61.76	42.86	torradovirus	Squash chlorotic leaf spot virus
Cus-003	JQ971975	9714	9131 (94)	2.6	95.1	98.89	90.24	potyvirus	Potato virus Y
Cus-004	M84483	2265	1334 (58.9)	11.3	83.44	93.22	81.02	comovirus	Andean potato mottle virus
	AF522296	9704	9704 (100)	70.4	95.02	100	89.04	potyvirus	Potato virus Y
	ACF60613.1	1916	292 (15.2)	51.3	80.14	97.56	68.57	badnavirus	Cycad leaf necrosis virus
	ADN88180.1	175	164 (93.7)	99.7	50.61	76.74	41.32	badnavirus	Pineapple bacilliform
	ANN87818.1	1684	1442 (85.6)	73.7	71.89	94	56.36	badnavirus	Canna yellow mottle virus
	ANN87816.1	133	133 (100)	60.1	42.54	42.54	42.54	badnavirus	Canna yellow mottle virus
	ACS44708.1	1792	1038 (57.9)	63.4	72.26	96	41.9	badnavirus	Sugarcane bacilliform virus
	AAA42421.1	997	780 (78.2)	14.9	74.91	88.89	55.17	comovirus	Andean potato mottle virus
	ACX53644.1	1857	551 (29.7)	21.7	40.07	57.5	25.66	comovirus	Turnip ringspot virus
	ADI60054.1	1863	524 (28.1)	19.7	39.74	53.93	27.78	comovirus	Broad bean true mosaic virus
	AAA42422.1	729	245 (33.6)	10.1	94	95.83	84.21	comovirus	Andean potato mottle virus

	AIG59705.1	186	146 (78.5)	6.4	71.71	83.33	61.76	comovirus	Andean potato mottle virus
Cus-005	JQ971975	9714	9547 (98.3)	6.7	94.59	98.55	89.9	potyvirus	Potato virus Y
	ACT79984.1	1197	274 (22.9)	10	45.82	66.67	28.43	NA	Tomato chocolate spot virus
Cus-006	JQ971975	9714	9138 (94.1)	5.4	94.71	100	89.13	potyvirus	Potato virus Y
Cus-007	KJ534601	6445	6202 (96.2)	4.6	96.61	100	84.5	potexvirus	Potato virus X
	KR270797	9726	9133 (93.9)	5.1	95.37	100	89.66	potyvirus	Potato virus Y
	AJ131403	9572	1362 (14.2)	3.6	92.46	100	84.29	potyvirus	Potato virus A
	AMN91910.1	2207	344 (15.6)	8.9	59.3	100	33.33	torradovirus	Squash chlorotic leaf spot virus
	AMN91912.1	1054	171 (16.2)	10	49.71	63.64	32.65	torradovirus	Squash chlorotic leaf spot virus
	AHA91818.1	973	151 (15.5)	12.2	47.68	61.9	37.78	torradovirus	Cassava Torrado-like virus
	AHA85557.1	1167	126 (10.8)	9.4	49.22	64.1	40	torradovirus	Carrot torradovirus
	AHA91817.1	659	126 (19.1)	10.2	63.49	67.61	55.56	torradovirus	Cassava Torrado-like virus
Cus-008	JQ971975	9714	9669 (99.5)	19.5	93.41	98.44	88.78	potyvirus	Potato virus Y
	AHA91818.1	973	213 (21.9)	19.2	47.25	56.52	32.84	torradovirus	Cassava Torrado-like virus
	AMN91912.1	1054	166 (15.7)	19	45.29	54.55	33.33	torradovirus	Squash chlorotic leaf spot virus
Cus-009	D00461	3552	384 (10.8)	5.8	97.66	100	95.45	carlavirus	Potato virus S
	KJ534601	6445	6442 (100)	278.1	98.09	98.67	90.28	potexvirus	Potato virus X
	JQ971975	9714	9525 (98.1)	28.6	93.41	100	86.57	potyvirus	Potato virus Y
	AHM92948.1	1832	326 (17.8)	92.7	70.25	84.78	35.42	badnavirus	Banana streak virus.
	AEC49887.1	1836	320 (17.4)	108.5	79.69	84.78	60.34	badnavirus	Banana streak virus.
	ACF60611.1	175	175 (100)	115.4	51.43	51.43	51.43	badnavirus	Cycad leaf necrosis virus
	ANN87818.1	1684	1355 (80.5)	102.3	73.64	85.22	60.71	badnavirus	Canna yellow mottle virus
	ACB14289.1	282	126 (44.7)	63.2	50.39	76.32	39.33	badnavirus	Pineapple bacilliform
	ACS44708.1	1792	1118 (62.4)	103.2	73.27	88.31	40.57	badnavirus	Sugarcane bacilliform virus
	AAA42421.1	997	975 (97.8)	57	72.86	75.63	72.47	comovirus	Andean potato mottle virus

	AAA42422.1	729	715 (98.1)	61.4	94.55	94.55	94.55	comovirus	Andean potato mottle virus
	ABV44417.1	190	40 (21.1)	15.5	52.5	52.5	52.5	torradovirus	Tomato marchitez virus
	AMN91910.1	2207	291 (13.2)	16.3	50	80	27.1	torradovirus	Squash chlorotic leaf spot virus
Cus-010	JQ647830	8482	6014 (70.9)	25.2	86.46	95.12	81.81	carlavirus	Potato virus S
	KJ534601	6445	6443 (100)	255.1	93.67	98.67	81.82	potexvirus	Potato virus X
	KU586450	9600	3423 (35.7)	1.6	98.36	100	95.24	potyvirus	Potato virus A
	AFJ22628.1	1974	759 (38.4)	21	77.73	79.67	43.9	carlavirus	Potato virus S
	AAF89749.1	115	19 (16.5)	92.1	94.74	94.74	94.74	potexvirus	Potato virus X
	AIS39931.1	1456	157 (10.8)	223.4	85.99	100	83.82	potexvirus	Potato virus X
	AHA91818.1	973	231 (23.7)	6.2	54.74	66.67	39.53	torradovirus	Cassava Torrado-like virus
	AMN91910.1	2207	1462 (66.2)	12	45.57	81.94	25.26	torradovirus	Squash chlorotic leaf spot virus
	AIT59086.1	205	104 (50.7)	6.9	37.74	37.74	37.74	torradovirus	Motherwort yellow mottle virus
Cus-010B	KJ534601	6445	6356 (98.6)	7.8	94.86	100	83.64	potexvirus	Potato virus X
	JQ971975	9714	9667 (99.5)	7.6	94.82	100	86.92	potyvirus	Potato virus Y
	AOE46790.1	1231	812 (66)	18.9	55.29	75.51	41.46	enamovirus	Grapevine enamovirus
	AOE46789.1	813	512 (63)	20.8	41.7	48.57	40	enamovirus	Grapevine enamovirus
	AOE46792.1	544	380 (69.9)	23.5	54.01	55.76	46.34	enamovirus	Grapevine enamovirus
Cus-011	JQ971975	9714	9617 (99)	45.2	93.15	99.14	89.3	potyvirus	Potato virus Y
Cus-012	AJ238607	6043	3257 (53.9)	2.6	98.89	100	95.65	pomovirus	Potato mop-top virus
	JQ971975	9714	9631 (99.1)	197.6	92.88	100	87.16	potyvirus	Potato virus Y
	AOE46790.1	1231	720 (58.5)	24.2	59.36	82.22	30.77	enamovirus	Grapevine enamovirus
	AOE46792.1	544	353 (64.9)	22	55.21	70	49.09	enamovirus	Grapevine enamovirus
	AOE46788.1	314	167 (53.2)	31.3	23.39	23.39	23.39	enamovirus	Grapevine enamovirus
Cus-013	KX656670	7148	4067 (56.9)	12.2	94.71	98.58	85.94	nepovirus	Potato virus B
	JQ971975	9714	9591 (98.7)	27.9	93.34	100	89.58	potyvirus	Potato virus Y

