



## **Máster en Bioinformática**

**Reconstrucción del Genoma e  
Identificación Taxonómica de *Avipoxvirus*  
*sp.* en lesiones cutáneas de curruca (*Sylvia  
atricapilla*) mediante Secuenciación con  
Oxford Nanopore**

**Autor: Blga. Luisa Daniela Espinosa Delgado**

**Tutor: PhD. Fernando Esperón Fajardo**

**Curso 2023-24**

## **AGRADECIMIENTOS**

"En primer lugar, quiero expresar mi más sincero agradecimiento al PhD. Fernando Esperón, por compartir su conocimiento, así como por su constante apoyo y profesionalismo que han sido pilares fundamentales en esta investigación. Su guía constante y confianza en mis capacidades me han impulsado a alcanzar logros que nunca imaginé posibles.

También extendo mi agradecimiento al Mgtr. Manuel Remensal, por su valiosa orientación en el trabajo bioinformático, por compartir sus conocimientos y por dedicar tiempo al desarrollo de este trabajo.

A las personas que forman parte del Hospital Clínico Veterinario de la Universidad Europea de Madrid, especialmente a Elena, Bárbara, Fátima, Verónica y Laura, les agradezco profundamente por compartir su experiencia en distintas áreas y por acompañarme durante el trabajo de laboratorio. Su apoyo, confianza, respaldo y cariño han sido invaluable.

Finalmente, quiero agradecer a mi madre Jenny, a mi compañero de vida Mauricio y a toda mi familia, cuya fe inquebrantable en mí, incluso en los momentos más difíciles, ha sido el pilar de este logro. En especial, mi más sincero reconocimiento a Jorge, Gianna y Mely, quienes se convirtieron en mi segundo hogar durante mi estancia en España. Gracias por acompañarme en esta etapa, por escucharme y por brindarme su apoyo incondicional. Sin ustedes, todo esto no habría sido posible."

# ÍNDICE DE CONTENIDO

CARÁTULA.....	1
AGRADECIMIENTOS.....	2
ÍNDICE DE CONTENIDO.....	3
GLOSARIO DE TÉRMINOS Y ABREVIATURAS.....	5
RESUMEN.....	7
ABSTRACT.....	8
1. INTRODUCCIÓN.....	9
2. HIPÓTESIS Y OBJETIVOS.....	10
3. METODOLOGÍA.....	11
3.1 MATERIAL BIOLÓGICO.....	11
3.2 PREPARACIÓN DE LA MUESTRA.....	11
3.3 EXTRACCIÓN Y AMPLIFICACIÓN DE ADN.....	12
3.4 PREPARACIÓN DE LA LIBRERÍA Y SECUENCIACIÓN.....	12
3.5 ANÁLISIS BIOINFORMÁTICO.....	12
3.5.1 DISEÑO INICIAL DEL PIPELINE.....	13
3.5.2 AJUSTES DEL PIPELINE.....	13
3.5.3 PIPELINE IMPLEMENTADO.....	14
3.6 ANÁLISIS FILOGENÉTICO.....	15
4. RESULTADOS.....	16
4.1 ANÁLISIS DEL PROTOCOLO DE LABORATORIO.....	16
4.2 DATOS OBTENIDOS DE LA SECUENCIACIÓN.....	17
4.3 ANÁLISIS DEL PIPELINE.....	17
4.4 PREPROCESAMIENTO.....	19
4.5 ENSAMBLAJE DEL GENOMA.....	22
4.6 ANÁLISIS FILOGENÉTICO.....	23
5 DISCUSIÓN.....	25
6 CONCLUSIONES.....	28
7 BIBLIOGRAFÍA.....	29

8 ANEXOS .....	36
ANEXO 1.....	36
ANEXO 2.....	37
ANEXO 3.....	37
9 DECLARACIÓN OBLIGATORIA DEL USO DE HERRAMIENTAS DE IA.....	37

## GLOSARIO DE TÉRMINOS Y ABREVIATURAS

**ADN:** Ácido desoxirribonucleico.

**APV:** Avipoxvirus (nombre general para los virus del género *Avipoxvirus*).

**BLAST:** Herramienta básica de búsqueda de alineación local (Basic Local Alignment Search Tool).

**ChPV:** Subfamilia *Chordopoxvirinae*.

**Contigs:** Conjunto de segmentos o secuencias de ADN que se superponen parcialmente para generar una representación continua de una región o de todo un genoma.

**CAM:** Membranas corioalantoideas.

**dsDNA:** ADN de doble cadena (double-stranded DNA).

**EPI2ME:** Plataforma de análisis en la nube proporcionada por Oxford Nanopore para procesar y analizar datos de secuenciación.

**Fastq:** Formato de archivo que almacena secuencias de ADN junto con sus calidades de lectura.

**kpb:** Kilo pares de bases, unidad de medida de longitud de secuencias de ADN equivalente a 1000 pares de bases.

**MinION:** Dispositivo de secuenciación portátil basado en la tecnología Oxford Nanopore.

**MAFFT:** Software de alineamiento múltiple de secuencias (Multiple Alignment using Fast Fourier Transform).

**MEGA:** Software para análisis evolutivo y filogenético (Molecular Evolutionary Genetics Analysis).

**NGS:** Secuenciación de Nueva Generación (Next-Generation Sequencing).

**NCBI:** Centro Nacional de Información Biotecnológica.

**ONT:** Tecnología Oxford Nanopore.

**POD5:** Formato de archivo para datos de secuenciación en tecnología ONT que reemplaza el formato FAST5.

**pb:** Pares de bases, la unidad de medida para la longitud de secuencias de ADN.

**PCR:** Reacción en cadena de la polimerasa.

**Ph:** índice Phred, medida de la calidad de las bases nucleotídicas en secuenciación de ADN.

**Pipeline:** Conjunto de herramientas, programas y código para procesar información compleja, generando un flujo de trabajo bioinformático para el manejo eficiente y reproducible.

**SMRT:** Secuenciación en tiempo real de molécula única (Single molecule real time).

**WGS:** Secuenciación del Genoma Completo (Whole Genome Sequencing).

## RESUMEN

Los avipoxvirus (APV) son patógenos de amplia distribución que afectan a la avifauna silvestre, cautiva y doméstica a nivel mundial. Se han identificado aproximadamente doce especies, sin embargo, la mayoría de avipoxvirus que infectan aves silvestres no han sido completamente caracterizados, y su diversidad genética permanece poco estudiada. En este contexto, la presente investigación desarrolló un protocolo optimizado para la secuenciación directa de lesiones causadas por viruela aviar, empleando la tecnología de Oxford Nanopore, y diseñó un flujo bioinformático (pipeline) que permitió la reconstrucción del genoma completo de un *Avipoxvirus* sp. extraído de curruca (*Sylvia atricapilla*).

La metodología propuesta logró una concentración inicial de ADN de 7.43 µg/ml, que incrementó a 139 µg/ml tras los procesos de amplificación y purificación. La secuenciación generó un total de 627.46 Mb, correspondientes a 14.230 lecturas. El flujo bioinformático incluyó ocho etapas: análisis de calidad, eliminación de adaptadores, filtrado de secuencias de baja calidad, identificación taxonómica, selección de secuencias de interés, ensamblaje, pulido y visualización. Para el ensamblaje genómico, se empleó un enfoque de reensamblaje basado en el genoma de referencia de *Canarypox virus*, obteniendo una secuencia de 351.122 pb. A partir de ello, se obtuvo una región contigua de 249.264 pb que fue utilizada en el análisis filogenético. Esta secuencia mostró una similitud del 98.88% con respecto al *Canarypox virus*, clasificándola dentro del clado B. La secuencia estudiada ha sido denominada como *Warblerpox virus*.

Los resultados obtenidos demuestran que tanto el protocolo de laboratorio como el flujo bioinformático diseñado son eficientes y viables, facilitando la reconstrucción del genoma. Esta aportación es de gran relevancia para la vigilancia, el control y la identificación taxonómica de nuevos patógenos aviarios.

### Palabras clave:

APV, análisis bioinformático, genoma viral, Oxford Nanopore MinION, WGS (Whole-Genome Sequencing).

## ABSTRACT

Avipoxviruses (APV) are widely distributed pathogens that affect wild, captive, and domestic avifauna globally. Approximately twelve species have been identified; however, most avipoxviruses infecting wild birds have not been fully characterized, and their genetic diversity remains underexplored. In this context, the present study developed an optimized protocol for the direct sequencing of lesions caused by avian pox using Oxford Nanopore technology and designed a bioinformatic pipeline that enabled the reconstruction of the complete genome of an *Avipoxvirus* sp. isolated from a warbler (*Sylvia atricapilla*).

The proposed methodology achieved an initial DNA concentration of 7.43 µg/ml, which increased to 139 µg/ml following amplification and purification processes. Sequencing generated a total of 627.46 Mb, corresponding to 14,230 reads. The bioinformatic pipeline included eight key stages: quality analysis, adapter removal, filtering of low-quality sequences, taxonomic identification, selection of sequences of interest, assembly, polishing, and visualization. For genomic assembly, a reassembly approach based on the reference genome of *Canarypox virus* was employed, resulting in a sequence of 351,122 bp. From this, a contiguous region of 249,264 bp was obtained and utilized in the phylogenetic analysis. This sequence exhibited a similarity of 98.88% to the *Canarypox virus*, classifying it within clade B. The sequence studied has been designated *Warblerpox virus*.

