

# Genes, Genomas y Proteomas

---

## Resultados de aprendizaje del curso

---

1. Integrar datos genómicos, transcriptómicos y proteómicos para analizar sistemas biológicos, con una comprensión crítica de los mecanismos moleculares que subyacen a la regulación génica, a la arquitectura del genoma y a la función proteica.
2. Aplicar métodos estadísticos y computacionales a datos biológicos de alto rendimiento, incluyendo RNA-seq, espectrometría de masas y anotación genómica, utilizando herramientas como R/Bioconductor.
3. Evaluar críticamente la literatura científica en genómica funcional y evolutiva, y liderar de manera efectiva discusiones académicas sobre métodos y hallazgos de vanguardia.
4. Diseñar experimentos rigurosos basados en ómicas, incorporando controles apropiados, replicación y estrategias de análisis de datos para abordar preguntas biológicas complejas a nivel de sistemas.

## Sección 1: Evolución, homología y genómica comparada

---

### 1. Quest for Orthologs in the Era of Data Deluge and AI: Challenges and Innovations in Orthology Prediction and Data Integration

Majidian S. et al. | *Journal of Molecular Evolution* (2025)

<https://doi.org/10.1007/s00239-025-10272-6>

**Resumen:** Síntesis colaborativa del 8.º encuentro Quest for Orthologs que revisa el estado actual de la inferencia de ortología y paralogía, examinando cómo la IA, las estructuras proteicas predichas, las isoformas de splicing y la evolución de los dominios proteicos están remodelando la detección de ortólogos y la anotación funcional. Cataloga métodos, recursos y prácticas de benchmarking, y plantea desafíos abiertos como la anotación de parálogos.

**Pregunta de discusión:** *La clase distingue ortología, paralogía y xenología mediante las pruebas de Patterson ¿Cómo operacionalizan los métodos computacionales modernos de detección de ortólogos esa lógica conceptual, y qué supuestos evolutivos podrían fallar al inferir ortología a escala genómica con herramientas basadas en IA?*

### 2. Human Endogenous Retroviruses: Our Genomic Fossils and Companions

Stein R.A. & DePaola R.V. | *Physiological Genomics* (2023)

<https://doi.org/10.1152/physiolgenomics.00171.2022>

**Resumen:** Esta revisión explica cómo cerca del 8% del genoma humano corresponde a retrovirus endógenos (HERV) dejados por antiguas infecciones de la línea germinal por retrovirus hoy extintos, integrados en posiciones ortólogas hace decenas de millones de años y heredados verticalmente. Muestra que, lejos de ser 'ADN basura', algunos HERV han sido cooptados para funciones esenciales (p. ej., las sincitinas en la placentación).

**Pregunta de discusión:** *En la clase, los ERV en loci ortólogos compartidos son una de las líneas de evidencia molecular más sólidas de ascendencia común. ¿Por qué una integración independiente en el mismo sitio ortólogo es astronómicamente improbable, y qué advertencias (p. ej., el caso de PTERV1) deben tenerse en cuenta al usar ERV como marcadores filogenéticos?*

### 3. The Last Universal Common Ancestor of Ribosome-Encoding Organisms: Portrait of LUCA

Forterre P. | *Journal of Molecular Evolution* (2024)

<https://doi.org/10.1007/s00239-024-10186-9>

**Resumen:** Revisión crítica que reconstruye la biología inferida del Último Ancestro Común Universal (LUCA) a partir de análisis filogenómicos de genes conservados, y argumenta que LUCA fue un organismo relativamente simple que ya realizaba el metabolismo central y la síntesis de proteínas. Destaca cómo las preguntas no resueltas sobre las ramas profundas y las decisiones de modelo/datos moldean lo que inferimos sobre la raíz del árbol de la vida.

**Pregunta de discusión:** *La clase presenta el código genético casi universal y la maquinaria central compartida como evidencia de un único ancestro común (prueba formal de la UCA de Theobald, 2010). ¿Cómo dependen las reconstrucciones de LUCA de los supuestos de los modelos filogenómicos y qué señal distingue la verdadera ascendencia común de la mera similitud de secuencia?*

### 4. Evolutionary Sparse Learning Reveals the Shared Genetic Basis of Convergent Traits

Allard J.B. et al. | *Nature Communications* (2025)

<https://doi.org/10.1038/s41467-025-58428-8>

**Resumen:** Los autores presentan ESL-PSC, un método de aprendizaje automático basado en la filogenia que utiliza contrastes entre pares de especies para identificar genes y sitios subyacentes a rasgos convergentes evolucionados de forma independiente, validado en la fotosíntesis C4 en gramíneas y en la ecolocalización en mamíferos. El método descuenta explícitamente la convergencia aparente debida a la ascendencia compartida.