Cus-014	JQ971975	9714	9148 (94.2)	9.7	93.21	100	83.08	potyvirus	Potato virus Y
Cus-015	KJ534601	6445	5197 (80.6)	1.6	98.36	100	91.84	potexvirus	Potato virus X
	JQ971975	9714	9623 (99.1)	7.2	94.78	100	84.85	potyvirus	Potato virus Y
Cus-016	M72416	7568	7563 (99.9)	66.3	96.82	100	94.51	potexvirus	Potato virus X
	JQ971975	9714	9408 (96.8)	7.2	95.58	100	90.48	potyvirus	Potato virus Y
Cus-017	M84483	2265	921 (40.7)	36.1	82.89	89.8	81.73	comovirus	Andean potato mottle virus
	KJ534601	6445	3501 (54.3)	1.9	98.46	100	93.33	potexvirus	Potato virus X
	KJ634024	9724	9724 (100)	560.2	95.52	100	85.71	potyvirus	Potato virus Y
	AAA42421.1	997	969 (97.2)	80.4	72.48	79.37	70.56	comovirus	Andean potato mottle virus
	AKL83120.1	1856	549 (29.6)	129	37.63	42.57	30.23	comovirus	Bean rugose mosaic virus
	AAA42422.1	729	446 (61.2)	65.1	94.62	94.62	94.62	comovirus	Andean potato mottle virus
	AAO23120.1	261	91 (34.9)	15.6	62.64	62.64	62.64	potyvirus	Potato virus Y
Cus-018	KJ534601	6445	2912 (45.2)	10.8	98	100	91.38	potexvirus	Potato virus X
	KU586450	9600	3843 (40)	11.2	98.08	100	91.67	potyvirus	Potato virus A
Cus-019	JQ971975	9714	9706 (99.9)	22.3	93.67	97.22	87.69	potyvirus	Potato virus Y
Cus-020	JQ971975	9714	9657 (99.4)	10.4	93.99	100	87.5	potyvirus	Potato virus Y
Cus-20B	JQ971975	9714	9577 (98.6)	54.2	91.99	100	82.07	potyvirus	Potato virus Y
	AOE46790.1	1231	522 (42.4)	66.5	45.49	55.36	41.07	enamovirus	Grapevine enamovirus
	AOE46788.1	314	246 (78.3)	130	24.4	24.4	24.4	enamovirus	Grapevine enamovirus
	AOE46792.1	544	215 (39.5)	63.1	55.96	55.96	55.96	enamovirus	Grapevine enamovirus
	ABV01912.1	542	341 (62.9)	70.4	51.31	51.31	51.31	polerovirus	Tobacco vein distorting virus
Cus-021	KJ534601	6445	1936 (30)	4.2	98.24	100	92.59	potexvirus	Potato virus X
	JQ971975	9714	9714 (100)	788	91.46	98.08	82.69	potyvirus	Potato virus Y
Cus-022	KX656670	7148	2938 (41.1)	13.7	96.06	100	92.86	nepovirus	Potato virus B
	KJ534601	6445	1081 (16.8)	6.8	97.96	100	93.48	potexvirus	Potato virus X
	KF770835	9700	987 (10.2)	520	91.83	94.34	85.84	potyvirus	Potato virus Y

Cus-023	KJ534601	6445	1497 (23.2)	10.1	98.46	100	93.33	potexvirus	Potato virus X
	JQ971975	9714	9714 (100)	666.4	91.6	100	81.44	potyvirus	Potato virus Y
Cus-024	KJ534601	6445	1597 (24.8)	7	98.69	100	96.61	potexvirus	Potato virus X
	JQ971975	9714	9709 (99.9)	732.1	91.77	100	82.85	potyvirus	Potato virus Y
Cus-025	FJ643622	885	138 (15.6)	24.6	96	97.06	94.64	carlavirus	Potato virus S
	KJ534601	6445	1759 (27.3)	5.5	98.58	100	95.45	potexvirus	Potato virus X
	JQ971975	9714	9580 (98.6)	891.8	92.82	100	84.56	potyvirus	Potato virus Y
Cus-026	KJ534601	6445	2028 (31.5)	5.2	98.22	100	94.55	potexvirus	Potato virus X
	JQ971975	9714	9714 (100)	790.3	91.56	96.83	81.31	potyvirus	Potato virus Y
	KU586450	9600	3022 (31.5)	25.5	97.98	100	90.48	potyvirus	Potato virus A
Cus-027	KX656670	7148	6589 (92.2)	82.8	90.93	96.72	86.21	nepovirus	Potato virus B
	KJ634024	9724	9614 (98.9)	652.1	90.13	100	86.44	potyvirus	Potato virus Y
Cus-028	KR270797	9726	9607 (98.8)	694.4	93.6	100	81.21	potyvirus	Potato virus Y
Cus-029	JQ971975	9714	9714 (100)	739.4	91.3	100	87.7	potyvirus	Potato virus Y
Cus-030	GU563327	8995	919 (10.2)	6	97.75	100	92.68	ipomovirus	Cassava brown streak virus
	KJ634024	9724	9540 (98.1)	669	90.43	100	86.23	potyvirus	Potato virus Y
Cus-030B	JQ971975	9714	9652 (99.4)	773.3	93.08	100	89.33	potyvirus	Potato virus Y
Cus-031	KR108833	8742	896 (10.2)	3.5	98.1	100	95.12	ipomovirus	Cassava brown streak virus
	KX656670	7148	6747 (94.4)	47.6	89.61	98.25	85.55	nepovirus	Potato virus B
	JQ971975	9714	9622 (99.1)	445.7	91.81	97.67	84.21	potyvirus	Potato virus Y
Cus-032	KX656670	7148	4843 (67.8)	64.8	92.49	100	85.95	nepovirus	Potato virus B
	KJ534601	6445	2128 (33)	10.1	98.78	100	95.24	potexvirus	Potato virus X
	JQ971975	9714	9671 (99.6)	852	92.83	100	86.49	potyvirus	Potato virus Y
Cus-033	KJ534601	6445	2035 (31.6)	5	98.62	100	95.24	potexvirus	Potato virus X
	JQ971975	9714	9584 (98.7)	732.1	92.77	100	87.36	potyvirus	Potato virus Y
Cus-034	KJ534601	6445	2162 (33.5)	3.8	98.66	100	95.65	potexvirus	Potato virus X

