

## Datos

- Nombre: Elisa Luque Jiménez
- Curso Académico: 2024/2025
- Título provisional de la Tesis Doctoral: Búsqueda de artrópodos reservorios de bacterias y virus patógenos mediante el análisis de datos de secuenciación NGS públicos.
- Línea de Investigación: Ingeniería, Ciencia de Datos y Bioinformática.
- Director de la tesis doctoral: Antonio Jesús Pérez Pulido.
- Financiación: Contrato con duración del proyecto (cuatro años) de Titulada Superior de Apoyo a la Investigación.
- Lugar donde se llevará a cabo la investigación: Centro Andaluz de Biología del Desarrollo (CABD).
- Situación de la tesis: Comienzo de tesis.

## Plan de Investigación

- Introducción:

En las bases de datos públicas de datos genómicos existen actualmente miles de genomas borradores de especies de artrópodos. Cuando se secuencian estos genomas, pueden estar contaminados por genomas de microorganismos, principalmente virus y bacterias, que se encuentran colonizando al animal de partida (Cornet L et al, 2022). Las secuencias de estos otros genomas son normalmente filtrados de los proyectos de secuenciación del artrópodo, pero se encontrarían disponibles en las lecturas de los datos crudos (*reads* en inglés). Así, estudiando estas secuencias, utilizando herramientas de bioinformática y biología computacional, se podrían encontrar reservorios animales de microorganismos patógenos oportunistas de humanos. El conocimiento de estos reservorios sería de utilidad a la hora de manejar infecciones de patógenos conocidos o de nuevo surgimiento en nuestro entorno.

En concreto, el género *Acinetobacter* es de especial interés actualmente en clínica, debido principalmente a su resistencia a antibióticos (Bharadwaj A et al, 2022) (A Hernández et al, 2010). Y, precisamente, miembros de este grupo de bacterias han sido recientemente encontrados en el medio ambiente, aislados en aves, que a su vez conviven con invertebrados (M Dahiru et al, 2015).

Dentro de un proyecto de secuenciación, cuyos datos son públicos, hemos encontrado secuencias de una bacteria del género *Acinetobacter* en un ácaro del género *Oppiella*, por lo que usaremos los datos crudos de este proyecto como control positivo.

- Estado de la cuestión:

En la actualidad, se han publicado trabajos en los que se ha reportado la presencia de bacterias en artrópodos, centrados en algunas especies en concreto o en artrópodos que se encuentran en hospitales.

El estudio del microbioma de artrópodos está revelando una compleja red de interacciones entre estos animales y sus microorganismos asociados. Se ha documentado la presencia de bacterias simbióticas pertenecientes a los géneros *Wolbachia*, *Rickettsia* o *Sodalis*, que pueden influir en la reproducción, inmunidad y capacidad como vector del hospedador (Zug R et al, 2015). Sin embargo, más allá de estos simbioses ya conocidos, existe una diversidad microbiana aún poco caracterizada, especialmente en especies no hematófagas o no asociadas a enfermedades humanas.

Diversos estudios han demostrado que los artrópodos pueden albergar bacterias con genes de resistencia a antibióticos, incluso en entornos no clínicos (Founou L et al, 2016). Por ejemplo, se han detectado genes de resistencia a antibióticos, betalactamasas, y bombas de flujo en insectos urbanos y rurales, lo que sugiere una circulación ambiental de estos determinantes genéticos. Esta resistencia puede ser adquirida por transferencia horizontal y facilitada por la convivencia de múltiples especies bacterianas en el mismo hospedador (McGann P et al, 2016).

El uso de datos de secuenciación públicos para detectar microorganismos en hospedadores no humanos ha sido validado en varios estudios. Por ejemplo, en el artículo de Li et al. (2015) utilizaron datos de transcriptomas de insectos para identificar virus previamente no descritos. Asimismo, herramientas como Kraken2 permiten realizar análisis taxonómicos y funcionales de genomas y metagenomas a partir de datos no diseñados originalmente para ese fin (Wood et al., 2019).

En cuanto al género *Acinetobacter*, su presencia en artrópodos ha sido documentada en cucarachas, moscas y garrapatas, tanto en entornos urbanos como rurales. Pueden formar parte del microbioma intestinal o cuticular, y su papel como vectores o reservorios aún no está completamente descrito. La detección de genes de resistencia en estas cepas refuerza la necesidad de estudiar su ecología y evolución en hospedadores no humanos.

A pesar de estos avances, persisten importantes lagunas de conocimiento. La mayoría de estudios se centran en especies de interés médico o agrícola ya conocidas, dejando de lado la enorme diversidad de artrópodos presentes en ecosistemas naturales de los que aún se desconoce su posible papel como vector de patógenos.