**Pregunta de discusión:** *El contraste central de la clase es entre homología y convergencia (las 'alas' de aves y murciélagos como ejemplo). ¿Cómo separa metodológicamente este enfoque la convergencia molecular genuina de la similitud heredada, y qué riesgos de falsos positivos persisten al buscar firmas de evolución convergente en todo el proteoma?*

## Sección 2: Estructura de genes, organización del genoma y secuencias repetitivas

---

### 5. Repetitive DNA sequence detection and its role in the human genome

Liao X. et al. | *Communications Biology* (2023)

<https://doi.org/10.1038/s42003-023-05322-y>

**Resumen:** Esta revisión ofrece un recorrido completo del ADN repetitivo (repeticiones en tándem, elementos transponibles y duplicaciones segmentarias), describiendo cómo se clasifican, cómo se detectan computacionalmente y cómo impulsan la evolución del genoma y contribuyen a enfermedades humanas complejas.

**Pregunta de discusión:** *Dado que las secuencias repetitivas ocupan una gran fracción del genoma humano y se asocian con enfermedades, ¿cómo debería integrarse su detección y anotación en el diseño de un experimento de genómica funcional? ¿Qué desafíos metodológicos hay que considerar?*

### 6. Complete sequencing of ape genomes

Yoo D. et al. | *Nature* (2025)

<https://doi.org/10.1038/s41586-025-08816-3>

**Resumen:** Mediante ensamblajes telómero a telómero de seis especies de grandes simios, este estudio histórico resuelve regiones repetitivas antes inaccesibles (duplicaciones

segmentarias, centrómeros y loci del MHC), aportando una nueva línea base para entender la evolución y la diversidad del genoma humano.

**Pregunta de discusión:** *Los autores solo pudieron resolver ciertas regiones genómicas mediante el ensamblaje de telómero a telómero. ¿Qué implica esto para interpretar conclusiones evolutivas extraídas de genomas de referencia previos con vacíos? ¿Cómo cambia nuestra comprensión de la evolución del genoma?*

## Sección 3: Replicación del ADN, recombinación, elementos transponibles y retrovirus

### 7. [Bridge RNAs direct programmable recombination of target and donor DNA](#)

Durrant M.G. et al. | *Nature* (2024)

<https://doi.org/10.1038/s41586-024-07552-4>

**Resumen:** Este estudio descubre que las secuencias de inserción IS110 codifican un 'ARN puente' no codificante que contacta simultáneamente el ADN donante y el diana para guiar la recombinación sitio-específica, introduciendo un mecanismo enteramente nuevo de edición de ADN guiada por ARN con potencial programable.

**Pregunta de discusión:** *Los ARN puente representan un paradigma fundamentalmente nuevo en la biología guiada por ácidos nucleicos, distinto del de CRISPR. Compara la lógica mecánica de la recombinación por ARN puente con la edición mediante CRISPR-Cas9. ¿Qué implicaciones tiene este descubrimiento para la ingeniería de genomas y la biología de los elementos transponibles?*

### 8. [Retrotransposon Life Cycle and Its Impacts on Cellular Responses](#)

Luqman-Fatah A. et al. | *RNA Biology* (2024)

<https://doi.org/10.1080/15476286.2024.2409607>

**Resumen:** Recorre el ciclo de vida de los retrotransposones humanos (LINE-1, Alu, HERV) y los mecanismos del hospedero que controlan sus intermediarios de ARN y ADNc. Su reactivación durante la tumorigénesis, el desarrollo, el estrés y el envejecimiento compromete la estabilidad genómica, a la vez que estos elementos operan como motores de diversidad genética y de la inmunidad innata.