	JQ971975	9714	9614 (99)	453.1	92.7	100	89.22	potyvirus	Potato virus Y
Cus-035	KJ534601	6445	2052 (31.8)	12.7	98	100	91.67	potexvirus	Potato virus X
	JQ971975	9714	9495 (97.7)	699.7	92.56	100	88.16	potyvirus	Potato virus Y
Cus-036	KR108833	8742	1196 (13.7)	57.5	97.5	100	93.18	ipomovirus	Cassava brown streak virus
	JQ971975	9714	9637 (99.2)	829.2	92.24	100	86.08	potyvirus	Potato virus Y
Cus-037	JQ971975	9714	9552 (98.3)	630.9	92.99	97.37	88.46	potyvirus	Potato virus Y
Cus-038	KR108833	8742	1401 (16)	4.7	97.41	100	89.36	ipomovirus	Cassava brown streak virus
	JQ971975	9714	9631 (99.1)	454.7	92.82	100	87.69	potyvirus	Potato virus Y
Cus-039	JF909090	2802	518 (18.5)	6.2	96.43	100	91.94	begomovirus	East African cassava mosaic virus
	KR108833	8742	2144 (24.5)	5.6	97.48	100	93.48	ipomovirus	Cassava brown streak virus
	JQ971975	9714	9714 (100)	756.4	96.51	100	92.06	potyvirus	Potato virus Y
Cus-040	M84483	2265	1343 (59.3)	43.5	81.81	81.81	81.81	comovirus	Andean potato mottle virus
	KR108833	8742	1982 (22.7)	2.3	96.88	100	90.91	ipomovirus	Cassava brown streak virus
	KX656670	7148	7073 (99)	19.9	97.2	100	89.04	nepovirus	Potato virus B
	JQ971975	9714	9714 (100)	357.4	94.25	100	85	potyvirus	Potato virus Y
	ANN87817.1	173	96 (55.5)	77.9	54.9	61.67	45.24	badnavirus	Canna yellow mottle virus
	AEC49888.1	176	78 (44.3)	85.1	62.03	68	59.26	badnavirus	Banana streak virus.
	ACS44708.1	1792	1518 (84.7)	106	69.96	97.96	40	badnavirus	Sugarcane bacilliform virus
	ANN87818.1	1684	1466 (87.1)	97.5	71.97	95.92	41.18	badnavirus	Canna yellow mottle virus
	ANN87816.1	133	111 (83.5)	56.2	41.96	41.96	41.96	badnavirus	Canna yellow mottle virus
	AAA42421.1	997	954 (95.7)	57.3	72.16	72.62	67.44	comovirus	Andean potato mottle virus
	AKL83120.1	1856	1201 (64.7)	67	39.93	45.16	34.45	comovirus	Bean rugose mosaic virus
	CBL43396.1	1499	782 (52.2)	27.4	45.44	78.57	31.3	nepovirus	Grapevine Bulgarian latent virus
	AAD09329.2	1626	456 (28)	20.3	57.61	82.14	38.21	nepovirus	Blackcurrant reversion virus

	CBL43395.1	2095	1606 (76.7)	28.2	68.67	100	44.19	nepovirus	Grapevine Bulgarian latent virus
Cus-040B	KM659859	6487	1869 (28.8)	3.3	96.51	100	93.33	potexvirus	Potato virus X
	KJ634024	9724	9722 (100)	642.2	91.06	100	86.72	potyvirus	Potato virus Y
Cus-041	KX656671	4527	2883 (63.7)	8	91	97.73	84.4	nepovirus	Potato virus B
	JQ971975	9714	9346 (96.2)	15.5	93.31	100	86.32	potyvirus	Potato virus Y
Cus-042	DQ355155	219	48 (21.9)	11.1	89.58	89.58	89.58	caulimovirus	Cauliflower mosaic virus
	KJ534601	6445	970 (15.1)	8.6	98.56	100	94.23	potexvirus	Potato virus X
	JQ971975	9714	9526 (98.1)	716.6	92.65	100	88	potyvirus	Potato virus Y
Cus-043	JQ971975	9714	9479 (97.6)	10.1	94.12	100	88.33	potyvirus	Potato virus Y
Cus-044	KR270797	9726	9355 (96.2)	11.7	94.21	100	85.56	potyvirus	Potato virus Y
Cus-045	KR270797	9726	9119 (93.8)	18.9	94.09	100	87.74	potyvirus	Potato virus Y
Cus-046	KJ534601	6445	1442 (22.4)	9.5	97.71	100	92.86	potexvirus	Potato virus X
	JQ971975	9714	9569 (98.5)	753.4	92.84	100	83.52	potyvirus	Potato virus Y
Cus-047	KJ534601	6445	2036 (31.6)	2.3	98.67	100	96.23	potexvirus	Potato virus X
	JQ971975	9714	9375 (96.5)	517.3	92.87	100	88.07	potyvirus	Potato virus Y
	KP849483	9859	5457 (55.4)	28.2	97.6	100	91.3	potyvirus	Potato virus V
Cus-048	AY763582	714	104 (14.6)	7.2	98.08	100	96.55	potexvirus	Potato virus X
	JQ971975	9714	9594 (98.8)	622.5	92.76	98.99	85.59	potyvirus	Potato virus Y
	KP849483	9859	5521 (56)	28.6	97.52	100	91.22	potyvirus	Potato virus V
Cus-049	JQ971975	9714	9713 (100)	441	96.44	100	90.91	potyvirus	Potato virus Y
Cus-050	KJ534601	6445	1235 (19.2)	3.4	98.3	100	94.74	potexvirus	Potato virus X
	JQ971975	9714	9712 (100)	482.1	96.7	100	89.55	potyvirus	Potato virus Y
	KP849483	9859	4128 (41.9)	26.3	97.77	100	92.16	potyvirus	Potato virus V
Cus-050B	KX656671	4527	1324 (29.2)	5.7	94.53	100	88.41	nepovirus	Potato virus B
	KR265317	1902	866 (45.5)	2.4	94.43	96.92	92.08	potyvirus	Potato virus Y
Cus-051	KJ534601	6445	678 (10.5)	11.2	97.94	100	90.91	potexvirus	Potato virus X