The results indicate that both the protocol for the laboratory protocol and the designed bioinformatic pipeline are efficient and viable, facilitating genome reconstruction. This contribution is highly relevant for the surveillance, control, and taxonomic identification of emerging avian pathogens.

### **Palabras clave:**

APV, bioinformatic analysis, Oxford Nanopore MinION, viral genome, WGS (Whole-Genome Sequencing).

## 1. Introducción

El agente viral responsable de la viruela aviar pertenece al grupo de los Avipoxvirus (APV) clasificados en la subfamilia *Chordopoxvirinae* (ChPV), de la familia *Poxviridae* (Giotis & Skinner, 2019). Se consideran patógenos relevantes de avifauna silvestre, cautiva y doméstica a nivel mundial, afectando a más de 300 especies de aves entre las que destacan principalmente pollos, pavos, palomas y rapaces (Giotis & Skinner, 2019; Liu et al., 2024). La transmisión de estos virus ocurre mecánicamente por insectos o por contacto directo, se propagan lentamente e inducen lesiones cutáneas con crecimientos proliferativos en las partes sin plumas o lesiones diftéricas asociadas con necrosis en las vías respiratorias superiores y el tracto digestivo (Lebdah et al., 2019; Net et al., 2020).

A nivel molecular, los Avipoxvirus poseen un genoma complejo de ADN de doble cadena (dsDNA), con un tamaño que oscila entre los 130 a 360 kilo pares de bases (kpb). Su genoma incluye aproximadamente 260 marcos de lectura abiertos y 65 genes homólogos conservados entre los *Chordopoxvirinae* (Asif et al., 2021; Croville et al., 2018; Deng et al., 2023; Lebdah et al., 2019; Liu et al., 2024). Además, se caracterizan por una alta frecuencia de recombinación y mutación, lo que resulta en la presencia de repeticiones, deleciones o inserciones a lo largo del genoma (Sasani et al., 2018).

En la actualidad, el diagnóstico y vigilancia de la viruela aviar se basa en la identificación histopatológica, inmunotransferencia, el aislamiento del virus en membrana corioalantoidea (CAM), análisis de enzimas de restricción y PCR dirigida a uno o varios genes virales como: P4b y la ADN polimerasa (Asif et al., 2021; Croville et al., 2018; Lebdah et al., 2019). Por otra parte, la secuenciación del genoma completo (WGS) y la secuenciación de nueva generación (NGS), surgen como una alternativa moderna para la identificación de posibles modificaciones genéticas *in situ* ante la evolución del virus. Tal es el caso de la secuenciación de lecturas largas, por medio de la tecnología de Oxford Nanopore (ONT), misma que permite un análisis de alta resolución para el estudio del genoma viral, permitiendo la secuenciación en tiempo real de molécula única (SMRT) y la identificación precisa de variantes (Brinkmann et al., 2024; Croville et al., 2018; Fixsen et al., 2020; Sasani et al., 2018; Saud et al., 2021).

Dentro del género *Avipoxvirus*, se han identificado 12 especies entre las que destacan: *Canarypox*, *Flamingopox*, *Fowlpox*, *Juncopox*, *Mynahpox*, *Pigeonpox*, *Penguinpox*, *Psittacinepox*, *Quailpox*, *Sparrowpox*, *Starlingpox* y *Turkeypox* (Sarker, Athukorala, & Raidal, 2021; Walker et al., 2020). Estas se encuentran clasificadas en tres clados: el clado A, asociado con *Fowlpox virus* y subdividido en A1 a A7; el clado B relacionado con *Canarypox virus*, subdividido en B1, B2 y B3; y el clado C correspondiente a *Psittacinepox virus*. Además existen nuevos reportes que sugieren la inclusión del clado D, relacionado con la viruela de codornices japonesas (*Coturnix japonica*), y el clado E asociado con pavos (*Meleagris gallopavo*) y pollos (*Gallus gallus domesticus*) (Chacón et al., 2022; Ribeiro et al., 2020). No obstante, la mayoría de *Avipoxvirus* que infectan aves silvestres aún no han sido caracterizados, y su diversidad genética sigue sin estar clara. Por lo tanto, la falta de conocimiento sobre la estructura filogenética del género, dificulta una evaluación precisa del riesgo que representa este patógeno (He et al., 2023; Lebdah et al., 2019; Sarker, Athukorala, & Raidal, 2021).

Ante ello, la presente investigación se enfocó en el desarrollo de un protocolo optimizado de laboratorio, para secuenciar directamente lesiones de viruela aviar, utilizando la tecnología de secuenciación Oxford Nanopore. Así mismo, buscó generar un flujo bioinformático para reconstruir el genoma completo de *Avipoxvirus* sp., con el fin de identificar la especie presente en currucas silvestres (*Sylvia atricapilla*), esperando que este flujo de trabajo sea aplicable para el control y vigilancia de este patógeno en distintas especies de aves.

## 2. Hipótesis y Objetivos

### Hipótesis

- Oxford Nanopore es un método efectivo para la secuenciación del genoma a partir de lesiones, permitiendo obtener la secuencia completa del virus, para su posterior reconstrucción e identificación taxonómica.
- La secuencia de *Avipoxvirus* reconstruida presenta una identidad inferior al 97% con respecto a la secuencia de referencia, lo que la convierte en una especie nueva.

### Objetivo primario

- Reconstruir e identificar la secuencia de *Avipoxvirus sp.* originario de una lesión cutánea de curruca (*Sylvia atricapilla*).

### Objetivos secundarios

- Diseñar un protocolo de laboratorio para la secuenciación completa del genoma viral, a partir de lesiones cutáneas.
- Desarrollar un pipeline para el procesamiento de datos y reconstrucción del genoma completo viral, a partir de fragmentos largos obtenidos mediante tecnología Oxford Nanopore.
- Identificar taxonómicamente la especie de *Avipoxvirus* y determinar si se trata de una especie conocida o posiblemente nueva.

## 3. Metodología

### 3.1 Material biológico

El material biológico empleado en este estudio proviene del Laboratorio de Biología Molecular del Hospital Clínico Veterinario de la Universidad Europea de Madrid (Tabla 1), mismo que se encontraba almacenado a -80 °C.

**Tabla 1**

*Información de la muestra*

<b>Código de la muestra</b>	<b>Tipo</b>	<b>Método de conservación</b>	<b>Organismo</b>
22324	Costra	Tubo Eppendorf	<i>Sylvia atricapilla</i>

Nota. Muestra empleada para la identificación de *Avipoxvirus sp.*

### 3.2 Preparación de la muestra

En un tubo estéril se colocó 0.2 gramos de la costra con 100 µl de buffer de lisis, 800 µl de agua destilada autoclavada (H<sub>2</sub>O), y perlas de cerámica, para la trituración de la muestra en el molino Fisherbrand™. Seguidamente, se extrajo el sobrenadante y se

centrifugó a 2500 rpm durante 15 minutos. A continuación, se filtró mediante una membrana de 0.22  $\mu\text{m}$  y el filtrado se utilizó para la extracción de ADN viral.

### **3.3 Extracción y Amplificación de ADN**

Para extraer el ADN genómico se empleó el QuickGene DNA tissue kit (Fujifilm Life Science, Japón) siguiendo las indicaciones del fabricante. Posteriormente, se realizó una amplificación isotérmica del ADN empleando el Repli-G mini kit (Quiagen, Alemania) de acuerdo con el protocolo establecido. La reacción consistió en (50  $\mu\text{L}$ ): 10  $\mu\text{L}$  de agua libre de nucleasas, 29  $\mu\text{L}$  de REPLI-g Mini Reaction Buffer, 1  $\mu\text{L}$  REPLI-g Mini DNA Polymerase y 10  $\mu\text{L}$  de ADN. El producto resultante fue purificado mediante una adaptación del DNA Extration Kit (Biotools, España) (Anexo 1).

Finalmente, la concentración del ADN se determinó mediante el espectrofotómetro Genova Nano (Jenway™, Staffordshire, Reino Unido).

### **3.4 Preparación de la librería y secuenciación**

La preparación de la librería se realizó con el Ligation Sequencing Kit v14 (SQK-LSK114, Oxford Nanopore Technologies) y el módulo complementario NEBNext® (*New England Biolabs, EE.UU*), siguiendo los protocolos estándar sugeridos por el fabricante.

La librería fue secuenciada en una celda de flujo R10.4.1 (FLO-MIN114, ONT) en el dispositivo MinION Mk1C (ONT, Reino Unido) durante 48 h, mediante el programa MinKnow versión 24.02.16, Dorado versión 7.3.11 y el modelo high accuracy basecalling; el resto de parámetros se dejaron por defecto.

Para transformar las lecturas POD5 se utilizó MinKNOW Core 5.9.12, que permitió realizar la llamada de base (basecalling) para obtener lecturas en formato fastq estándar.

### **3.5 Análisis bioinformático**

El análisis bioinformático se realizó en un sistema operativo GNU/Linux a través del terminal, utilizando herramientas instaladas localmente en el sistema o mediante Conda

(<https://www.anaconda.com/>) y el paquete BioConda (Grüning et al., 2018). Además, se empleó el lenguaje de programación Python versión 3 (Schacksen et al., 2024).

Se diseñó tres pipelines, los cuales fueron adaptados en función de los datos, los objetivos de la investigación y la capacidad computacional disponible. El pipeline A fue la base para el pipeline B, mismo que se optimizó para desarrollar el pipeline C. El código de los pipelines, la documentación y los repotes se encuentran disponibles en el Anexo 2 y 3.