**Pregunta de discusión:** *Los retrotransposones son a la vez parásitos genómicos y elementos reguladores. ¿Qué factores del hospedero restringen su movilización en las células somáticas y cómo su desregulación conecta inestabilidad genómica, inflamación y autoinmunidad?*

### 9. [Biochemical Mechanisms of Genetic Recombination and DNA Repair](#)

Raina V.B. et al. | *Annual Review of Biochemistry* (2025)

<https://doi.org/10.1146/annurev-biochem-083024-113931>

**Resumen:** Esta revisión traza los mecanismos bioquímicos de la recombinación homóloga, desde la invasión de cadena hasta la resolución de las uniones de Holliday, integrando hallazgos recientes de experimentos con moléculas únicas que han revelado la cinética y los intermediarios estructurales del proceso.

**Pregunta de discusión:** *¿Cómo revisan o refinan los hallazgos de los estudios de molécula única sobre la recombinación los modelos clásicos de HR? ¿Qué implicaciones tienen estos detalles mecánicos para entender la recombinación meiótica y su papel en la generación de diversidad genética?*

## Sección 4: Regulación de la expresión génica, splicing, ARN catalítico y código genético

---

### 10. [The regulation and function of post-transcriptional RNA splicing](#)

Choquet K. et al. | *Nature Reviews Genetics* (2025)

<https://doi.org/10.1038/s41576-025-00836-z>

**Resumen:** Esta revisión revela que hasta el 40% de los intrones de mamíferos se procesan (splicing) después de la terminación de la transcripción y no durante ella, lo que replantea el splicing postranscripcional como una capa reguladora importante del control de la expresión génica durante el desarrollo, el estrés y la enfermedad.

**Pregunta de discusión:** *La visión tradicional considera el splicing como mayormente cotranscripcional. ¿Cómo cambia el descubrimiento del splicing postranscripcional generalizado nuestra comprensión mecánica de cómo las células regulan el momento y la abundancia de isoformas proteicas específicas en respuesta a señales celulares?*

### 11. [Genome organization around nuclear speckles drives mRNA splicing efficiency](#)

Bhat P. et al. | *Nature* (2024)

<https://doi.org/10.1038/s41586-024-07429-6>

**Resumen:** Este estudio demuestra que la proximidad física de un gen a los nuclear speckles, condensados enriquecidos en reguladores del splicing, determina directamente la eficiencia con la que se procesa su pre-ARNm, vinculando la organización tridimensional del núcleo con la expresión génica.

**Pregunta de discusión:** *Este artículo sostiene que la posición genómica en el núcleo, no solo la secuencia del pre-ARNm, controla la eficiencia del splicing. ¿Cómo se integra este modelo espacial de regulación génica con las señales de splicing basadas en secuencia? ¿Qué diseños experimentales se necesitarían para distinguir ambos mecanismos?*

### 12. [mRNA Regulation by RNA Modifications](#)

Gilbert W.V. et al. | *Annual Review of Biochemistry* (2023)

<https://doi.org/10.1146/annurev-biochem-052521-035949>

**Resumen:** Las modificaciones químicas del ARNm (como m6A, m5C y pseudouridina) constituyen un 'epitranscriptoma' que influye en cada etapa de la vida del ARNm, desde la transcripción en el núcleo hasta la degradación citoplasmática, lo que abre nuevas perspectivas sobre el control postranscripcional de la expresión génica.

**Pregunta de discusión:** *¿Cómo se comparan las modificaciones epitranscriptómicas con la metilación del ADN, mejor estudiada, como mecanismos reguladores? Al diseñar un experimento de RNA-seq o proteómica, ¿qué pasos tomarías para tener en cuenta los posibles efectos de confusión de las modificaciones del ARNm en la cuantificación?*

## Sección 5: Tecnologías de secuenciación masiva — Illumina, PacBio y Nanopore

---

### 13. [Next-Generation Sequencing Technology: Current Trends and Advancements](#)

Satam H. et al. | *Biology* (2023)

<https://doi.org/10.3390/biology12070997>

**Resumen:** Esta amplia revisión mapea el panorama de las plataformas NGS y sus aplicaciones en genómica, transcriptómica, epigenómica y metagenómica, y cubre los

principios técnicos, los desafíos del análisis de datos y la trayectoria hacia la secuenciación de moléculas únicas y de lecturas largas.

