



Curso pre-congreso — Congreso Nacional de Genética 2026
De humanos, ratones, genes y algoritmos: Bioinformática e IA
para un uso responsable de modelos animales
Sociedad Mexicana de Genética · Facultad de Ciencias, UNAM

PLAN DEL CURSO

Fechas: 1 al 5 de junio de 2026
Horario: 10:00 a 14:00 h
Modalidad: teórico-práctica, presencial

Profesores

Dra. Sandra Cabrera Benítez

Laboratorio de Fibrosis, Facultad de Ciencias, UNAM. Laboratorio de Biopatología Pulmonar Ciencias-INER

Dra. Yalbi Itzel Balderas-Martínez

Laboratorio de Biología Computacional, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER)

Ing. Luis Alberto Meza Cova

Laboratorio de Biología Computacional, I Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER)

Versión 2.0 · Abril 2026

1. Presentación y justificación

La citogenética clásica y los modelos murinos han sido, durante décadas, los dos pilares del estudio de las enfermedades genéticas raras. La primera, apoyada en el microscopio y la interpretación experta de bandas cromosómicas; el segundo, basado en la generación y fenotipado de ratones *knock-in*, *knock-out* y condicionales. Ambos enfoques continúan siendo indispensables, pero el panorama del 2026 no es el mismo del 2010.

La disponibilidad de bases de datos genómicas y citogenéticas de acceso libre, los navegadores genómicos de última generación, los paquetes bioinformáticos de R y Python, y más recientemente la irrupción de modelos de aprendizaje profundo para análisis de imágenes de cariotipo y predicción de fenotipos, permiten al investigador actual obtener información que hace una década sólo podía conseguirse con experimentación adicional. Este curso plantea una pregunta que tiene implicaciones éticas, científicas y económicas directas: ¿podemos usar estas herramientas computacionales para reducir el número de experimentos en modelos animales que requieren nuestras preguntas de investigación?

La respuesta no es “sí, los reemplazamos”. Es más matizada. Las herramientas bioinformáticas permiten priorizar mejor qué preguntas merecen un ratón, diseñar cohortes más pequeñas con mayor poder estadístico, predecir fenotipos antes de generar un modelo, rescatar información de experimentos ya publicados e integrar niveles de evidencia que antes estaban dispersos. El curso enseña a usar estas herramientas de manera práctica, siempre con el marco ético de las 3R (Russell y Burch, 1959: Replace, Reduce, Refine) como brújula explícita.

Las sesiones de bioinformática (miércoles 3 y jueves 4 de junio) complementan las sesiones de la Dra. Sandra Cabrera (lunes 1, martes 2 y viernes 5), que abordan la biología del ratón genéticamente modificado, los componentes genéticos de las enfermedades raras, la ingeniería genética y los aspectos bioéticos. Mientras la Dra. Cabrera cubre los fundamentos biológicos y experimentales, nuestras sesiones operacionalizan ese conocimiento con las herramientas computacionales que el profesional moderno debe dominar para tomar decisiones experimentales informadas.

2. Objetivos

2.1 Objetivo general del curso

Proveer herramientas al participante para que tenga una visión general de cómo tanto los modelos in vivo desarrollados en ratones genéticamente modificados como la bioinformática han contribuido al entendimiento de enfermedades raras y al refinamiento de estrategias experimentales de investigación. Al finalizar las sesiones de bioinformática, los participantes serán capaces de navegar las principales bases de datos genómicas y citogenéticas de acceso libre, realizar análisis básicos de variantes y regiones cromosómicas con R/Bioconductor y Python/Colab, interpretar resultados bajo criterios estandarizados

internacionales (ACMG/ClinGen), y articular cómo estas herramientas computacionales contribuyen a reducir el número de experimentos en modelos murinos necesarios para el estudio de enfermedades genéticas raras, en el marco de las 3R.

2.2 Objetivos específicos

De conocimiento

- Se revisará la organización del genoma humano y los conceptos de homología y sintenia, para después lograr entender porqué la especie *Mus musculus* ha sido ampliamente utilizada para el estudio de enfermedades humanas. Se profundizará en la revisión de ejemplos de genes y mutaciones específicas que generan enfermedades raras, se analizarán los síndromes o manifestaciones fenotípicas de las mutaciones que conducen a las enfermedades raras.
- Describir la arquitectura y el alcance de las bases de datos genómicas y citogenéticas de acceso libre más relevantes: UCSC Genome Browser, Ensembl, DECIPHER, ClinVar, OMIM, DGV, MGI (Mouse Genome Informatics) e IMPC (International Mouse Phenotyping Consortium).
- Explicar los criterios ACMG/ClinGen para la clasificación de variantes de número de copias (CNVs) como patogénicas, probablemente patogénicas, de significado incierto, probablemente benignas o benignas.
- Reconocer los principios básicos de las redes neuronales convolucionales (CNN) aplicadas al análisis de imágenes de metafases y clasificación automática de cromosomas.
- Articular el marco de las 3R (reemplazar, reducir, refinar) de Russell y Burch y sus implicaciones concretas para el uso de modelos murinos en la era de la bioinformática.