	JQ971975	9714	9418 (97)	770.6	92.74	100	89.33	potyvirus	Potato virus Y
Cus-052	JQ971975	9714	9511 (97.9)	8.2	93.36	100	85.51	potyvirus	Potato virus Y
Cus-053	KR270797	9726	8482 (87.2)	3.4	95.43	100	90.34	potyvirus	Potato virus Y
Cus-054	KJ534601	6445	1280 (19.9)	2.1	97.81	100	95.24	potexvirus	Potato virus X
	KR270797	9726	8046 (82.7)	1.9	95.85	100	89.29	potyvirus	Potato virus Y
Cus-055	JQ971975	9714	9153 (94.2)	10.8	93.78	100	88.46	potyvirus	Potato virus Y
Cus-056	KR270797	9726	8471 (87.1)	10.8	95.34	100	90.11	potyvirus	Potato virus Y
Cus-057	JQ971975	9714	8587 (88.4)	2.9	95.14	100	91.67	potyvirus	Potato virus Y
Cus-058	JQ971975	9714	8875 (91.4)	4	94.79	100	88.16	potyvirus	Potato virus Y
Cus-059	KJ534601	6445	2291 (35.5)	5.1	98.08	100	92.68	potexvirus	Potato virus X
	JQ971975	9714	9488 (97.7)	757.5	92.61	100	87.01	potyvirus	Potato virus Y
Cus-060	KJ534601	6445	2366 (36.7)	3.7	98.01	100	92.86	potexvirus	Potato virus X
	JQ971975	9714	9673 (99.6)	335.7	93.32	98.21	87.88	potyvirus	Potato virus Y
	KP849483	9859	6883 (69.8)	12.9	97.79	100	92.5	potyvirus	Potato virus V
Cus-060B	JQ971975	9714	8555 (88.1)	8.8	95.25	100	89.92	potyvirus	Potato virus Y
Cus-061	JQ971975	9714	9016 (92.8)	753.9	92.78	97.47	91.64	potyvirus	Potato virus Y
Cus-062	KJ534601	6445	686 (10.6)	19.7	98.69	100	95.74	potexvirus	Potato virus X
	JQ971975	9714	9446 (97.2)	799.2	92.93	97.16	88.24	potyvirus	Potato virus Y
Cus-063	JQ971975	9714	8194 (84.4)	1.6	95.4	100	89.8	potyvirus	Potato virus Y
	KP849483	9859	5374 (54.5)	1.6	98.64	100	94.83	potyvirus	Potato virus V
Cus-064	JQ971975	9714	9009 (92.7)	2.4	95.2	100	90.2	potyvirus	Potato virus Y
Cus-065	KJ534601	6445	1178 (18.3)	3.6	98.39	100	95.24	potexvirus	Potato virus X
	JQ971975	9714	9652 (99.4)	430.4	92.73	100	83.78	potyvirus	Potato virus Y
	AOE46792.1	544	324 (59.6)	21.8	55.05	56.11	52.83	enamovirus	Grapevine enamovirus
	AOE46790.1	1231	308 (25)	21.3	53.07	82.76	35.59	enamovirus	Grapevine enamovirus
	AHA91816.1	152	41 (27)	14.5	51.22	51.22	51.22	polerovirus	Cassava Polero-like virus

Cus-066	KJ534601	6445	959 (14.9)	7.8	98.12	100	94.74	potexvirus	Potato virus X
	KJ634024	9724	9628 (99)	761.7	93.49	100	87.57	potyvirus	Potato virus Y
Cus-067	JQ971975	9714	9650 (99.3)	12	93.25	100	85.09	potyvirus	Potato virus Y
Cus-068	KJ534601	6445	2692 (41.8)	6.4	98.48	100	96	potexvirus	Potato virus X
	KJ634024	9724	9694 (99.7)	608.7	92.14	98.25	85.71	potyvirus	Potato virus Y
Cus-069	KJ534601	6445	2165 (33.6)	5.6	98.2	100	95.35	potexvirus	Potato virus X
	KJ634024	9724	9721 (100)	681.8	95.69	100	85.71	potyvirus	Potato virus Y
Cus-070	JQ971975	9714	9702 (99.9)	13.3	94.12	100	87.5	potyvirus	Potato virus Y
Cus-070B	AF172259	6435	1582 (24.6)	1.7	96.27	98.86	90.74	potexvirus	Potato virus X
	JQ971975	9714	9685 (99.7)	134.5	92.67	100	86.78	potyvirus	Potato virus Y
Cus-071	AF172259	6435	6241 (97)	140.1	89.46	96.08	80.6	potexvirus	Potato virus X
	JQ971975	9714	9650 (99.3)	438	92.27	100	82.78	potyvirus	Potato virus Y
	AAW50963.1	190	130 (68.4)	34.1	52.31	52.31	52.31	NA	Banana streak virus.
	AGN48134.1	239	181 (75.7)	42	34.18	34.18	34.18	allexivirus	Garlic virus C
	AJS10757.1	2090	316 (15.1)	17.1	33.13	43.48	29.22	carlavirus	Sweet potato chlorotic fleck virus
	CAC39422.1	1962	546 (27.8)	14.3	37.81	68.42	28.48	citrivirus	Citrus leaf blotch virus
	AHF50167.1	2183	439 (20.1)	23.3	39.37	65.79	27.27	foveavirus	Apple stem pitting virus
	AKN08994.1	1986	709 (35.7)	20.9	40.99	65.79	27.85	prunevirus	Caucasus prunus virus
	CDF66416.2	1679	449 (26.7)	21.3	51.1	66.07	36.26	prunevirus	Apricot vein clearing associated virus
	BAB47271.2	1867	288 (15.4)	19.9	37.84	69.64	24.68	trichovirus	Grapevine berry inner necrosis virus
Cus-072	KJ534601	6445	735 (11.4)	2.3	97.69	100	95.79	potexvirus	Potato virus X
	JQ971975	9714	9376 (96.5)	40.6	93.36	100	89.87	potyvirus	Potato virus Y
Cus-073	JQ647830	8482	3663 (43.2)	17.2	92.76	100	83.56	carlavirus	Potato virus S
	KX656671	4527	2193 (48.4)	34.4	91.27	98.45	85	nepovirus	Potato virus B
	KJ534601	6445	1016 (15.8)	7.1	98.23	100	94	potexvirus	Potato virus X