- Hipótesis y/o preguntas de investigación:

- Hay una presencia desconocida y relevante de bacterias y virus en artrópodos, con potencial de transmisión y patogenicidad en humanos, por lo que el estudio de sus datos genómicos sería de gran interés.
- Existen diferencias entre las bacterias presentes en artrópodos domésticos y de ambientes urbanos (que habitan en nuestras residencias y espacios habitables) y las bacterias presentes en artrópodos de ambiente rural.
- Existen diferencias entre las bacterias presentes en artrópodos que habitan hospitales y las bacterias presentes en artrópodos de ambiente urbano y rural.
- Los artrópodos pueden actuar como vectores de bacterias resistentes a antibióticos tanto para humanos como para otros animales.
- Hay animales afectados por la transmisión de bacterias patógenas resistentes a antibióticos afectan indirectamente a humanos, como en la industria cárnica.

- Objetivos que se pretenden alcanzar:

- Describir el género de bacteria *Acinetobacter* encontrada en una especie de ácaro del género *Opplia*.
- Buscar secuencias de bacterias en otros ácaros de la misma familia Oppiidae.
- Descargar datos crudos de secuenciación de genomas de artrópodos aislados en diferentes ámbitos.
- Buscar secuencias de bacterias y virus en los genomas de artrópodos descargados, restringiendo a grupos taxonómicos concretos, en caso de existir un gran número.
- Realizar pangenomas de las especies de bacterias y virus encontradas para caracterizarlas filogenéticamente dentro de la especie y estudiar su potencial patogenicidad y virulencia.
- Buscar y analizar genes de resistencia a antibióticos y de virulencia en las secuencias de bacterias y virus encontrados.
- Investigar posibles diferencias entre la presencia de bacterias en artrópodos de hábitat rural y de hábitat urbano.
- Distinguir entre contaminación y transferencia horizontal bacteriana entre los datos de artrópodos. Dilucidar si, en caso de contaminación, se debe a una relación mutualista de simbiosis, tanto de endosimbiosis como de ectosimbiosis.

Estas búsquedas de datos genómicos serán realizadas en bases de datos con datos públicos de acceso disponible, como las bases de datos del National Center for Biotechnological Information (NCBI).

- Metodología:

- Realización de listado de especies de artrópodos urbanos y rurales a analizar.
- Descarga de datos crudos y metadatos.
- Análisis de calidad y limpieza de datos crudos.
- Identificación taxonómica.

- Mapeado entre lecturas y genomas de referencia.
- Ensamblado de genomas.
- Anotación estructural y funcional de genomas.
- Identificación de proteínas y dominios proteicos de interés.
- Identificación de virus y profagos.
- Construcción de filogenias.
- Medios materiales:
  - Material informático:
    - PCs
    - Servidor C3UPO
    - Servidor NAS de grupo
    - Servidor CICA Hércules, Junta de Andalucía
  - Herramientas bioinformáticas:
    - NCBI datasets
    - FastQC
    - BLAST
    - MAGICBLAST
    - Kraken2
    - Nxtrim
    - Trimmomatic
    - FastP
    - Hisat2
    - Bowtie2
    - Salmon
    - Samtools
    - Bakta
    - PHASTEST
    - Unicycler
    - EggNOG-mapper
    - Rag-tag
    - Mummer
    - Panaroo
    - MAFFT
    - IQ-Tree
    - Phylopic
  - Programas informáticos:
    - RStudio
    - Visual Studio Code
    - InkScape
    - Notepad ++
    - IGV
    - FileZilla
    - Bitvise SSH Client
  - Lenguajes informáticos:
    - Linux
    - Python
    - Perl
    - R

- Planificación temporal:

Objetivos	2024 - 2025	2025 - 2026	2027- 2028	2028 - 2029
Revisión bibliográfica y formación en diversas herramientas bioinformáticas	X	X	X	X
Validación metodológica con caso control ( <i>Oppeilla</i> y <i>Acinetobacter</i> )	X			
Selección y creación del listado de especies	X	X		
Realización completa del pipeline de análisis de datos	X	X		
Comparación entre resultados de distintos hábitats		X	X	
Estudio de contaminación, transferencia horizontal y relaciones simbióticas	X	X	X	
Construcción de pangenomas y análisis filogenético		X	X	X
Redacción de la tesis				X

- Referencias Bibliográficas:

- Cornet L, Baurain D. Contamination detection in genomic data: more is not enough. *Genome Biol.* 2022 Feb 21;23(1):60. doi: 10.1186/s13059-022-02619-9. PMID: 35189924; PMCID: PMC8862208.
- Bharadwaj A, Rastogi A, Pandey S, Gupta S, Sohal JS. Multidrug-Resistant Bacteria: Their Mechanism of Action and Prophylaxis. *Biomed Res Int.* 2022 Sep 5;2022:5419874. doi: 10.1155/2022/5419874. PMID: 36105930; PMCID: PMC9467707.
- A. Hernández Torres, E. García Vázquez, G. Yagüe, J. Gómez Gómez. Multidrug resistant *Acinetobacter baumannii*: clinical update and new highlights. *Revista Española de Quimioterapia*, ISSN-e 0214-3429, Vol. 23, Nº. 1, 2010, págs. 12-19.