De habilidades

- Consulta de bases de datos para la identificación de mutaciones en genes específicos y los ratones modificados genéticamente que han sido incorporados a dichas bases de datos.
- Traducir una descripción citogenética en nomenclatura ISCN a coordenadas genómicas en GRCh38/hg38 y viceversa, usando las tablas de bandas citogenéticas de UCSC y los recursos de Ensembl.
- Consultar programáticamente Ensembl, ClinVar y DECIPHER desde R con biomaRt y manipular regiones genómicas con GenomicRanges.
- Ejecutar un modelo pre-entrenado de clasificación de cromosomas en Google Colab a partir de un *dataset* abierto de imágenes de metafases.
- Identificar ortólogos humano-ratón para un conjunto de genes de interés clínico y evaluar la existencia de modelos murinos publicados (MGI, IMPC) para variantes específicas.
- Conocimiento de estrategias experimentales de reemplazo de modelos animales.

- Sensibilización en aspectos bioéticos sobre el uso de generación de ratones mutantes y aplicación de las 3R (Reemplazar, Reducir, Refinar).

De actitudes

- Valorar el acceso abierto a datos y herramientas (*open data, open source*) como un bien común científico que acelera la investigación y democratiza el conocimiento.
- Adoptar una postura crítica y basada en evidencia sobre los alcances y limitaciones reales de la inteligencia artificial en citogenética clínica y predicción de fenotipos.
- Integrar el marco de las 3R en la toma de decisiones experimentales cotidianas, como principio ético y como criterio de rigor científico.

3. Perfil de ingreso y perfil de egreso

3.1 Perfil de ingreso

El curso está dirigido a estudiantes de posgrado, clínicos e investigadores con formación en ciencias biológicas o biomédicas. Los requisitos previos son:

- Conocimientos básicos de biología molecular y genética: replicación, transcripción, estructura cromosómica, herencia mendeliana.
- Familiaridad con los conceptos de cariotipo humano, bandeo G y nomenclatura citogenética básica (estos temas se revisan en las sesiones del lunes y martes con la Dra. Cabrera).
- No se requiere experiencia previa en programación. Los ejercicios están diseñados para ejecutarse con un clic y con scripts completamente comentados.
- Laptop propia con navegador web actualizado (Chrome, Firefox o Edge) y conexión a internet estable.

3.2 Perfil de egreso

Al finalizar las sesiones de bioinformática, los participantes:

- Podrán consultar bases de datos genómicas para responder preguntas concretas sobre variantes cromosómicas (localización, frecuencia poblacional, clasificación clínica, genes involucrados, fenotipos asociados).
- Podrán identificar lo que hace un script básico de R y un notebook de Python (no necesariamente escribir desde cero).
- Podrán articular un flujo de trabajo bioinformático que parta de una hipótesis clínica o de una alteración citogenética y llegue a una decisión informada sobre la pertinencia de generar un modelo murino, incluyendo la exploración previa de modelos ya publicados.
- Contarán con el marco ético de las 3R como herramienta explícita para la toma de decisiones experimentales, y con la conciencia de cómo las herramientas computacionales permiten implementarlo de manera concreta.

4. Estructura general del curso

El curso se imparte de manera virtual durante la semana del 1 al 5 de junio de 2026, con sesiones diarias de 10:00 a 14:00 h. La Dra. Sandra Cabrera coordina el bloque experimental y conceptual (lunes, martes y viernes); la Dra. Yalbi Balderas-Martínez y el Ing. Luis Alberto Meza Cova coordinan el bloque computacional (miércoles y jueves, co-impartidos). La tabla a continuación resume el panorama general; el detalle de las sesiones de bioinformática se desarrolla en la sección 5.