	EU563512	9699	970 (10)	404.2	90.93	95.12	88.12	potyvirus	Potato virus Y
	AAW50958.1	301	36 (12)	17.6	77.78	84.21	70.59	NA	Banana streak virus.
	AAL17724.1	400	131 (32.8)	69.1	36.76	50	30.43	botrexvirus	Botrytis virus X
	CAI30423.2	302	59 (19.5)	63.9	33.9	33.9	33.9	cytorhabdovirus	Lettuce necrotic yellows virus
Cus-074	M72416	7568	2852 (37.7)	169.6	82.96	97.78	79.86	potexvirus	Potato virus X
	JQ971975	9714	9582 (98.6)	198.4	92.73	100	88	potyvirus	Potato virus Y
Cus-075	KM659859	6487	2488 (38.4)	354.1	86.08	97.62	79.39	potexvirus	Potato virus X
	JQ971975	9714	9695 (99.8)	458.2	91.91	100	83.26	potyvirus	Potato virus Y
	ALZ41789.1	358	106 (29.6)	27.6	39.81	39.81	39.81	citrivirus	Citrus leaf blotch virus
	AKN08995.1	438	156 (35.6)	20.1	30.13	30.13	30.13	prunevirus	Caucasus prunus virus
Cus-076	M72416	7568	6367 (84.1)	73.7	86.68	98.53	81.37	potexvirus	Potato virus X
	JQ971975	9714	9645 (99.3)	531.5	92.92	100	85.83	potyvirus	Potato virus Y
	KU586450	9600	3620 (37.7)	5.7	98.23	100	90.99	potyvirus	Potato virus A
Cus-077	JQ971975	9714	8786 (90.4)	3	95.37	100	88.24	potyvirus	Potato virus Y
Cus-078	KJ534601	6445	2209 (34.3)	12.7	97.96	100	89.47	potexvirus	Potato virus X
	JQ971975	9714	9714 (100)	724	91.92	97.5	81.88	potyvirus	Potato virus Y
	KU586450	9600	7568 (78.8)	22.8	98.3	100	92.65	potyvirus	Potato virus A
Cus-079	KX656670	7148	6297 (88.1)	58.9	91.3	97.75	83.96	nepovirus	Potato virus B
	KJ534601	6445	2577 (40)	1.9	98.22	100	94.31	potexvirus	Potato virus X
	JQ971975	9714	9678 (99.6)	444.8	93.2	100	87.74	potyvirus	Potato virus Y
	KU586450	9600	9577 (99.8)	217.3	98.13	99.4	92.45	potyvirus	Potato virus A
Cus-080	KJ534601	6445	2592 (40.2)	2.2	97.8	100	94.59	potexvirus	Potato virus X
	KU586450	9600	9600 (100)	446.1	98.1	98.57	91.3	potyvirus	Potato virus A
	JQ971975	9714	9360 (96.4)	11.8	93.79	98.53	88.89	potyvirus	Potato virus Y
Cus-080B	JQ647830	8482	4396 (51.8)	25.1	91.45	98	83.02	carlavirus	Potato virus S

	KJ534601	6445	3381 (52.5)	2.3	98.36	100	91.84	potexvirus	Potato virus X
	EU482153	9691	977 (10.1)	479.9	90	98.48	86.92	potyvirus	Potato virus Y
	BAC16226.1	2040	381 (18.7)	20.5	56.1	81.48	38.89	varicosavirus	Lettuce big-vein virus
Cus-081	KJ534601	6445	687 (10.7)	7	98.4	100	95.16	potexvirus	Potato virus X
	FJ214726	9773	979 (10)	549.1	89.91	97.14	85.71	potyvirus	Potato virus Y
	AGC24022.1	306	157 (51.3)	16	37.5	37.5	37.5	NA	Rhabdovirus
	AJP67515.1	427	157 (36.8)	14.8	27.67	27.67	27.67	cytorhabdovirus	Barley yellow striate mosaic virus
	AFA36173.1	290	84 (29)	18.8	29.76	29.76	29.76	varicosavirus	Lettuce big-vein virus
	BAC16226.1	2040	1383 (67.8)	65	46.65	60	31.01	varicosavirus	Lettuce big-vein virus
Cus-082	JQ647830	8482	7145 (84.2)	40.7	87.75	98.67	78.63	carlavirus	Potato virus S
	JQ971975	9714	9452 (97.3)	636.4	93.03	100	85.33	potyvirus	Potato virus Y
	JX508291	6337	5870 (92.6)	79	83.52	95.56	80.74	tymovirus	Andean potato mottle virus
Cus-083	JQ647830	8482	5993 (70.7)	30.9	89.09	100	83.33	carlavirus	Potato virus S
	KJ534601	6445	658 (10.2)	24.4	98.18	100	91.67	potexvirus	Potato virus X
	JQ971975	9714	9488 (97.7)	723.8	93.04	100	87.5	potyvirus	Potato virus Y
	KT218527	227	42 (18.5)	1.3	92.86	92.86	92.86	tospovirus	Tomato chlorotic spot virus
Cus-084	JQ647830	8482	7127 (84)	48.5	88	94.81	80	carlavirus	Potato virus S
	JQ971975	9714	9562 (98.4)	801.8	92.66	100	88.89	potyvirus	Potato virus Y
Cus-085	JQ647830	8482	5272 (62.2)	46.9	89.75	97.67	80.56	carlavirus	Potato virus S
	JQ971975	9714	9577 (98.6)	805.3	92.66	100	88.89	potyvirus	Potato virus Y
	KU586450	9600	5628 (58.6)	7.4	98.45	100	94.44	potyvirus	Potato virus A
Cus-086	JQ647830	8482	2296 (27.1)	6.6	92.54	100	84.28	carlavirus	Potato virus S
	KJ534601	6445	6445 (100)	94.8	92.82	100	77.56	potexvirus	Potato virus X
	JQ971975	9714	9506 (97.9)	13.2	93.59	100	88.24	potyvirus	Potato virus Y
	KU586450	9600	6181 (64.4)	1.7	98.58	100	92.68	potyvirus	Potato virus A

Cus-087	KX656670	7148	6274 (87.8)	27.1	90.44	98.25	83.87	nepovirus	Potato virus B
	AB714135	9703	9156 (94.4)	5.5	94.18	98	89.83	potyvirus	Potato virus Y
	AAD09329.2	1626	466 (28.7)	25.9	61.64	77.19	40.68	nepovirus	Blackcurrant reversion virus
	CBL43396.1	1499	442 (29.5)	30.3	58.61	72.41	48.15	nepovirus	Grapevine Bulgarian latent virus
	CBL43395.1	2095	1519 (72.5)	30.9	67.47	97.78	30.18	nepovirus	Grapevine Bulgarian latent virus
Cus-088	JQ647830	8482	2644 (31.2)	9.1	92.38	99.19	85.25	carlavirus	Potato virus S
	KJ534601	6445	6444 (100)	139.4	93.48	100	81.9	potexvirus	Potato virus X
	JQ971975	9714	9424 (97)	5.5	94.63	100	89.83	potyvirus	Potato virus Y
Cus-089	JQ647830	8482	5574 (65.7)	14.4	87.82	97.83	81.56	carlavirus	Potato virus S
	M72416	7568	7568 (100)	72.5	95.3	100	80.85	potexvirus	Potato virus X
	JQ971975	9714	9448 (97.3)	5	94.43	100	87.5	potyvirus	Potato virus Y
Cus-090	KJ534601	6445	2234 (34.7)	3.3	97.9	100	95	potexvirus	Potato virus X
	JN936418	9479	7958 (84)	3.3	96.31	100	88.46	potyvirus	Potato virus Y
	KP849483	9859	5870 (59.5)	1.5	98.42	100	95.16	potyvirus	Potato virus V
Cus-090B	KX656670	7148	6033 (84.4)	26.3	92.19	98.46	83.95	nepovirus	Potato virus B
	JQ971975	9714	9278 (95.5)	4.7	94.55	100	89.05	potyvirus	Potato virus Y
Cus-091	M84806	2303	2109 (91.6)	38.9	83.67	97.62	81.2	comovirus	Andean potato mottle virus
	KJ534601	6445	1884 (29.2)	2.4	98.25	100	94.74	potexvirus	Potato virus X
	KJ634024	9724	9724 (100)	223.9	96.16	100	86.27	potyvirus	Potato virus Y
	AAA42421.1	997	976 (97.9)	57.5	72.78	90.91	62.93	comovirus	Andean potato mottle virus
	CAA46104.1	1864	976 (52.4)	84.6	38.65	43.26	27.53	comovirus	Red Clover Mottle Virus
Cus-092	KR270797	9726	9346 (96.1)	10.7	94.46	100	83.44	potyvirus	Potato virus Y
	AFA36173.1	290	39 (13.4)	4.1	69.23	70.59	68.18	varicosavirus	Lettuce big-vein virus
	BAC16226.1	2040	221 (10.8)	6.1	79.37	100	65.79	varicosavirus	Lettuce big-vein virus
Cus-093	M84806	2303	2176 (94.5)	76.1	83.29	89.13	81.24	comovirus	Andean potato mottle virus