**Pregunta de discusión:** *El artículo describe la NGS como facilitadora de la medicina de precisión y la genética de poblaciones. Para una pregunta de investigación de tu área (p. ej., estudiar la regulación génica en un organismo no modelo), ¿qué plataforma NGS elegirías y cómo tendría en cuenta tu diseño experimental las limitaciones de esa plataforma?*

#### 14. [The long and short of it: benchmarking viromics using Illumina, Nanopore and PacBio sequencing technologies](#)

Cook R. et al. | *Microbial Genomics* (2024)

<https://doi.org/10.1099/mgen.0.001198>

**Resumen:** Usando una comunidad mock controlada, este estudio compara rigurosamente Illumina, PacBio y Nanopore en la recuperación de genomas, las tasas de error y la predicción génica posterior, y encuentra que los ensamblajes híbridos (Illumina + Nanopore) superan a cualquier tecnología por separado.

**Pregunta de discusión:** *El estudio concluye que ninguna tecnología de secuenciación es óptima en todos los contextos. En el contexto de secuenciar un genoma eucariota complejo con muchas repeticiones, ¿cómo diseñarías una estrategia de secuenciación híbrida y qué métricas usarías para evaluar la calidad del ensamblaje?*

### Sección 6: Diseño experimental de alto rendimiento, ensamblaje y anotación de genomas

#### 15. [BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA](#)

Gabriel L. et al. | *Genome Research* (2024)

<https://doi.org/10.1101/gr.278090.123>

**Resumen:** BRAKER3 es un pipeline automatizado de anotación de genomas que integra RNA-seq y bases de datos de proteínas con modelos génicos entrenados de forma iterativa, superando a enfoques previos, especialmente en genomas eucariotas grandes y complejos.

**Pregunta de discusión:** *La calidad de la anotación de un genoma depende en gran medida del tipo y la calidad de la evidencia aportada. Compara las contribuciones relativas de la evidencia de RNA-seq frente a la de la homología proteica al entrenar buscadores de genes. En ausencia de datos transcriptómicos de una especie nueva, ¿cómo priorizarías tu estrategia de anotación?*

#### 16. [Building better genome annotations across the tree of life](#)

Freedman A.H. et al. | *Genome Research* (2025)

<https://doi.org/10.1101/gr.280377.124>

**Resumen:** Esta comparación sistemática de 12 métodos de anotación en 21 especies halla que la herramienta de transferencia de anotación TOGA, junto con BRAKER3 y StringTie, son consistentemente los de mejor desempeño, y que los datos de RNA-seq mejoran sustancialmente la calidad de anotación en organismos no modelo.

**Pregunta de discusión:** *La transferencia de anotación entre especies depende de la conservación evolutiva. ¿En qué condiciones filogenéticas esperarías que la transferencia de anotación falle, y cómo combinarías evidencia de transferencia con evidencia de novo (RNA-seq, proteínas) para anotar un linaje sin parientes cercanos bien anotados?*

## Sección 7: Modelos estadísticos para datos ómicos — modelos generativos, mixtos y RNA-seq

---

### 17. [Statistical Strategies for Differential Expression Analysis of RNA Sequencing Data: Model Comparison and Benchmarking](#)

Luo Y. | *Theoretical and Natural Science* (2025)

<https://doi.org/10.54254/2753-8818/2025.au28984>

**Resumen:** Esta revisión de benchmarking compara enfoques binomiales negativos, bayesianos, de modelos lineales y no paramétricos para el análisis de expresión diferencial en RNA-seq, discutiendo los compromisos en sensibilidad, especificidad y robustez según el tamaño muestral y el diseño experimental.

**Pregunta de discusión:** *El artículo destaca que los tamaños muestrales pequeños representan un desafío persistente en los estudios de RNA-seq. Para el diseño experimental de un seminario, ¿cómo justificarías tu elección del modelo estadístico para la expresión diferencial, y qué análisis de potencia harías a priori para determinar el número mínimo de réplicas biológicas?*

### 18. [Exploring and mitigating shortcomings in single-cell differential expression analysis with a new statistical paradigm](#)

Wu C.-H. et al. | *Genome Biology* (2025)

<https://doi.org/10.1186/s13059-025-03525-6>

**Resumen:** Este artículo identifica cuatro 'maldiciones' fundamentales de la expresión diferencial en single-cell (exceso de ceros, problemas de normalización, efectos de donante y sesgos acumulativos) y propone un nuevo marco de efectos mixtos Poisson/Binomial generalizado (GLIMES) para abordarlas.