El flujo bioinformático incluyó ocho etapas principales: análisis de calidad, eliminación de adaptadores, filtrado de secuencias de baja calidad, identificación taxonómica, selección de secuencias de interés, ensamblaje, pulido y visualización (Figura 1).

### 3.5.1 Diseño inicial del pipeline

Las lecturas crudas se analizaron en EPI2ME Agent versión 3.7.3, aplicando los protocolos: Fastq WIMP (What's in my pot) versión 2023.06.13-1865548 y Fastq Custom Aligment versión 2023.06.14-1866743. Posteriormente, se realizó un análisis de calidad con NanoPlot versión 1.43.0 (<https://github.com/wdecoster/NanoPlot>), se eliminaron adaptadores con Porechop versión 0.2.4 ([www.github.com/rrwick/Porechop](http://www.github.com/rrwick/Porechop)) y las lecturas largas de alta calidad se filtraron con Chopper versión 0.8.0 (<https://github.com/wdecoster/chopper?tab=readme-ov-file>) (Schacksen et al., 2024). Finalmente, las lecturas se ensamblaron de *novo* usando Flye versión 2.8 (<https://github.com/mikolmogorov/Flye>) (Baker, 2022; Hu et al., 2021) (Figura 1,A).

### 3.5.2 Ajustes del pipeline

Se desplazó el análisis Fastq WIMP (What's in my pot) de EPI2ME Agent versión 3.7.3, antes del ensamblaje. Este paso permitió filtrar las secuencias víricas en base a su readID. Las lecturas se ensamblaron con Flye versión 2.8 (<https://github.com/mikolmogorov/Flye>) y los contigs obtenidos se analizaron con Quast versión 5.2.0 (<https://github.com/ablab/quast>). El pulido del ensamblaje se llevó a cabo con Minimap2 versión 2.28-r1209 y Racon versión 1.5.0 (<https://github.com/isovic/racon>), repitiendo el proceso por triplicado o hasta

estabilizar el número de contigs. Finalmente, se utilizó Medaka versión 1.9.1 (<https://github.com/nanoporetech/medaka>) y la secuencia final fue verificada nuevamente con Quast y visualizada en IGV (Saud et al., 2021) (Figura 1,B).

### 3.5.3 Pipeline implementado

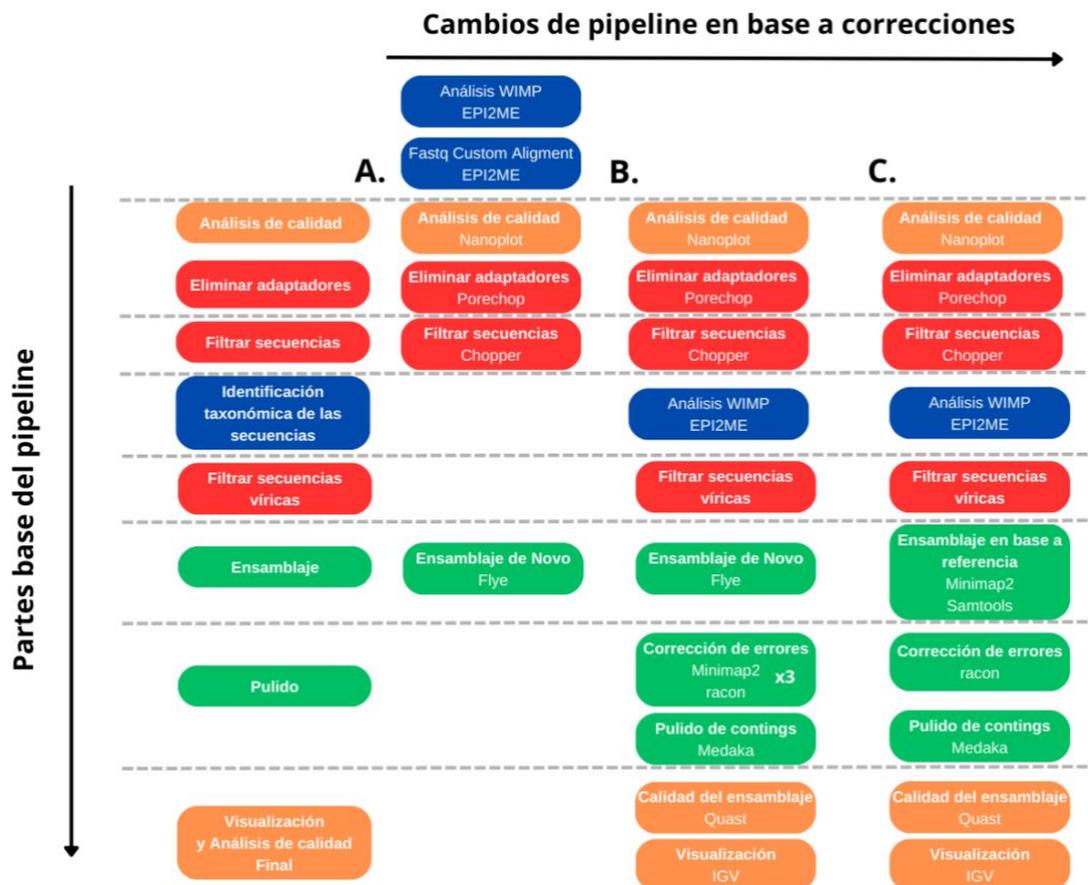
En base a la estructura de los pipelines anteriores, se empleó NanoPlot versión 1.43.0 (<https://github.com/wdecoster/NanoPlot>) para evaluar el nivel de calidad de la secuenciación y se eliminó los adaptadores con Porechop versión 0.2.4 ([www.github.com/rrwick/Porechop](http://www.github.com/rrwick/Porechop)) (Saud et al., 2021). Además, se empleó Chopper versión 0.8.0 (<https://github.com/wdecoster/chopper?tab=readme-ov-file>) para filtrar y mantener lecturas largas de alta calidad bajo los parámetros de longitud mínima >1000 pb y valor de calidad Ph >8 .

Posteriormente, las secuencias fueron cargadas en EPI2ME Agent Versión 3.7.3 aplicando el protocolo Fastq WIMP (What's in my pot) versión 2023.06.13-1865548, para clasificar y extraer las secuencias virales. A diferencia del pipeline B (Figura 1), el ensamblaje se realizó por mapeo, utilizando como referencia la secuencia de *Canarypox virus* (NC\_005309.1). Las lecturas fueron alineadas utilizando Minimap2 versión 2.28-r1209, y los alineamientos obtenidos fueron ordenados, indexados y filtrados usando Samtools versión 1.18 (Khrenova et al., 2022; Riaz et al., 2021).

El ensamblaje fue refinado mediante la herramienta Racon versión 1.5.0 (<https://github.com/isovic/racon>) y Medaka versión 1.9.1 (<https://github.com/nanoporetech/medaka>) (Schacksen et al., 2024). Finalmente, para verificar la secuencia obtenida se empleó Quast versión 5.2.0 (<https://github.com/ablab/quast>) y se visualizó en IGV (Figura 1,C) (Khrenova et al., 2022).

**Figura 1**

Diseño del pipeline bioinformático para el análisis de secuencias y reconstrucción del genoma de *Avipoxvirus sp.*



Nota. En la parte izquierda se muestra los pasos base del pipeline. (A) Pipeline inicial, (B) Pipeline ajustado y (C) Pipeline implementado. El color naranja indica controles de calidad y visualización, el rojo señala modificaciones como la eliminación de segmentos o secuencias completas, el azul corresponde a la identificación taxonómica de las secuencias y el verde representa el ensamblaje y pulido de la secuencia.

### 3.6 Análisis Filogenético

Se realizó el alineamiento entre la secuencia reconstruida y las secuencias de *Albatrosspox virus*, *Canarypox virus*, *Cheloniidpox virus* y *Shearwaterpox virus*, disponibles en la base de datos del Centro Nacional de Información Biotecnológica (NCBI) (Tabla 2). Para ello, se utilizó MAFFT versión 7.525 con la opción `-auto`,

permitiendo que el software detecte automáticamente el algoritmo más adecuado de alineamiento. El resultado fue visualizado en MEGA versión 11.0.11, con el fin de verificar y corregir posibles errores de secuenciación por medio de un ajuste manual.

El análisis filogenético se llevó a cabo utilizando el fragmento de la secuencia corregida, a la cual se le designó el nombre de *Warblerpox virus*. La secuencia se alineó nuevamente con las secuencias de ocho poxvirus disponibles en la base de datos NCBI (Tabla 2). Finalmente, se calcularon las distancias-p y se construyó el árbol filogenético aplicando el algoritmo Neighbor-Joining, con 500 replicaciones (Bootstrap) para validar la robustez del árbol.