	KJ634024	9724	9724 (100)	623	95.13	100	86.27	potyvirus	Potato virus Y
	ACX53646.1	1860	981 (52.7)	161.7	41.92	68.18	34.34	comovirus	Turnip ringspot virus
	AAA42421.1	997	970 (97.3)	119.5	72.76	77.63	68	comovirus	Andean potato mottle virus
	AIG59705.1	186	25 (13.4)	35.5	60	60	60	comovirus	Andean potato mottle virus
Cus-094	JQ971975	9714	9476 (97.5)	11.4	93.55	100	88.35	potyvirus	Potato virus Y
Cus-095	KX656670	7148	3022 (42.3)	3.7	94.67	100	84.27	nepovirus	Potato virus B
	JQ971975	9714	9487 (97.7)	5.4	94.58	98	89.13	potyvirus	Potato virus Y
	AFA36171.1	397	84 (21.2)	6.5	63.1	69.57	55.56	varicosavirus	Lettuce big-vein virus
	BAC16226.1	2040	820 (40.2)	7.1	72.76	100	45.71	varicosavirus	Lettuce big-vein virus
	AFA36173.1	290	120 (41.4)	8	76.67	87.5	69.49	varicosavirus	Lettuce big-vein virus
Cus-096	KT160432	1065	542 (50.9)	33.5	89.93	90.45	88.83	ilarvirus	Fragaria chiloensis latent virus
	KX656671	4527	662 (14.6)	1.9	96.82	100	92.59	nepovirus	Potato virus B
	JQ971975	9714	9265 (95.4)	6.3	94.86	100	89.36	potyvirus	Potato virus Y
	AAF89747.1	1456	435 (29.9)	1.5	97.84	100	94.74	potexvirus	Potato virus X
Cus-097	AY707771	2570	892 (34.7)	32	85.9	94.44	82.66	ilarvirus	Fragaria chiloensis latent virus
	JQ971975	9714	9271 (95.4)	6.7	94.34	100	88.8	potyvirus	Potato virus Y
	KP849483	9859	9246 (93.8)	3	98.2	100	89.13	potyvirus	Potato virus V
Cus-098	KJ534601	6445	2260 (35.1)	7.3	98.32	100	93.85	potexvirus	Potato virus X
	KP849483	9859	9846 (99.9)	699.1	97.82	100	91.3	potyvirus	Potato virus V
	JQ971975	9714	9286 (95.6)	38.6	94.77	100	88.24	potyvirus	Potato virus Y
Cus-099	AY682102	3431	740 (21.6)	15	95.73	100	92.98	ilarvirus	Fragaria chiloensis latent virus
	KP849483	9859	9583 (97.2)	9.4	98.06	100	94.12	potyvirus	Potato virus V
	JQ971975	9714	7187 (74)	3.3	95.64	100	91.67	potyvirus	Potato virus Y
Cus-100	KM659859	6487	6175 (95.2)	2.5	95.87	100	87.61	potexvirus	Potato virus X
	JQ971975	9714	9310 (95.8)	3.8	94.7	100	89.23	potyvirus	Potato virus Y

Cus-100B	AF493951	949	522 (55)	22.7	90.63	96.72	86.51	carlavirus	Potato virus S
	M84483	2265	1583 (69.9)	66.2	84.29	96.83	80.52	comovirus	Andean potato mottle virus
	M72416	7568	6844 (90.4)	3.2	97.77	100	91.53	potexvirus	Potato virus X
	KJ534601	6445	3458 (53.7)	2.8	97.86	100	93.75	potexvirus	Potato virus X
	ACV41391.1	1860	1006 (54.1)	141	39.03	43.07	34.25	comovirus	Turnip ringspot virus
Cus-101	KJ534601	6445	660 (10.2)	6.8	98.64	100	96.08	potexvirus	Potato virus X
	KJ634024	9724	9528 (98)	513.8	90.51	100	83.18	potyvirus	Potato virus Y
Cus-102	JQ971975	9714	9055 (93.2)	3.6	95.34	100	88.89	potyvirus	Potato virus Y
Cus-103	JQ971975	9714	8937 (92)	4.1	95.07	100	88.39	potyvirus	Potato virus Y
	AOE46790.1	1231	810 (65.8)	35.2	57.66	74.19	38.86	enamovirus	Grapevine enamovirus
Cus-104	FJ643622	885	108 (12.2)	9.8	98.15	100	97.1	carlavirus	Potato virus S
	KJ634024	9724	9589 (98.6)	766.8	92.47	98.68	86.72	potyvirus	Potato virus Y
Cus-105	M72416	7568	7029 (92.9)	8.4	97.49	100	88.73	potexvirus	Potato virus X
	JQ971975	9714	9560 (98.4)	733.1	92.51	100	87.67	potyvirus	Potato virus Y
Cus-106	M72416	7568	6881 (90.9)	8.3	97.67	100	95.12	potexvirus	Potato virus X
	JQ971975	9714	9639 (99.2)	761.1	92.66	100	86.51	potyvirus	Potato virus Y
Cus-107	M72416	7568	6938 (91.7)	2.2	97.76	100	90.77	potexvirus	Potato virus X
	JQ971975	9714	9403 (96.8)	6.1	93.95	100	87.9	potyvirus	Potato virus Y
Cus-108	HE817771	276	42 (15.2)	86.5	90.48	90.48	90.48	NA	Maize red stripe virus
	AF149101	313	47 (15)	1.6	91.67	91.67	91.67	ophiovirus	Citrus psorosis virus
	M72416	7568	7008 (92.6)	10.8	97.32	100	88	potexvirus	Potato virus X
	JQ971975	9714	9509 (97.9)	711.4	93.07	100	86.4	potyvirus	Potato virus Y
Cus-109	KJ534601	6445	2837 (44)	8	97.85	100	92.55	potexvirus	Potato virus X
	KJ634024	9724	9638 (99.1)	904.6	90.83	100	87.85	potyvirus	Potato virus Y
Cus-110	M72416	7568	7168 (94.7)	9	97.75	100	92.5	potexvirus	Potato virus X
	JQ971975	9714	9609 (98.9)	718.7	92.74	97.92	85.71	potyvirus	Potato virus Y