**Pregunta de discusión:** *El modelo GLIMES usa expresión absoluta de ARN en lugar de abundancia relativa. ¿Qué implicaciones tiene esta elección para la interpretación biológica de los resultados de expresión diferencial? ¿Cómo decidirías entre un marco de cuantificación relativa y uno absoluto en tu propia investigación?*

## Sección 8: Genómica funcional — RNA-seq, single-cell, ChIP-seq, ATAC-seq y metilación del ADN

---

### 19. [The technological landscape and applications of single-cell multi-omics](#)

Baysoy A. et al. | *Nature Reviews Molecular Cell Biology* (2023)

<https://doi.org/10.1038/s41580-023-00615-w>

**Resumen:** Esta revisión integral abarca las tecnologías de célula única que perfilan simultáneamente el transcriptoma, el epigenoma, el proteoma y otras capas moleculares de una misma célula, y describe sus aplicaciones en la construcción de atlas celulares, la inmunología tumoral y el trazado de linajes.

**Pregunta de discusión:** *El perfilado simultáneo del transcriptoma y de la accesibilidad de la cromatina en células individuales aporta información complementaria, aunque a veces contradictoria, sobre el estado celular. Diseña un experimento multiómico de célula única para investigar cambios en la regulación génica durante la diferenciación celular, especificando qué modalidades medirías y por qué.*

### 20. [Spatially resolved genome-wide joint profiling of epigenome and transcriptome](#)

Li H. et al. | *Nature Protocols* (2025)

<https://doi.org/10.1038/s41596-025-01145-9>

**Resumen:** Este artículo de protocolos describe spatial-ATAC-RNA-seq y spatial-CUT&Tag-RNA-seq, dos métodos que coperefilan la expresión génica y las marcas epigenómicas (accesibilidad de la cromatina o modificaciones de histonas) dentro de secciones de tejido intactas a una resolución casi de célula única.

**Pregunta de discusión:** *¿Cómo cambian, con la preservación de la información espacial junto con las mediciones moleculares, las conclusiones que se pueden extraer en comparación con los enfoques de célula única disociada? Propón una pregunta biológica de tu área que requiera específicamente datos multiómicos con resolución espacial para responderla.*

## 21. [A DNA methylation atlas of normal human cell types](#)

Loyfer N. et al. | *Nature* (2023)

<https://doi.org/10.1038/s41586-022-05580-6>

**Resumen:** Mediante secuenciación profunda con bisulfito de genoma completo de 39 tipos celulares humanos purificados, este estudio construye un atlas integral del metiloma que revela enhancers reguladores específicos de cada tipo celular y loci regulados por Polycomb, y aporta una referencia para la biopsia líquida y los estudios de enfermedad.

**Pregunta de discusión:** *El atlas muestra que los loci no metilados específicos de un tipo celular suelen encontrarse en enhancers específicos de tejido. ¿Cómo podrías usar este atlas de referencia para inferir la composición celular de una muestra de tejido heterogéneo a partir de datos de secuenciación con bisulfito en bulk? ¿Qué limitaciones tendría este enfoque?*

## Sección 9: Proteómica y cuantificación de péptidos — espectrometría de masas y biomarcadores

### 22. [Comprehensive Overview of Bottom-Up Proteomics Using Mass Spectrometry](#)

Jiang Y. et al. | *ACS Measurement Science Au* (2024)

<https://doi.org/10.1021/acsmeasuresciau.3c00068>

**Resumen:** Esta revisión tipo manual cubre el flujo de trabajo completo de la proteómica bottom-up (desde la extracción y digestión de proteínas hasta la adquisición por LC-MS/MS y la interpretación bioinformática), aportando conocimiento fundamental para quien diseñe un experimento de proteómica.

**Pregunta de discusión:** *La proteómica bottom-up pierde inherentemente información sobre las proteoformas intactas debido a la digestión de las proteínas. ¿En qué contextos biológicos sería más consecuente esta pérdida de información de proteoformas, y cómo complementarías un enfoque bottom-up para abordar esta limitación?*

### 23. [Data-Independent Acquisition: A Milestone and Prospect in Clinical Mass Spectrometry-Based Proteomics](#)

Fröhlich K. et al. | *Molecular & Cellular Proteomics* (2024)

<https://doi.org/10.1016/j.mcpro.2024.100800>

**Resumen:** Esta revisión explica por qué la adquisición independiente de datos (DIA) ha transformado la proteómica al permitir la medición no sesgada y reproducible de todos los péptidos detectables de una muestra, y discute las herramientas de software, las aplicaciones clínicas y los desafíos pendientes para implementar DIA en diagnóstico de rutina.