**Tabla 2**

*Información de las secuencias de poxvirus del NCBI*

<b>Número de accesoión</b>	<b>Nombre del virus</b>	<b>Tamaño</b>
MW365933.1	<i>Albatrosspox virus</i>	351909 pb
NC005309.1	<i>Canarypox virus</i>	359853 pb
MT799800.1	<i>Cheloniidpox virus</i>	343132 pb
KX857215.1	<i>Shearwaterpox virus</i>	351108 bp
NC036582.1	<i>Flamingopox virus</i>	293123 pb
MG702259.1	<i>Flowpox virus</i>	308827 pb
NC024446.1	<i>Penguinpox virus</i>	306862 pb
NC028238.1	<i>Turkeypox virus</i>	188534 pb

Nota. Genomas empleados en el alineamiento y análisis filogenético

## 4. Resultados

### 4.1 Análisis del protocolo de laboratorio

En este estudio, se trabajó con una muestra de 0.2 g, utilizando volúmenes reducidos de reactivos para optimizar el manejo del material biológico. Para maximizar la concentración de ADN, se recomienda ajustar los volúmenes a 50 ul de buffer lisis y 400 ul de agua destilada, durante la preparación de la muestra.

Adicionalmente, se implementó un paso importante de filtrado para reducir la carga microbiana y eliminar restos de tejido del ave, minimizando la contaminación durante las etapas subsecuentes.

La extracción de ADN resultó en una concentración de 7.43 µg/ml, que se incrementó a 139 µg/ml tras los procesos de amplificación y purificación de ADN.

#### **4.2 Datos obtenidos de la secuenciación**

El secuenciador registró un estimado de 627.46 Mb equivalentes a 148.230 lecturas con una longitud media de 3.56 kb. Sin embargo, solo 106.873 lecturas superaron el umbral de calidad, presentando un puntaje mínimo Phred de 8.

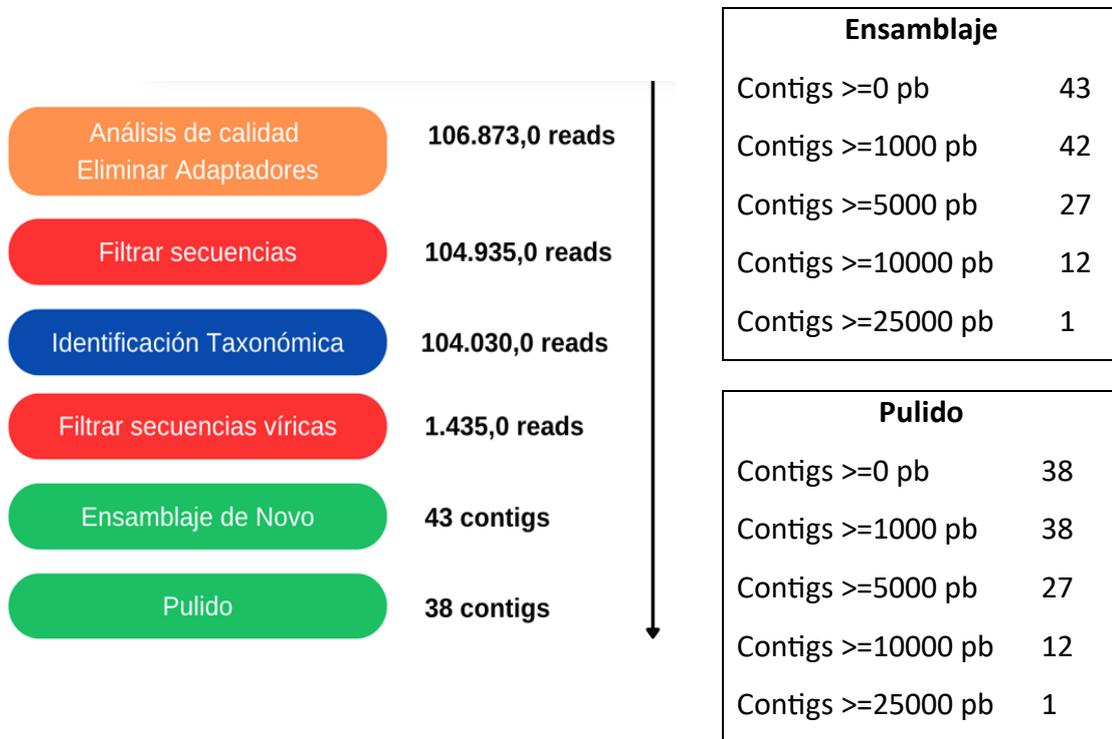
#### **4.3 Análisis del pipeline**

El pipeline inicial permitió establecer el análisis de calidad, la eliminación de adaptadores y el filtrado de secuencias. No obstante, el ensamblaje no fue viable debido al elevado volumen de datos contenidos en la muestra, superando los recursos computacionales disponibles. Para optimizar el proceso, se implementó la identificación taxonómica de EPI2ME como herramienta para seleccionar específicamente las secuencias de interés previo al ensamblaje.

Con el pipeline ajustado, se logró un ensamblaje de *novo*, obteniendo 43 contigs reducidos a 38 debido a la corrección de errores (Figura 2). Sin embargo, la visualización en IGV reveló una baja cobertura de estos, en comparación con el genoma de referencia del *Canarypox virus* (Figura 3).

**Figura 2**

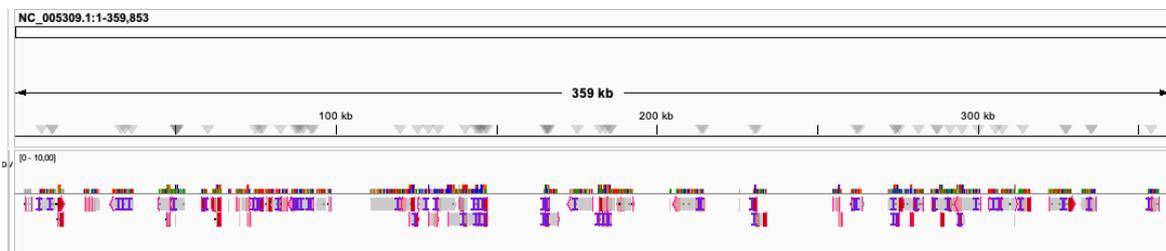
*Resultado del pipeline ajustado*



Nota. El diagrama muestra cómo el proceso de filtrado de secuencias víricas facilitó el ensamblaje de *novo*. Los resultados del análisis Quast, presentados en los cuadros de la derecha, muestran los tamaños de los contigs.

**Figura 3**

*Visualización de contigs en IGV*

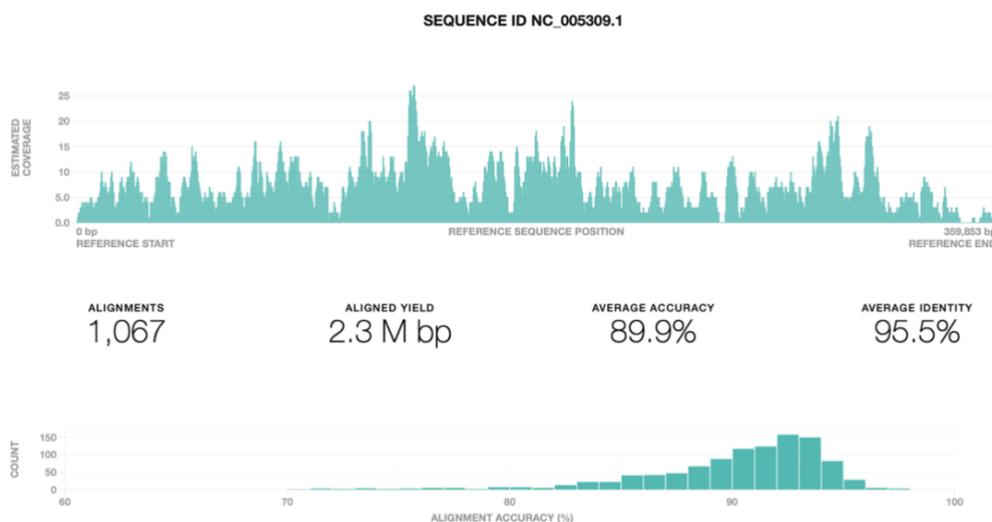


Nota. El siguiente gráfico muestra como los contigs del ensamblaje de *novo* no logran cubrir por completo el genoma de referencia de *Canarypox virus*.

A partir de este resultado, se propuso un cambio de estrategia para el ensamblaje del genoma, utilizando como referencia una secuencia de avipoxvirus. Esta decisión se basó en los resultados obtenidos con Fastq WIMP (What's in my pot) versión 2023.06.13-1865548 y Fastq Custom Aligment versión 2023.06.14-1866743 en el pipeline inicial, los cuales indicaron un 89.9% de precisión y un 95.5% de identidad con la secuencia de *Canaripox virus* (Figura 4). Adicionalmente, la corrección de errores se realizó una sola vez, en lugar de hacerlo por triplicado, con el objetivo de evitar que la secuencia se ajuste demasiado con la referencia.

**Figura 4**

*Reporte de EPI2ME, Fastq Custom Aligmen*



Nota. El siguiente gráfico indica los parámetros analizados por los cuales se seleccionó la secuencia de referencia.

A continuación, se presentan los resultados obtenidos en la reconstrucción del genoma empleando el pipeline establecido.

#### 4.4 Preprocesamiento

El análisis de calidad mediante Nanoplot proporcionó la información resumida en la Tabla 3. Además, indicó que la mayoría de las lecturas presentan longitudes entre 1.000 y 10.000 pb. El valor promedio de calidad (Q) fue de 9.6 y el 59.8% de las lecturas

alcanzaron una precisión de Q10, lo que sugiere que más de la mitad de las lecturas son de calidad adecuada para su uso.