Día	Fecha	Contenido	Responsable
Lunes	1 de junio	1. Organización del genoma humano y del genoma del ratón 2. Componentes genéticos de las enfermedades raras 3. ¿Cómo se genera un ratón genéticamente modificado 4. Ingeniería genética, DNA recombinante, estrategias para inducir deleciones o mutaciones puntuales 5. Ratones <i>knock-in</i>, <i>knock-out</i> y condicionales. Ejemplos de genes específicos y estrategias de modificación del genoma.	Dra. Sandra Cabrera
Martes	2 de junio	1. Fenotipos de ratones genéticamente modificados y su asociación con enfermedades raras. 2. Revisión de mutaciones específicas de enfermedades raras y síndromes, genes asociados y análisis y características de los fenotipos de ratón. 3. Terapia génica y estudio de caso (reparación y modificación del DNA). 4. Exploración de bases de datos de laboratorios generadores de ratones genéticamente modificados. 5. Discusión grupal	Dra. Sandra Cabrera
Miércoles	3 de junio	Bioinformática — análisis de datos (teoría). Navegadores genómicos. Variantes cromosómicas y criterios ACMG/ClinGen. Del humano al ratón con datos (MGI, IMPC). IA aplicada a citogenética: un vistazo.	Dra. Yalbi Balderas + Ing. Luis Meza
Jueves	4 de junio	Bioinformática — taller práctico y aplicaciones. R/Bioconductor con biomaRt y GenomicRanges. Python/Colab con CNN para clasificación de cromosomas. Del gen al ratón con datos. Caso integrador: las 3R en acción.	Dra. Yalbi Balderas + Ing. Luis Meza

Día	Fecha	Contenido	Responsable
Viernes	5 de junio	1. Videos y aplicaciones del diseño de organismos modificados (iGEM, protocolos especializados). 2. Aspectos bioéticos asociados al estudio de enfermedades humanas en ratones 3. Alternativas de investigación en estudios biomédicos para comprender enfermedades genéticas raras. Foro de cierre: implicaciones éticas y sociales.	Dra. Sandra Cabrera

Las sesiones de bioinformática (miércoles y jueves) están resaltadas en rosa.

5. Detalle de las sesiones de bioinformática

Las dos sesiones de bioinformática están diseñadas como una unidad pedagógica: el miércoles siembra los conceptos y las herramientas; el jueves los pone a trabajar sobre un caso integrador. La distribución de tiempos fue ajustada siguiendo dos principios pedagógicos explícitos: (1) menos temas por bloque, con tiempo real para asentar bases, ya que se trata de contenido nuevo para la mayoría de los participantes; y (2) cada concepto nuevo recibe contexto, definición, ejemplo, demo guiada, y espacio de preguntas, antes de pasar al siguiente.

5.1 Miércoles 3 de junio — Teoría y análisis de datos

Sesión teórica de 4 horas efectivas (con receso). Objetivo específico de la sesión: que los participantes sepan qué bases de datos existen, qué información contienen, cómo se consultan (manualmente, en este día), y cómo se relacionan entre sí para responder preguntas clínicas o de investigación.

Hora	Min	Bloque	Contenido	Responsable
10:00–10:20	20	Apertura	Bienvenida y objetivos del día. Recapitulación guiada de los contenidos cubiertos con la Dra. Cabrera lunes y martes (genoma humano/murino, enfermedades raras, modelos genéticamente modificados). Encuesta rápida de expectativas y nivel del grupo.	Yalbi
10:20–11:30	70	A · Navegadores genómicos	¿Qué es un navegador genómico y qué problema resuelve? Concepto de coordenadas genómicas. Traducción ISCN ↔ coordenadas (ejemplo: 46,XX,del(7)(q11.23) → chr7:72,726,578–74,142,672). Demo guiada de	Yalbi

Hora	Min	Bloque	Contenido	Responsable
			UCSC Genome Browser (cytoBand, genes, variantes). Demo guiada de Ensembl. Mini-práctica en vivo: los alumnos abren el navegador y exploran una región compartida.	
11:30–11:45	15	Receso	—	—
11:45–13:00	75	B · Variantes cromosómicas	Puente: ¿qué es una CNV y cómo se relaciona con lo que vieron en citogenética clásica? Microarreglos cromosómicos (CMA) como tecnología puente. Recorrido por DECIPHER (variantes y fenotipos). ClinVar y OMIM (evidencia clínica y correlación genotipo-fenotipo). DGV (variantes en población sana). Criterios ACMG/ClinGen para clasificación de CNVs: cómo se lee un reporte. Espacio amplio de preguntas y síntesis.	Yalbi
13:00–13:40	40	C · Del humano al ratón con datos	¿Por qué ratón? Recapitulación desde la perspectiva de datos disponibles. Mouse Genome Informatics (MGI): qué ofrece y cómo navegarlo. International Mouse Phenotyping Consortium (IMPC): catálogo de fenotipos de knockouts. Conexión explícita con los modelos y ejemplos que dio la Dra. Cabrera.	Luis
13:40–13:55	15	D · IA aplicada: un vistazo	Demostración sin código de sistemas de IA aplicados a citogenética: HiBand (ASI), Ikaros v7 (MetaSystems), Karyotype AI con Vision Transformers. Un cariotipo real, la IA lo clasifica. Discusión: qué resolvió (velocidad, reproducibilidad) y qué no (casos ambiguos, variantes raras, juicio clínico). Gancho para el taller del jueves.	Yalbi
13:55–14:00	5	Cierre	Síntesis del día. Qué haremos mañana. Preguntas pendientes registradas para retomar al inicio del jueves.	Yalbi