- M. Dahiru & O. I. Enabulele. *Acinetobacter baumannii* in Birds' Feces: A Public Health Threat to Vegetables and Irrigation Farmers. *Advances in Microbiology*, Vol.5 No.10, 2015. doi: 10.4236/aim.2015.510072.
  - Zug R, Hammerstein P. Bad guys turned nice? A critical assessment of *Wolbachia* mutualisms in arthropod hosts. *Biol Rev Camb Philos Soc*. 2015 Feb;90(1):89-111. doi: 10.1111/brv.12098. Epub 2014 Mar 11. PMID: 24618033.
  - Founou LL, Founou RC, Essack SY. Antibiotic Resistance in the Food Chain: A Developing Country-Perspective. *Front Microbiol*. 2016 Nov 23;7:1881. doi: 10.3389/fmicb.2016.01881. PMID: 27933044; PMCID: PMC5120092.
  - McGann P, Snesrud E, Maybank R, Corey B, Ong AC, Clifford R, Hinkle M, Whitman T, Lesho E, Schaecher KE. *Escherichia coli* Harboring *mcr-1* and *blaCTX-M* on a Novel *IncF* Plasmid: First Report of *mcr-1* in the United States. *Antimicrob Agents Chemother*. 2016 Jun 20;60(7):4420-1. doi: 10.1128/AAC.01103-16. Erratum in: *Antimicrob Agents Chemother*. 2016 Jul 22;60(8):5107. doi: 10.1128/AAC.01353-16. PMID: 27230792; PMCID: PMC4914657.
  - Ci-Xiu Li, Mang Shi, Jun-Hua Tian, Xian-Dan Lin, Yan-Jun Kang, Liang-Jun Chen, Xin-Cheng Qin, Jianguo Xu, Edward C Holmes, Yong-Zhen Zhang (2015). Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* 4:e05378. doi: 10.7554/eLife.05378
  - Wood, D.E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 20, 257 (2019). doi: 10.1186/s13059-019-1891-0
  - Bae JM. Methodological issues for determining intervals of subsequent cancer screening. *Epidemiol Health*. 2014;36.e2014010. doi: 10.4178/epih/e2014010
- Modalidad: Tesis monográfica.
  - Menciones: Tesis con mención internacional. Se prevé que la estancia dure alrededor de tres meses y que el trabajo científico que se realice esté directamente relacionado con los objetivos de la tesis mencionados anteriormente o con los de alguna publicación. Actualmente se está llevando a cabo una búsqueda de grupos de investigación y proyectos con los que colaborar para la estancia.

## Plan de Formación

- Plan de formación de la Escuela de Doctorado:

- Curso “La documentación científica para la elaboración de la Tesis”, impartido del 13 al 31 enero de 2025, con un total de 8 horas.
- Curso “La revisión bibliográfica como método de investigación”, impartido del 19 de mayo al 6 junio de 2025, con un total de 6 horas.
- Curso “Técnicas avanzadas de gestión de la información”, impartido del 3 al 27 de febrero de 2025, con un total de 10 horas.

Se espera la participación de más cursos de la Escuela de Doctorado para el siguiente curso 2025/2026.

- Plan de formación de los Programas de Doctorado:

Se planea participar en los siguientes cursos en el próximo curso 2025-2026:

- Data processing using R: basic statistics and graph generation (duración de 10 horas).
- Introduction to Data Analysis with Orange3 (duración de 10 horas).

- Seminarios de Investigación:

Asistencia a seminarios internos (Centro Andaluz de Biología del Desarrollo, CABD):

- Antonio Giráldez. Yale University. Understanding the mechanisms of gene regulations through genomics and imaging. [09-04-2025].
- Modesto Orozco. Instituto de Investigación Biomédica (IRB) Barcelona. Expanding the frontiers of nucleic acid simulations. [21-03-2025].
- Dirk Shübeler. Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland. Finding your place: transcription factors and modifiers of chromatin. [11-04-2025].