**Tabla 3**

*Métricas del análisis de calidad en Nanoplot*

<b>Longitud media de lectura</b>	3.781.9
<b>Calidad media de lectura</b>	9.6
<b>Mediana de Longitud de lectura</b>	3.542.0
<b>Mediana de calidad de lectura</b>	10.3
<b>Longitud de lectura N50</b>	3.561
<b>Bases totales</b>	404.182.699.0

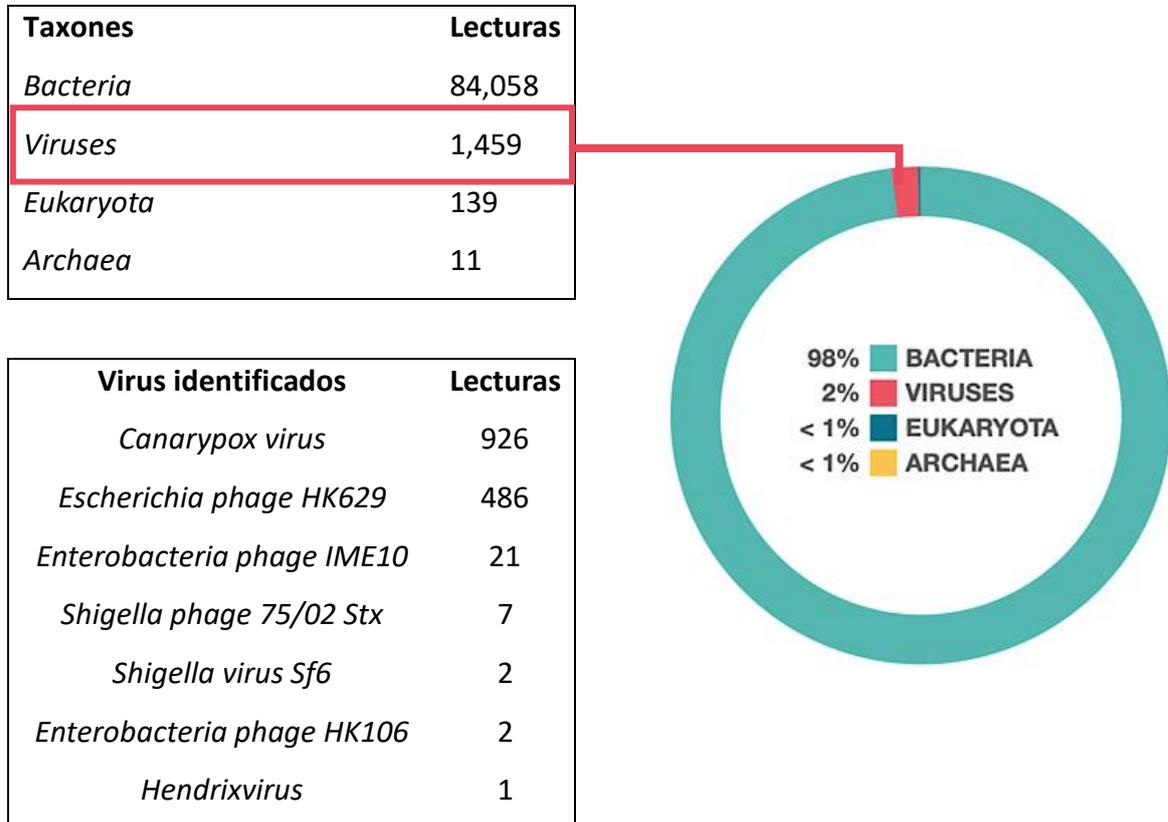
Nota. Resumen del análisis de calidad de las lecturas obtenido con Nanoplot

Se eliminaron los adaptadores empleados durante la secuenciación, estableciendo un umbral del 98% en Porechop. En total, se eliminaron 2.190.781 pb correspondientes a los adaptadores de inicio y 578.144 pb de los adaptadores finales. Posteriormente se eliminó las secuencias de tamaño menor a 1000 pb y Ph < 8 con Chopper, obteniendo un total de 104.935 lecturas.

Las lecturas filtradas fueron analizadas en EPI2ME bajo el análisis WIMP que funciona con una base de datos de referencia que cubre todos los genomas de bacterias, virus y hongos disponibles en RefSeq (*EPI2ME WIMP Workflow*, 2019). En este análisis se clasificaron con éxito 104.030 lecturas, obteniendo que el 2% de las lecturas corresponden a virus como: *Avipoxvirus*, *Lambdavirus*, *Lederbergvirus*, entre otros (Figura 5).

**Figura 5**

*Identificación y cuantificación de lecturas en EPI2ME*

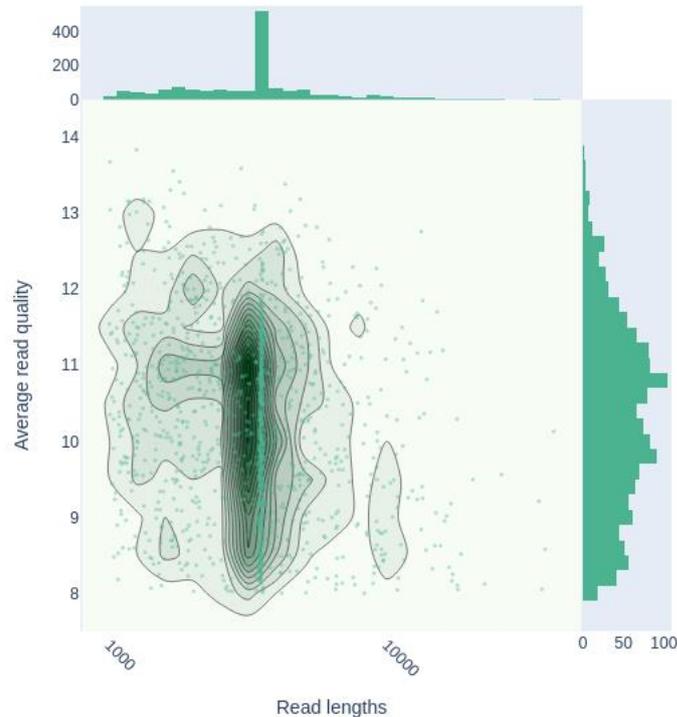


Nota. El diagrama destaca los reinos taxonómicos presentes en la muestra. También se desglosan algunos de los virus identificados según el análisis de EPI2ME.

En base a este análisis se filtró las secuencias que coincidían como virus obteniendo un total de 1.435 lecturas. La calidad de las secuencias se analizó en Nanoplot obteniendo la Figura 6, donde se observa claramente una distribución densa y variable de lecturas, mismas que cumplen con los estándares de calidad y tamaño previamente establecidos en el pipeline.

**Figura 6**

*Evaluación de las lecturas identificadas como virus*



Nota. En el siguiente gráfico de estimación de densidad de kernel (KDE), se visualiza la distribución de las lecturas de acuerdo con su tamaño y la calidad.

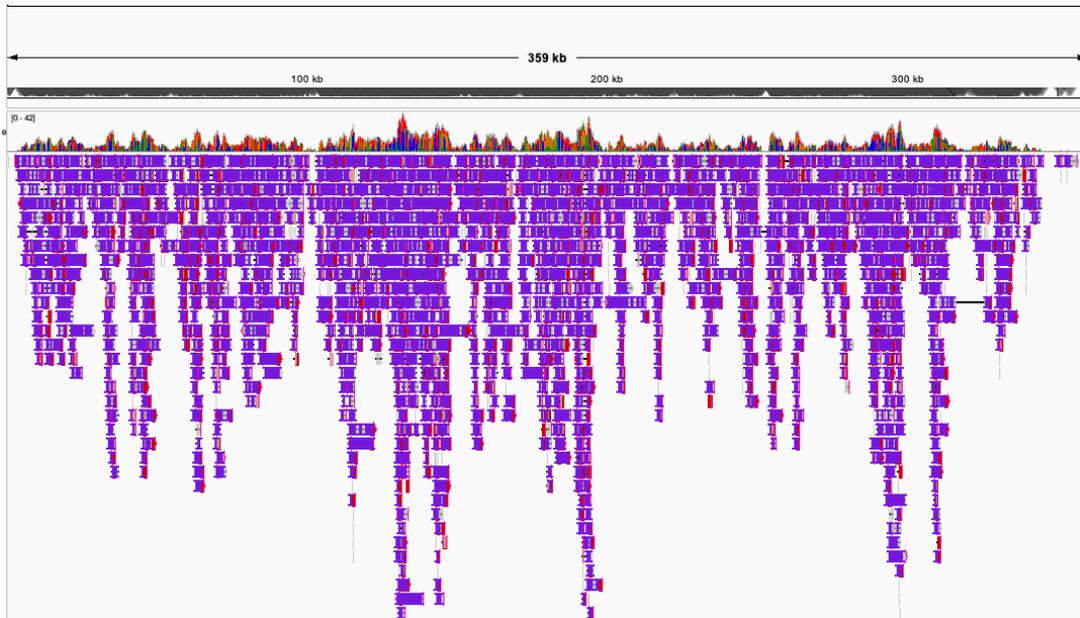
#### **4.5 Ensamblaje del Genoma**

Previamente, se determinó que las secuencias víricas abarcan una cobertura del 42x y una profundidad de secuenciación variable que se refleja en los picos y valles de la Figura 7. Esto sugiere que la cobertura no es uniforme y presenta inconsistencias a lo largo del genoma.