Cus-110B	M72416	7568	6969 (92.1)	4.4	97.62	100	92.16	potexvirus	Potato virus X
	JQ971975	9714	9658 (99.4)	756.8	92.36	98.88	87.69	potyvirus	Potato virus Y
Cus-111	KJ534601	6445	1047 (16.2)	23.2	98.09	100	92.5	potexvirus	Potato virus X
	JQ971975	9714	9570 (98.5)	843.1	92.35	100	85.62	potyvirus	Potato virus Y
Cus-112	KJ534601	6445	1321 (20.5)	25.6	98.86	100	96.08	potexvirus	Potato virus X
	JQ971975	9714	9487 (97.7)	768.4	92.65	100	86.85	potyvirus	Potato virus Y
	AHA91817.1	659	70 (10.6)	25	68.57	73.17	62.07	torradovirus	Cassava Torrado-like virus
	ABV44417.1	190	69 (36.3)	24.2	40	40	40	torradovirus	Tomato marchitez virus
	AMN91910.1	2207	336 (15.2)	27.6	51.01	90.91	27.61	torradovirus	Squash chlorotic leaf spot virus
Cus-113	KJ534601	6445	1230 (19.1)	3.9	98.46	100	95.12	potexvirus	Potato virus X
	JQ971975	9714	9586 (98.7)	734	92.86	97.92	84.09	potyvirus	Potato virus Y
	ACT79984.1	1197	134 (11.2)	41	34.33	45.65	28.41	NA	Tomato chocolate spot virus
	AMN91910.1	2207	237 (10.7)	26.8	57.32	100	36	torradovirus	Squash chlorotic leaf spot virus
	AHA91818.1	973	234 (24)	40.9	45.3	50.79	41.67	torradovirus	Cassava Torrado-like virus
	AMN91912.1	1054	207 (19.6)	42	41.15	46.94	35	torradovirus	Squash chlorotic leaf spot virus
	AIT59086.1	205	110 (53.7)	54.5	38.18	38.18	38.18	torradovirus	Motherwort yellow mottle virus
Cus-114	KX656671	4527	2359 (52.1)	37.5	91.86	100	85.47	nepovirus	Potato virus B
	KJ534601	6445	1991 (30.9)	8.1	98.54	100	94.29	potexvirus	Potato virus X
	JQ971975	9714	9533 (98.1)	630.7	92.58	100	88.64	potyvirus	Potato virus Y
Cus-115	KX656671	4527	2889 (63.8)	53.1	90.43	98.28	85.16	nepovirus	Potato virus B
	JQ971975	9714	9451 (97.3)	722.4	93	100	90.48	potyvirus	Potato virus Y
	ACT79984.1	1197	131 (10.9)	30.8	41.22	57.69	30	NA	Tomato chocolate spot virus
	AHA91818.1	973	253 (26)	39.7	47.27	55.41	33.33	torradovirus	Cassava Torrado-like virus
Cus-116	M72416	7568	7126 (94.2)	3.9	97.3	100	89.89	potexvirus	Potato virus X
	JQ971975	9714	9643 (99.3)	479.5	92.87	100	86.03	potyvirus	Potato virus Y

Cus-117	M72416	7568	7226 (95.5)	2	97.7	100	92.77	potexvirus	Potato virus X
	JQ971975	9714	9699 (99.8)	7.9	94.63	100	86.96	potyvirus	Potato virus Y
	AMN91912.1	1054	685 (65)	34.4	41.07	62.5	30	torradovirus	Squash chlorotic leaf spot virus
	ACB47565.1	1198	192 (16)	36.4	50	56.14	42.86	torradovirus	Tomato torrado virus
	ACA49852.1	1074	176 (16.4)	29.1	49.43	56.14	42.86	torradovirus	Tomato torrado virus
	AMM63581.1	190	157 (82.6)	46.6	31.45	31.45	31.45	torradovirus	Tomato marchitez virus
	AMN91910.1	2207	1266 (57.4)	20.8	52.52	87.18	30	torradovirus	Squash chlorotic leaf spot virus
Cus-118	AY682102	3431	2408 (70.2)	40.5	85.61	95.24	80.33	ilarvirus	Fragaria chiloensis latent virus
	M72416	7568	7193 (95)	3.9	97.65	100	93.59	potexvirus	Potato virus X
	JQ971975	9714	9714 (100)	649	95.42	100	89.13	potyvirus	Potato virus Y
	AAV40823.1	314	303 (96.5)	186.1	71.02	85.51	63.11	ilarvirus	Fragaria chiloensis latent virus
	AIT59087.1	1195	632 (52.9)	29.1	40.28	50	33.69	torradovirus	Motherwort yellow mottle virus
	AHA85557.1	1167	577 (49.4)	25.8	42	70.83	33.33	torradovirus	Carrot torradovirus
	AMN91910.1	2207	498 (22.6)	20.8	52.61	93.1	23.61	torradovirus	Squash chlorotic leaf spot virus
	AHA91818.1	973	343 (35.3)	30.2	46.97	51.54	40.71	torradovirus	Cassava Torrado-like virus
Cus-119	M72416	7568	7242 (95.7)	8.3	97.37	100	88.54	potexvirus	Potato virus X
	KR270797	9726	9392 (96.6)	754.2	93.99	100	87.76	potyvirus	Potato virus Y
	AIT59086.1	205	73 (35.6)	42.5	45.53	48	43.84	torradovirus	Motherwort yellow mottle virus
	ACB47565.1	1198	502 (41.9)	30	39.19	59.38	27.88	torradovirus	Tomato torrado virus
	AMN91910.1	2207	483 (21.9)	25.2	64.72	93.1	32.73	torradovirus	Squash chlorotic leaf spot virus
	AHA85557.1	1167	464 (39.8)	29.4	43.29	58.54	31.58	torradovirus	Carrot torradovirus
	AIT59087.1	1195	450 (37.7)	34.4	41.21	60	28.24	torradovirus	Motherwort yellow mottle virus
	ACU01026.1	1192	442 (37.1)	31.8	39.54	59.52	28.43	torradovirus	Tomato chocolate virus
	AMN91912.1	1054	407 (38.6)	31.1	45.39	70	36.25	torradovirus	Squash chlorotic leaf spot virus
	AHA91818.1	973	352 (36.2)	33.7	46.48	68.57	33.33	torradovirus	Cassava Torrado-like virus