**Pregunta de discusión:** Los autores recomiendan usar múltiples herramientas de software DIA en paralelo porque cada una tiene sus propios sesgos. ¿Qué implica esto para la reproducibilidad y la estandarización de los datos de proteómica en entornos clínicos? ¿Cómo diseñarías un estudio de validación para demostrar la utilidad clínica de un panel de biomarcadores proteicos identificado mediante DIA-MS?

#### **24. Recent developments in mass-spectrometry-based targeted proteomics of clinical cancer biomarkers**

Wenk D. et al. | *Clinical Proteomics* (2024)

<https://doi.org/10.1186/s12014-024-09452-1>

**Resumen:** Esta revisión abarca el estado de la proteómica dirigida por LC-MS/MS para medir biomarcadores proteicos de cáncer en muestras clínicas, contrastando este enfoque con los inmunoensayos y discutiendo el camino desde el descubrimiento hasta la validación clínica, e incluyendo los desafíos clave que aún limitan su implementación a gran escala.

**Pregunta de discusión:** El artículo describe una brecha entre el descubrimiento de biomarcadores mediante proteómica y su validación clínica. Dados los desafíos regulatorios, analíticos y biológicos descritos, propón un marco de validación por etapas para un biomarcador de cáncer recién descubierto mediante proteómica, especificando los criterios estadísticos y técnicos de cada etapa.

### **Sección 10: Integración de datos multiómicos**

---

#### **25. Methods and applications for single-cell and spatial multi-omics**

Vandereyken K. et al. | *Nature Reviews Genetics* (2023)

<https://doi.org/10.1038/s41576-023-00580-2>

**Resumen:** Esta revisión autorizada abarca el panorama experimental y computacional de la multiómica de célula única y espacial, y explica cómo la medición conjunta del transcriptoma, el epigenoma, el proteoma y el metaboloma de células individuales está transformando nuestra comprensión del desarrollo, la enfermedad y la identidad celular.

**Pregunta de discusión:** La revisión describe la multiómica como facilitadora de la inferencia causal sobre la regulación génica al vincular las distintas capas moleculares. ¿Cuáles son los supuestos computacionales y estadísticos clave que deben cumplirse para que tales inferencias causales sean válidas? ¿Cómo comprobarías si se cumplen en un conjunto de datos concreto?

#### **26. Multi-omics data integration using ratio-based quantitative profiling with Quartet reference materials**

Zheng Y. et al. | *Nature Biotechnology* (2023)

<https://doi.org/10.1038/s41587-023-01934-1>

**Resumen:** Usando un cuarteto familiar como sistema de muestras de referencia con verdad de base incorporada (basada en el parentesco), este estudio demuestra que la cuantificación absoluta, no el perfilado basado en razones, es la causa raíz de la irreproducibilidad en la integración multiómica y propone una solución práctica.

**Pregunta de discusión:** El artículo sostiene que el mayor obstáculo para la integración multiómica es la falta de materiales de referencia comunes, no los algoritmos. ¿Cómo diseñarías un estudio multiómico en tu institución para maximizar la reproducibilidad entre lotes y entre plataformas? ¿Qué muestras de referencia y métricas de control de calidad adoptarías?

## 27. [Deep learning-based approaches for multi-omics data integration and analysis](#)

Ballard J.L. et al. | *BioData Mining* (2024)

<https://doi.org/10.1186/s13040-024-00391-z>

**Resumen:** Esta revisión clasifica las arquitecturas de deep learning para la integración multiómica (redes feedforward, redes convolucionales de grafos, autoencoders y modelos generativos) y compara su capacidad para manejar datos faltantes, fusionar modalidades heterogéneas y producir conclusiones biológicas interpretables.

**Pregunta de discusión:** *Los modelos de deep learning suelen lograr un desempeño predictivo superior a costa de la interpretabilidad. En un estudio ómico clínico cuyo objetivo es tanto la predicción (p. ej., el pronóstico del paciente) como el descubrimiento de mecanismos, ¿cómo equilibrarías la complejidad del modelo con la necesidad de explicabilidad biológica?*