A partir del ensamblaje basado en la secuencia de referencia de *Canarypox virus* (NC\_005309.1) y tras la corrección de errores (polishing), se generó una secuencia de 351.122 pb, con un contenido de GC del 30.02% (Figura 8).

**Figura 7**

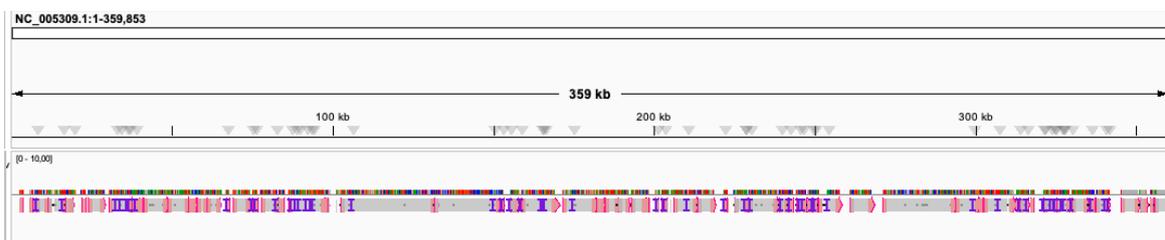
*Ensamblaje de los contigs en base a la secuencia de referencia*



Nota. En el siguiente gráfico generado en IGV, se pueden observar las lecturas alineadas al genoma de referencia, junto con los parámetros de profundidad y cobertura.

**Figura 8**

*Visualización del genoma reconstruido de Avipoxvirus sp.*



Nota. En el siguiente gráfico generado en IGV, se observa la reconstrucción del genoma *Avipoxvirus* sp. presente en la muestra de curruca (*Sylvia atricapilla*).

#### 4.6 Análisis filogenético

El software de alineamiento MAFFT seleccionó el algoritmo FFT-NS-2 "Fast but rough" como el más adecuado, dado el alto volumen de datos y la necesidad de obtener resultados de forma rápida.

La secuencia alineada fue ajustada manualmente para corregir errores en regiones conservadas entre los avipoxvirus, observándose cambios recurrentes de adenina (A) por guanina (G) y de guanina (G) por timina (T). Además, debido a la alta variabilidad y ausencia de información en una región específica, se decidió trabajar con la secuencia hasta la posición 249.264 pb.

El análisis filogenético de *Warblerpox virus* reveló una alta identidad con respecto a *Canarypox virus* (98.88%), *Shearwaterpox virus* (98.80%), *Albatrosspox virus* (98.74%) y *Cheloniid poxvirus* (88.95%) (Tabla 4). Sin embargo, la identidad con el resto de las secuencias de poxvirus baja notablemente, siendo la siguiente en proximidad la del *Penguinpox virus* (66.92%).

**Tabla 4**

*Valores de identidad filogenética de Warblerpox virus en comparación con otras especies de poxvirus*

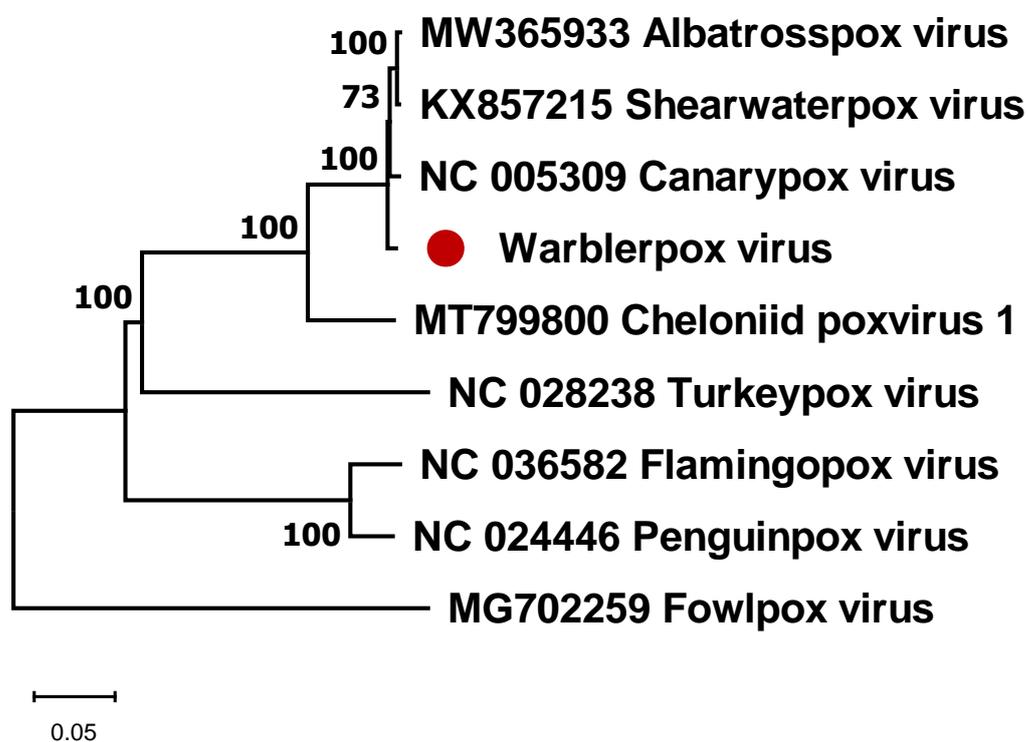
	1	2	3	4	5	6	7	8
<b><i>Warblerpox virus</i></b>								
(NC005309)	98,88							
<b><i>Canarypox virus</i></b>								
(KX857215)	98,80	98,88						
<b><i>Shearwaterpox virus</i></b>								
(MW365933)	98,74	98,78	99,73					
<b><i>Albatrosspox virus</i></b>								
(MT799800)	88,95	88,85	89,30	89,18				
<b><i>Cheloniid poxvirus</i></b>								
(NC036582)	66,27	65,82	65,84	65,89	66,14			
<b><i>Flamingopox virus</i></b>								
(MG702259)	50,34	50,24	50,32	50,33	50,62	50,27		
<b><i>Fowlpox virus</i></b>								
(NC024446)	66,92	66,34	66,39	66,50	66,69	94,28	50,51	
<b><i>Penguinpox virus</i></b>								
(NC028238)	66,81	66,10	66,13	66,14	66,60	64,31	48,36	64,44
<b><i>Turkeypox virus</i></b>								

Nota. Matriz de identidades entre las secuencias estudiadas.

El árbol filogenético (Figura 9), clasifica la secuencia obtenida en nuestro estudio dentro del clado B por su identidad con *Canarypox virus*. Finalmente, consideramos la nueva secuencia obtenida como una nueva especie viral, denominándola tentativamente como *Warblerpox virus*.

**Figura 9**

*Árbol filogenético*



Nota. Árbol filogenético generado bajo el algoritmo de Neighbor-Joining. La secuencia obtenida en nuestro estudio se marca con un círculo en rojo. Los valores de bootstrap por debajo de 70 se omiten en el árbol.

## 5 Discusión

La obtención de genomas completos es fundamental para el estudio de cualquier organismo (Baker, 2022). En la era posgenómica, hemos presenciado un crecimiento exponencial en el número de genomas secuenciados, por lo que actualmente contamos con más de 200.000 genomas en GenBank. Sin embargo, a pesar de este avance, la mayoría de los genomas disponibles provienen de organismos procariontes y fúngicos

(Nagy et al., 2020). Esta situación resalta la importancia de investigar los genomas virales, y en particular los avipoxvirus, para generar información que contribuya como base para comprender mejor su evolución y patogenicidad (Sarker, Bowden, & Boyle, 2021).

En este estudio, se logró extraer y amplificar ADN viral a partir de muestras de lesiones cutáneas de curruca. Este método de extracción directa ha demostrado ser eficiente en términos de tiempo y recursos, y ha sido empleado en numerosos estudios recientes (Sarker, Hannon, Athukorala, et al., 2021; Sarker & Sutherland, 2022). No obstante, existen metodologías alternativas para el aislamiento de ADN viral, como la extracción a partir de cuerpos de mosquitos (*Culex spp.*) (van der Meer et al., 2022) o el aislamiento del virus en membranas corioalantoideas (CAM) (Sarker, Bowden, & Boyle, 2021), que pueden ofrecer una mayor concentración del ADN viral. Pese a ello, esta investigación demuestra que una concentración inicial relativamente baja es suficiente para la detección y reconstrucción del genoma.

La secuenciación mediante tecnología Oxford Nanopore (ONT) ha revolucionado el campo de la genómica al permitir la obtención rápida, sencilla y rentable de lecturas largas. Esto ha facilitado el acceso a la secuenciación tanto en el ámbito de la investigación como en el diagnóstico clínico (Kumar et al., 2024; Tshiabuila et al., 2024). En este contexto, el protocolo de laboratorio propuesto podría ser fácilmente implementado en entornos rutinarios de laboratorio, ya que ha demostrado ser eficiente y reproducible en la extracción, amplificación y secuenciación de ADN viral.

Sin embargo, es crucial reconocer las limitaciones inherentes a la tecnología de secuenciación utilizada. Los errores no son completamente aleatorios, sino que tienden a ocurrir en tramos homopoliméricos o secuencias de repeticiones cortas, donde el software de identificación de bases enfrenta dificultades para determinar con precisión cuántas iteraciones consecutivas de una base han pasado a través del nanoporo (Espinosa et al., 2024). Esto da lugar a errores de inserción o eliminación de nucleótidos (indels), la aparición de errores en secuencias con extensiones de nucleótidos idénticos y la disminución de la calidad hacia el final de las lecturas (Baker, 2022). Estas inconsistencias fueron evidentes durante la revisión del alineamiento filogenético, lo

que conllevo a realizar ajustes manuales en las regiones conservadas entre los avipoxvirus analizados. Además, se detectaron áreas del ensamblaje con secuencias faltantes, probablemente causadas por problemas en la fase de secuenciación o ensamblaje.