5.2 Jueves 4 de junio — Taller práctico y aplicaciones

Sesión práctica de 4 horas efectivas (con receso). Objetivo específico de la sesión: que los participantes ejecuten, con sus propias manos, un flujo bioinformático completo que parta de una alteración cromosómica y llegue a una decisión informada sobre la pertinencia de un modelo murino, integrando IA para el análisis de imágenes y datos públicos de fenotipos murinos.

Hora	Min	Bloque	Contenido	Responsable
10:00–10:30	30	Apertura + setup	Bienvenida y retoma de preguntas del miércoles. Setup paso a paso de Posit Cloud (login, apertura del proyecto pre-configurado, ejecución del primer chunk para verificar entorno). Setup de Google Colab (login con cuenta Google, apertura del notebook). Verificación individualizada: no avanzamos hasta que todos tienen entorno funcional.	Yalbi
10:30–11:30	60	A · R/Bioconductor	¿Por qué acceso programático a bases de datos? Ventajas sobre la consulta manual. Ejercicio 1 (25 min): consulta de genes en una región cromosómica con biomaRt, anotación con GO terms, identificación de genes asociados a enfermedades. Ejercicio 2 (25 min): manipulación de regiones con GenomicRanges, visualización con karyoploteR. Síntesis (10 min).	Yalbi
11:30–11:45	15	Receso	—	—
11:45–12:30	45	B · IA para cariotipo	¿Qué es una red neuronal convolucional (CNN)? Explicación con analogías visuales, sin prerequisites matemáticos (10 min). Demo ejecutable: notebook pre-configurado en Colab que aplica un modelo pre-entrenado sobre imágenes del dataset abierto de Kuo et al. (2023) (20 min). Discusión crítica: dónde brilla la IA, dónde se equivoca, qué implica para la práctica clínica (15 min).	Yalbi
12:30–13:15	45	C · Del gen al ratón con datos	Ejercicio 3 (20 min): obtención de ortólogos humano-ratón con biomaRt para un conjunto de genes de interés clínico. Ejercicio 4 (25 min): consulta programática de IMPC para identificar qué fenotipos están ya reportados en ratones knockout de esos genes. Resultado concreto: una lista priorizada de genes para los cuales un nuevo modelo murino sería redundante, vs. genes donde aún hay preguntas abiertas.	Luis
13:15–13:55	40	Caso integrador	Paciente hipotético con alteración cromosómica detectada por CMA. En parejas, los participantes siguen el flujo completo: ISCN → coordenadas → DECIPHER/ClinVar → genes afectados → ortólogos humano-ratón → IMPC → decisión: ¿necesitamos generar un nuevo modelo murino o podemos responder la pregunta con datos públicos? Puesta en común.	Yalbi + Luis

Hora	Min	Bloque	Contenido	Responsable
13:55–14:00	5	Cierre	Síntesis de las 3R en acción: qué ejercicios del día ilustraron Reducción, Reemplazo y Refinamiento. Puente al foro del viernes con la Dra. Cabrera. Recursos para seguir aprendiendo.	Yalbi

Nota: el notebook de Python y el script de R estarán pre-configurados. Los participantes no escribirán código desde cero; leerán, modificarán parámetros y ejecutarán celdas con guía. Este diseño permite que personas sin experiencia previa en programación completen los ejercicios aun sin experiencia, y al mismo tiempo deja espacio para que quienes tengan experiencia exploren modificaciones.

6. Metodología didáctica

El curso combina cuatro estrategias didácticas que se alternan a lo largo de las sesiones:

1. Exposición magistral breve. Cada concepto nuevo se introduce con una exposición de 10–15 minutos, apoyada en diapositivas con ejemplos concretos. Nunca excedemos 20 minutos de exposición continua sin interacción.
2. Demostración guiada. Después de cada exposición, el ponente demuestra en vivo cómo usar la herramienta correspondiente, proyectando la pantalla y explicando cada paso. Los participantes observan y toman notas.
3. Práctica en vivo sincronizada. Los participantes replican la demostración en su propia laptop, con una región cromosómica o un gen compartido por todos. El ponente recorre el aula y resuelve dudas individuales.
4. Discusión y síntesis. Al cierre de cada bloque, se dedican 5–10 minutos a preguntas abiertas y a una síntesis breve del ponente que conecta el bloque con el anterior y anticipa el siguiente.