Asistencia a seminarios externos:

- Juan Pascual-Gil & Pedro Costales Maestre. Estación Biológica de Doñana. Fire effects on mediterranean communities from an ecosystem perspective / An integrative approach to the diversity of the Mediterranean Sonchus asper taxonomic complex. [13-02-2025].
- Ladislav Hamerlik. Estación Biológica de Doñana. Paleolimnology: Reading the lake sediments to reveal environmental history. [20-02-2025].

- Reuniones de seguimiento de los Proyectos:

Todos los jueves se han realizado reuniones de grupo de los integrantes del grupo de investigación a cargo de Antonio Jesús Pérez Pulido, en las que se ha seguido el progreso de la tesis semanalmente.

Se planea participar en los CABD Meetings y presentar una charla en el año 2025, que tratará sobre los inicios de la tesis doctoral y los avances hasta el momento.

- Cursos de especialización científica:

Todos los títulos obtenidos a lo largo de la carrera académica previa a la tesis han aportado conocimiento para ésta:

- Máster en Análisis de Datos Ómicos y Biología de Sistemas, Universidad Internacional de Andalucía, Sevilla (2023 - presente).
- Diploma de experta en Bioinformática, Universidad Pablo de Olavide, Sevilla (2023 - 2024).
- Curso de Experto Universitario en Entomología Aplicada, Universidad Nacional de Educación a Distancia (2021).
- Máster en Biodiversidad, Universidad Autónoma de Madrid, Madrid (2019 - 2020).
- Grado en Biotecnología, Universidad Pablo de Olavide, Sevilla (2014 - 2019).

- Asistencia y participación a congresos:

Se ha asistido y participado en las Jornadas Andaluzas de Bioinformática (JABI), en Málaga, del 12 al 14 de mayo de 2025, realizando una comunicación oral titulada “Bacteriophages in arthropods or contamination in genome databases?”.

Se planea participar además en las próximas JABI, presentando póster, y en los años siguientes que se realice la tesis.

Se planea participar en las Jornadas Doctorales en el curso 2024/2025 que se emplazarán en Carmona, en el Palacio de los Briones (sede de la Universidad Pablo de Olavide), en el mes de octubre. Se realizará una presentación de póster.

- Publicaciones:

En estos momentos se encuentra en proceso de publicar un artículo como primera autora titulado “A bacteriophage in a mite genome reveals bacterial contamination and opens new possibilities for exploring arthropod symbionts” en la revista *Microbial Genomics*, a la que enviamos el manuscrito en el mes de abril y recibimos primera revisión en junio (*Major revision*). Enviaremos la segunda versión corregida y ampliada antes de agosto con perspectiva de ser publicada en los próximos meses. Este trabajo incluye los resultados de la primera parte de la tesis.

Se prevé publicar más artículos y/o comunicaciones en los siguientes años de tesis, tanto como primera autora, como otras colaboraciones en el grupo.

- Estancias:

No se ha realizado ninguna estancia por el momento. Se planea hacerla en un centro de investigación internacional para el tercer año de tesis, en el curso 2026-2027, de una duración de alrededor de tres meses. El objetivo de la estancia será de un trabajo que deberá formar una aportación sustancial a la tesis, ya sea trabajo de campo y laboratorio o trabajo computacional.



- Formación para la docencia:

Se planea ejercer de ayudante de prácticas para prácticas tuteladas de la asignatura “Bioinformática” del grado de Biotecnología de la Universidad Pablo de Olavide en los próximos años de la tesis. Durante el curso 2025-2026 ya se han planificado alrededor de 21 horas.

Se planea ejercer como docente en el diploma de especialización de Bioinformática de la Universidad Pablo de Olavide a partir del curso 2026-2027.

- Iniciación a la Dirección de Proyectos Científicos y a la Transferencia de Conocimientos:

Como primera autora del artículo mencionado anteriormente, y para su completa realización entre varios miembros del grupo de investigación de mi director de tesis, se ha debido ejercer una labor de coordinadora de proyecto, de tal forma que las tareas ejercidas por los miembros participantes fueran coherentes las unas con las otras, se realizaran correctamente y finalmente se pusieran en común los resultados para construir el artículo.

- Complementos Formativos:

- Participación en el taller del Centro Andaluz de Biología del Desarrollo en la Feria de las Ciencias de Sevilla, el 22 de abril de 2025, duración de cuatro horas.
- Participación en el taller “¿Cómo se forman los seres vivos?” en el Colegio de Educación Infantil y Primaria Ana María Matute en Dos Hermanas, Sevilla, el día 21 de enero de 2025, duración de dos horas.
- Participación en el taller “Aprendiendo con los ratones” en el Colegio de Educación Infantil y Primaria Tartessos, en Sevilla Este, el día 28 de mayo de 2025, duración de tres horas.