	AHA85556.1	2192	273 (12.5)	25.7	71.89	89.47	41.67	torradovirus	Carrot torradovirus
Cus-120	M72416	7568	7306 (96.5)	3.6	97.44	100	90	potexvirus	Potato virus X
	KJ634024	9724	9671 (99.5)	766.8	92.75	100	86.97	potyvirus	Potato virus Y
	AOE46789.1	813	302 (37.1)	61.6	52.94	57.89	50	enamovirus	Grapevine enamovirus
	ABV44417.1	190	82 (43.2)	32.1	36.9	36.9	36.9	torradovirus	Tomato marchitez virus
	AMN91912.1	1054	307 (29.1)	31.7	41.1	48.24	32.86	torradovirus	Squash chlorotic leaf spot virus
	ACB47565.1	1198	213 (17.8)	27.5	47.2	61.11	39.44	torradovirus	Tomato torrado virus
	AHA91818.1	973	212 (21.8)	31.1	50.47	52.73	46.48	torradovirus	Cassava Torrado-like virus
Cus-120B	KX656671	4527	1779 (39.3)	15.2	95.24	98.53	90.91	nepovirus	Potato virus B
	M72416	7568	7525 (99.4)	10.6	97.39	100	88	potexvirus	Potato virus X
	KJ634024	9724	9697 (99.7)	980.7	92.87	100	86.72	potyvirus	Potato virus Y
Cus-120C	M72416	7568	7481 (98.9)	8.8	97.86	100	94.74	potexvirus	Potato virus X
	JQ971975	9714	9668 (99.5)	22.6	94.21	100	87.8	potyvirus	Potato virus Y
Cus-121	JQ647830	8482	3778 (44.5)	53.8	91.93	100	87.85	carlavirus	Potato virus S
	KX656671	4527	3121 (68.9)	27.3	91.29	100	85.71	nepovirus	Potato virus B
	JQ971975	9714	9047 (93.1)	8.5	94.1	100	89.19	potyvirus	Potato virus Y
Cus-122	KX656670	7148	3709 (51.9)	10.6	93.17	98.36	85.96	nepovirus	Potato virus B
	M72416	7568	7568 (100)	62.6	96.7	98.11	90.8	potexvirus	Potato virus X
	JQ971975	9714	9699 (99.8)	32.6	93.37	100	86.02	potyvirus	Potato virus Y
Cus-123	JQ971975	9714	8878 (91.4)	3.6	95.42	100	88.64	potyvirus	Potato virus Y
Cus-124	KT160431	713	74 (10.4)	2.6	97.3	97.3	97.3	ilarvirus	Fragaria chiloensis latent virus
	JQ971975	9714	9681 (99.7)	35.9	93.08	100	88.42	potyvirus	Potato virus Y
Cus-125	M72416	7568	7158 (94.6)	1.9	97.81	100	93.84	potexvirus	Potato virus X
	KR270797	9726	8919 (91.7)	4.7	94.9	100	86.9	potyvirus	Potato virus Y
Cus-126	KX656671	4527	4005 (88.5)	11.1	96.24	98.4	92.63	nepovirus	Potato virus B
	M72416	7568	6974 (92.2)	2	97.6	100	90	potexvirus	Potato virus X

	JQ971975	9714	8976 (92.4)	4.3	94.63	100	88.64	potyvirus	Potato virus Y
Cus-127	KJ534601	6445	6445 (100)	45.5	97.53	100	90.28	potexvirus	Potato virus X
Cus-128	M72416	7568	6730 (88.9)	3.2	97.62	100	90.38	potexvirus	Potato virus X
Cus-129	KJ534601	6445	5841 (90.6)	1.4	98.42	100	92	potexvirus	Potato virus X
Cus-130	M72416	7568	6866 (90.7)	2.2	97.52	100	91.07	potexvirus	Potato virus X
Cus-130B	KX656671	4527	826 (18.2)	5.8	98.06	100	93.75	nepovirus	Potato virus B
	KJ534601	6445	6445 (100)	277.6	90.42	97.18	84.91	potexvirus	Potato virus X
	KR270797	9726	1124 (11.6)	1.8	99.29	100	96.92	potyvirus	Potato virus Y

Anexo 6: Código de barras

Tabla 35: Código de barras utilizado para la preparación de librerías.

Primers PCR2_*	Secuencia de código de barra PCR2*-	Reversa del código de barra PCR2*_	Primers PCR2_*	Secuencia de código de barra PCR2*-	Reversa del código de barra PCR2*_
PCR2_1	ATCACG	CGTGAT	PCR2_25	ACTGAT	ATCAGT
PCR2_2	CGATGT	ACATCG	PCR2_26	ATGAGC	GCTCAT
PCR2_3	TTAGGC	GCCTAA	PCR2_27	ATTCCT	AGGAAT
PCR2_4	TGACCA	TGGTCA	PCR2_28	CAAAAG	CTTTTG
PCR2_5	ACAGTG	CACTGT	PCR2_29	CAACTA	TAGTTG
PCR2_6	GCCAAT	ATTGGC	PCR2_30	CACCGG	CCGGTG
PCR2_7	CAGATC	GATCTG	PCR2_31	CACGAT	ATCGTG
PCR2_8	ACTTGA	TCAAGT	PCR2_32	CACTCA	TGAGTG
PCR2_9	GATCAG	CTGATC	PCR2_33	CAGGCG	CGCCTG
PCR2_10	TAGCTT	AAGCTA	PCR2_34	CATGGC	GCCATG
PCR2_11	GGCTAC	GTAGCC	PCR2_35	CATTTT	AAAATG
PCR2_12	CTTGTA	TACAAG	PCR2_36	CCAACA	TGTTGG
PCR2_13	AGTCAA	TTGACT	PCR2_37	CGGAAT	ATTCCG
PCR2_14	AGTTCC	GGAACT	PCR2_38	CTAGCT	AGCTAG
PCR2_15	ATGTCA	TGACAT	PCR2_39	CTATAC	GTATAG
PCR2_16	CCGTCC	GGACGG	PCR2_40	CTCAGA	TCTGAG
PCR2_17	GTAGAG	CTCTAC	PCR2_41	GACGAC	GTCGTC
PCR2_18	GTCCGC	GCGGAC	PCR2_42	TAATCG	CGATTA
PCR2_19	GTGAAA	TTTCAC	PCR2_43	TACAGC	GCTGTA
PCR2_20	GTGGCC	GGCCAC	PCR2_44	TATAAT	ATTATA
PCR2_21	GTTTCG	CGAAAC	PCR2_45	TCATTC	GAATGA
PCR2_22	CGTACG	CGTACG	PCR2_46	TCCCGA	TCGGGA
PCR2_23	GAGTGG	CCACTC	PCR2_47	TCGAAG	CTTCGA
PCR2_24	GGTAGC	GCTACC	PCR2_48	TCGGCA	TGCCGA