Para mitigar estos errores, es recomendable aumentar la profundidad de cobertura (King et al., 2020). Esto se puede lograr optimizando la extracción de ADN y la amplificación de regiones específicas mediante PCR con primers diseñados específicamente para este fin (Hook & Timp, 2023). Además, la combinación de tecnologías de secuenciación, como la secuenciación de Illumina, puede validar y complementar los datos obtenidos, asegurando una mayor precisión en las regiones críticas del genoma (Wick et al., 2023).

Por otra parte, dado que la precisión en la secuenciación es esencial para obtener datos fiables, es fundamental considerar la selección de herramientas bioinformáticas adecuadas para su procesamiento. Para ello, se dispone de una amplia gama de herramientas diseñadas específicamente para los datos generados por las plataformas de secuenciación de Oxford Nanopore (ONT) (Lu et al., 2016). En el protocolo propuesto, se utilizaron varias de estas herramientas, incluyendo NanoPlot, Porechop, Chopper, Flye, Minimap2, Samtools, QUAST, Racon y Medaka, todas disponibles en GitHub. Estas herramientas han sido identificadas como las más utilizadas en investigaciones recientes (Baker, 2022; Hu et al., 2021; Khrenova et al., 2022; Riaz et al., 2021; Saud et al., 2021; Schacksen et al., 2024). Además, se empleó EPI2ME, una aplicación desarrollada por Oxford Nanopore.

En particular, el ensamblaje del genoma es una de las tareas más importantes en el ámbito de la genómica (Khrenova et al., 2022; Sutton et al., 2021). Para ello, se pueden utilizar dos estrategias: el reensamblaje, que alinea las lecturas a un genoma de referencia existente, o el ensamblaje de *novo* del genoma completo (Espinosa et al., 2024). Este último es particularmente valioso en el contexto de enfermedades infecciosas emergentes, donde los genomas de referencia pueden no estar disponibles (Hu et al., 2021). Sin embargo, el ensamblaje de *novo* presenta desafíos significativos de

tiempo y recurso computacional, motivo por el cual en este estudio se optó por un enfoque basado en secuencia de referencia.

En este contexto, la corrección de errores surge como un paso fundamental posterior al ensamblaje de lecturas largas, asegurando la precisión y fiabilidad de los datos genómicos secuenciados (Sutton et al., 2021). En este estudio, el uso de Medaka para la corrección de errores, permitiendo una representación más confiable del genoma viral.

Por otra parte, el análisis filogenético realizado sobre la secuencia ha revelado una notable identidad con otros avipoxvirus del clado B, principalmente con el virus de la viruela canaria (CNPV (Gong et al., 2019). Esta relación se extiende a otros poxvirus del mismo clado como *Shearwaterpox* (SWPV) y *Albatrosspox* como lo mencionan Sarker et al. (2021) y Sarker et al. (2017). Este hecho puede estar relacionado con una divergencia evolutiva a partir de un fenómeno de coevolución parásito-hospedador, procedente de un ancestro común. La identidad observada con *Cheloniid poxvirus* se debe a que este posee similitud con SWPV y CNPV (Sarker et al., 2021), lo que refuerza la idea de que *Warblerpox virus* comparte un ancestro común con otros miembros de poxvirus. Sin embargo, es importante señalar que la identidad con otras secuencias de clados supuestamente diferentes disminuye considerablemente, siendo la más cercana la de *Penguinpox virus*. Esto subraya la singularidad de *Warblerpox virus* y su clasificación dentro del clado B.

Finalmente, es importante tener en cuenta que una misma especie de virus de viruela aviar puede infectar a diferentes tipos de aves, y una sola ave puede ser infectada por múltiples cepas del virus (Liu et al., 2024). Por ello, se recomienda realizar estudios posteriores con análisis adicionales, como el estudio de genes específicos como la polimerasa o el gen P4b y análisis de anotación o expresión de proteínas, para confirmar si se trata efectivamente de una nueva especie.

## 6 Conclusiones

Este estudio permitió la obtención y análisis del genoma de un avipoxvirus a partir de lesiones cutáneas de curruca, empleando tecnología Oxford Nanopore (ONT). Los resultados demuestran que el protocolo de extracción directa de ADN viral es eficiente

y viable, incluso con concentraciones iniciales relativamente bajas, logrando la amplificación suficiente para la secuenciación y posterior análisis genómico. Este enfoque simplificado podría implementarse en estudios similares, facilitando la caracterización genética de virus emergentes en entornos con recursos limitados.

La tecnología ONT ha demostrado ser una herramienta poderosa en la secuenciación de lecturas largas, permitiendo el ensamblaje de genomas completos de manera rápida y rentable. No obstante, las limitaciones inherentes a la secuenciación, tales como los errores en tramos homopoliméricos y las regiones repetitivas, requieren estrategias adicionales de corrección de errores, como el uso de herramientas bioinformáticas avanzadas. En este sentido, se utilizó un pipeline con herramientas robustas en el procesamiento y análisis de los datos genómicos permitiendo reconstruir una secuencia de 351.122 pb .

Finalmente, tras considerar la identidad filogenética observada y la notable divergencia con otras secuencias de *Avipoxvirus*, se clasifica la secuencia obtenida de curruca como parte del clado B, a la que hemos denominado tentativamente como *Warblerpox virus*.

## 7 Bibliografía

Giotis, E. S., & Skinner, M. A. (2019). Spotlight on avian pathology: Fowlpox virus. *Avian Pathology: Journal of the W.V.P.A*, 48(2), 87-90.  
<https://doi.org/10.1080/03079457.2018.1554893>

Liu, H., Li, T., Tang, J., Ding, S., & Wang, D. (2024). Complete genome sequence analysis of Reticuloendotheliosis virus integrated in nonhomologous Avipoxvirus. *Microbial Pathogenesis*, 194, 106827.  
<https://doi.org/10.1016/j.micpath.2024.106827>

Lebdah, M., Ali, A. M., Ali, A. A., & Hassanin, O. (2019). Insights into pathological and molecular characterization of avipoxviruses circulating in Egypt. *British Poultry Science*, 60(6), 666-674. <https://doi.org/10.1080/00071668.2019.1639141>

Net, R. L., Provost, C., Lalonde, C., Régimbald, L., Vézina, F., Gagnon, C. A., & Lair, S. (2020). WHOLE GENOME SEQUENCING OF AN AVIPOXVIRUS ASSOCIATED WITH INFECTIONS IN A GROUP OF AVIARY-HOUSED SNOW BUNTINGS (PLECTROPHENAX NIVALIS). *Journal of Zoo and Wildlife Medicine*, 50(4), 803-812. <https://doi.org/10.1638/2018-0102>

Asif, K., O'Rourke, D., Legione, A. R., Shil, P., Marendá, M. S., & Noormohammadi, A. H. (2021). Whole-genome based strain identification of fowlpox virus directly from cutaneous tissue and propagated virus. *PLOS ONE*, 16(12), e0261122. <https://doi.org/10.1371/journal.pone.0261122>

Croville, G., Le Loc'h, G., Zanchetta, C., Manno, M., Camus-Bouclainville, C., Klopp, C., Delverdier, M., Lucas, M.-N., Donnadieu, C., Delpont, M., & Guérin, J.-L. (2018). Rapid whole-genome based typing and surveillance of avipoxviruses using nanopore sequencing. *Journal of Virological Methods*, 261, 34-39. <https://doi.org/10.1016/j.jviromet.2018.08.003>

Deng, L., Liu, C., Li, L., Hao, P., Wang, M., Jin, N., Yin, R., Du, S., & Li, C. (2023). Genomic characteristics of an avipoxvirus 282E4 strain. *Virus Research*, 336, 199218. <https://doi.org/10.1016/j.virusres.2023.199218>

Sasani, T. A., Cone, K. R., Quinlan, A. R., & Elde, N. C. (2018). Long read sequencing reveals poxvirus evolution through rapid homogenization of gene arrays. *eLife*, 7, e35453. <https://doi.org/10.7554/eLife.35453>

Brinkmann, A., Pape, K., Uddin, S., Woelk, N., Förster, S., Jessen, H., Michel, J., Kohl, C., Schaade, L., & Nitsche, A. (2024). Genome sequencing of the mpox virus 2022 outbreak with amplicon-based Oxford Nanopore MinION sequencing. *Journal of*



- Ribeiro, L. C., Monteiro, F. L., Chagas, D. B., Vargas, G. D., Lima, M. de, Fischer, G., & Hübner, S. de O. (2020). Identification of Clade E Avipoxvirus in Brazil. *Avian Diseases*, 64(2), 223-227. <https://doi.org/10.1637/0005-2086-64.2.223>
- He, L., Zhang, Y., Jia, Y., Li, Z., Li, J., Shang, K., Ding, K., Yu, H., & Sarker, S. (2023). A novel pathogenic avipoxvirus infecting oriental turtle dove (*Streptopelia orientalis*) in China shows a high genomic and evolutionary proximity with the pigeon avipoxviruses isolated globally. *Microbiology Spectrum*, 11(5), e01193-23. <https://doi.org/10.1128/spectrum.01193-23>
- Grüning, B., Dale, R., Sjödin, A., Chapman, B. A., Rowe, J., Tomkins-Tinch, C. H., Valieris, R., & Köster, J. (2018). Bioconda: Sustainable and comprehensive software distribution for the life sciences. *Nature Methods*, 15(7), 475-476. <https://doi.org/10.1038/s41592-018-0046-7>
- Schacksen, P. S., Østergaard, S. K., Eskildsen, M. H., & Nielsen, J. L. (2024). Complete pipeline for Oxford Nanopore Technology amplicon sequencing (ONT-AmpSeq): From pre-processing to creating an operational taxonomic unit table. *FEBS Open Bio*, n/a(n/a). <https://doi.org/10.1002/2211-5463.13868>
- Baker, J. L. (2022). Using Nanopore Sequencing to Obtain Complete Bacterial Genomes from Saliva Samples. *mSystems*, 7(5), e0049122. <https://doi.org/10.1128/msystems.00491-22>
- Hu, T., Li, J., Zhou, H., Li, C., Holmes, E. C., & Shi, W. (2021). Bioinformatics resources for SARS-CoV-2 discovery and surveillance. *Briefings in Bioinformatics*, 22(2), 631-641. <https://doi.org/10.1093/bib/bbaa386>
- Khrenova, M. G., Panova, T. V., Rodin, V. A., Kryakvin, M. A., Lukyanov, D. A., Osterman, I. A., & Zvereva, M. I. (2022). Nanopore Sequencing for De Novo Bacterial

- Genome Assembly and Search for Single-Nucleotide Polymorphism. *International Journal of Molecular Sciences*, 23(15), Article 15. <https://doi.org/10.3390/ijms23158569>
- Riaz, N., Leung, P., Barton, K., Smith, M. A., Carswell, S., Bull, R., Lloyd, A. R., & Rodrigo, C. (2021). Adaptation of Oxford Nanopore technology for hepatitis C whole genome sequencing and identification of within-host viral variants. *BMC Genomics*, 22(1), 148. <https://doi.org/10.1186/s12864-021-07460-1>
- EPI2ME WIMP workflow: Quantitative, real-time species identification from metagenomic samples.* (2019, noviembre 12). Oxford Nanopore Technologies. <https://nanoporetech.com/resource-centre/epi2me-wimp-workflow-quantitative-real-time-species-identification-metagenomic>
- Nagy, L. G., Merényi, Z., Hegedüs, B., & Bálint, B. (2020). Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. *Nucleic Acids Research*, 48(5), 2209-2219. <https://doi.org/10.1093/nar/gkz1241>
- Sarker, S., Bowden, T. R., & Boyle, D. B. (2021). Genomic characterisation of a novel avipoxvirus, magpiepox virus 2, from an Australian magpie (*Gymnorhina tibicen terraereginae*). *Virology*, 562, 121-127. <https://doi.org/10.1016/j.virol.2021.07.010>
- Sarker, S., Hannon, C., Athukorala, A., & Bielefeldt-Ohmann, H. (2021). Emergence of a Novel Pathogenic Poxvirus Infection in the Endangered Green Sea Turtle (*Chelonia mydas*) Highlights a Key Threatening Process. *Viruses*, 13(2), Article 2. <https://doi.org/10.3390/v13020219>

- Sarker, S., & Sutherland, M. (2022). Molecular characterisation of a novel pathogenic avipoxvirus from an Australian little crow (*Corvus bennetti*) directly from the clinical sample. *Scientific Reports*, *12*(1), 15053. <https://doi.org/10.1038/s41598-022-19480-2>
- van der Meer, C. S., Paulino, P. G., Jardim, T. H. A., Senne, N. A., Araujo, T. R., dos Santos Juliano, D., Massard, C. L., Peixoto, M. P., da Costa Angelo, I., & Santos, H. A. (2022). Detection and molecular characterization of Avipoxvirus in *Culex* spp. (Culicidae) captured in domestic areas in Rio de Janeiro, Brazil. *Scientific Reports*, *12*(1), 13496. <https://doi.org/10.1038/s41598-022-17745-4>
- Kumar, K. R., Cowley, M. J., & Davis, R. L. (2024). Next-Generation Sequencing and Emerging Technologies\*. *Seminars in Thrombosis and Hemostasis*, *50*(7), 1026-1038. <https://doi.org/10.1055/s-0044-1786397>
- Tshiabuila, D., Choga, W., James, S. E., Maponga, T., Preiser, W., van Zyl, G., Moir, M., van Wyk, S., Giandhari, J., Pillay, S., Anyaneji, U. J., Lessells, R. J., Naidoo, Y., Sanko, T. J., Wilkinson, E., Tegally, H., Baxter, C., Martin, D. P., & de Oliveira, T. (2024). An Oxford Nanopore Technology-Based Hepatitis B Virus Sequencing Protocol Suitable For Genomic Surveillance Within Clinical Diagnostic Settings. *medRxiv*, 2024.01.19.24301519. <https://doi.org/10.1101/2024.01.19.24301519>
- Hook, P. W., & Timp, W. (2023). Beyond assembly: The increasing flexibility of single-molecule sequencing technology. *Nature Reviews Genetics*, *24*(9), 627-641. <https://doi.org/10.1038/s41576-023-00600-1>
- Wick, R. R., Judd, L. M., & Holt, K. E. (2023). Assembling the perfect bacterial genome using Oxford Nanopore and Illumina sequencing. *PLOS Computational Biology*, *19*(3), e1010905. <https://doi.org/10.1371/journal.pcbi.1010905>

- Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*, 14(5), 265-279.  
<https://doi.org/10.1016/j.gpb.2016.05.004>
- Sutton, J. M., Millwood, J. D., Case McCormack, A., & Fierst, J. L. (2021). Optimizing experimental design for genome sequencing and assembly with Oxford Nanopore Technologies. *GigaByte*, 2021, gigabyte27.  
<https://doi.org/10.46471/gigabyte.27>
- Espinosa, E., Bautista, R., Larrosa, R., & Plata, O. (2024). Advancements in long-read genome sequencing technologies and algorithms. *Genomics*, 116(3), 110842.  
<https://doi.org/10.1016/j.ygeno.2024.110842>
- ong, Y., Feng, N., Bi, J., Zhang, X., Li, Y., Sun, W., Wang, T., Li, N., Zhao, Y., Yang, S., Zhang, W., Tang, J., Wang, J., Gao, Y., & Xia, X. (2019). Molecular Phylogeny of an Avipoxvirus Isolated from Red-Flanked Blue Robin in China. *Avian Diseases*, 64(1), 2-6. <https://doi.org/10.1637/0005-2086-64.1.2>
- Sarker, S., Athukorala, A., Nyandowe, T., Bowden, T. R., & Boyle, D. B. (2021). Genomic Characterisation of a Novel Avipoxvirus Isolated from an Endangered Northern Royal Albatross (*Diomedea sanfordi*). *Pathogens*, 10(5), Article 5.  
<https://doi.org/10.3390/pathogens10050575>
- Sarker, S., Das, S., Lavers, J. L., Hutton, I., Helbig, K., Imbery, J., Upton, C., & Raidal, S. R. (2017). Genomic characterization of two novel pathogenic avipoxviruses isolated from pacific shearwaters (*Ardenna* spp.). *BMC Genomics*, 18(1), 298.  
<https://doi.org/10.1186/s12864-017-3680-z>

## 8 Anexos

### Anexo 1

#### *Protocolo de purificación*

##### 1. Unión del ADN a la columna

- Añadir 210  $\mu$ l de etanol grado molecular al ADN y mezclar usando vortex.
- Colocar la mezcla en una columna ubicada en un tubo de colección
- Centrifugar 1 min a 11.000 x g. Descartar el filtrado y colocar la columna nuevamente en el tubo.

##### 2. Lavados de la membrana

- Añadir 500  $\mu$ l de Buffer BBW. Centrifugue 1 min a 11,000 x g. Desechar el filtrado y coloque la columna nuevamente en el Collection tube.
- Añadir 600  $\mu$ l de Buffer BB5. Centrifugue 1 min a 11,000 x g. Descartar el tubo y el filtrado.

##### 3. Secado de la membrana

- Coloque la columna en un microtubo de 1.5 mL y centrifugue 1 min a 11,000 x g para secar la membrana.

##### 4. Elución del ADN

- Transferir la columna a un microtubo de 1.5 mL y añada 50  $\mu$ l de Buffer BBE precalentado a (70°C) directamente en la membrana. Incube a temperatura ambiente durante 1 min y centrifugue 1 min a 11,000 x g. El eluido contiene el ADN puro extraído de la muestra.

## Anexo 2

*Inputs y Outputs del pipeline aprobado*

Link:  [TFM\\_DATA\\_223C6092](#)

## Anexo 3

*Pipelines*

Link: [Pipelines AVP ONT.html](#)

## 9 Declaración obligatoria del uso de herramientas de IA

El autor y tutor de este trabajo de investigación declaran el uso de ChatGPT durante la ejecución y redacción de este trabajo, con el fin de generar ideas relacionadas con el mismo. Después de utilizar la herramienta, el autor y el tutor revisaron y editaron minuciosamente el contenido según fuera necesario y asumen la total responsabilidad por el contenido de esta memoria.