Como hilo narrativo atraviesa el curso un caso integrador: un paciente hipotético con sospecha de enfermedad genética rara detectada por microarreglo cromosómico. Al inicio del miércoles se presenta; a lo largo de ambas sesiones se va enriqueciendo con la información obtenida de las distintas bases de datos; al final del jueves los participantes toman una decisión basada en evidencia sobre cómo continuar con la investigación del caso, incluyendo la pregunta de si un nuevo modelo murino aporta valor o si la evidencia existente es suficiente.

7. Recursos y herramientas

Todas las herramientas utilizadas en el curso son de acceso libre y gratuito. Este principio es pedagógicamente importante: garantiza que los participantes puedan seguir explorándolas después del curso, sin barreras económicas ni institucionales.

Categoría	Recurso	Qué aporta al curso
Navegadores genómicos	UCSC Genome Browser	Visualización de regiones, bandas citogenéticas, genes, variantes. Tabla cytoBand para ISCN.
Navegadores genómicos	Ensembl	Anotación genómica, genes, transcritos, variantes. Base de biomaRt.
Variantes (humano)	DECIPHER	Base de datos de variantes estructurales y síndromes asociados.
Variantes (humano)	ClinVar	Variantes de relevancia clínica con evidencia interpretativa.
Variantes (humano)	OMIM	Correlación genotipo-fenotipo y enfermedades mendelianas.
Variantes (humano)	DGV (Database of Genomic Variants)	Variantes estructurales en población control.
Modelo murino	MGI (Mouse Genome Informatics)	Anotación del genoma del ratón, ortólogos, fenotipos.
Modelo murino	IMPC (International Mouse Phenotyping Consortium)	Catálogo de fenotipos de knockouts murinos.
Plataforma R	Posit Cloud (posit.cloud)	Entorno R/RStudio preconfigurado, gratuito, sin instalación local.
Paquetes R	biomaRt, GenomicRanges, karyoploteR, AnnotationHub	Acceso programático a Ensembl y manipulación de regiones.
Plataforma Python	Google Colab	Notebook Python pre-configurado con GPU opcional, sin instalación local.
Dataset IA	Annotated Metaphase Cell Images (Kuo et al., Sci Data 2023)	5,000 imágenes anotadas de metafases, 24 clases cromosómicas.
Criterios clínicos	ACMG/ClinGen (Riggs et al. 2020)	Framework estandarizado para clasificación de CNVs.
Complemento tejido	Human Protein Atlas	Expresión proteica por IHC en tejidos normales y patológicos.

8. Requisitos técnicos

8.1 Equipo

- Laptop propia (no se proporciona equipo de cómputo).
- Navegador web actualizado: Chrome, Firefox o Edge (recomendado: la versión estable más reciente al momento del curso).
- Conexión a internet estable por Wi-Fi en la sede.
- Opcional: mouse externo (facilita la navegación en los navegadores genómicos).

8.2 Cuentas necesarias (gratuitas, creadas antes del curso)

- Cuenta en Posit Cloud (posit.cloud). El registro es gratuito, requiere un correo electrónico.
- Cuenta Google activa (para Google Colab). Si ya usa Gmail, Drive o Calendar, está cubierto.

8.3 Software local (opcional)

No se requiere instalar nada de manera local. Posit Cloud y Google Colab cubren todo el entorno. Como alternativa para quienes prefieran trabajar localmente:

- R $\geq$ 4.3 con RStudio Desktop.
- Paquetes de Bioconductor: BiocManager, biomaRt, GenomicRanges, AnnotationHub, karyoploteR.
- Python $\geq$ 3.10 con pip y los paquetes del notebook (se distribuye un requirements.txt).

9. Preparación previa de los participantes

Una semana antes del curso se envía a los inscritos un correo con la siguiente lista de verificación. La idea es que lleguen el miércoles con el entorno ya funcionando, para no perder los primeros 30 minutos de la sesión en problemas de login.

9.1 Lista de verificación (checklist)

- Crear cuenta gratuita en Posit Cloud (posit.cloud) y verificar acceso con el enlace al proyecto compartido que se enviará por correo.
- Confirmar acceso a una cuenta Google activa y abrir el notebook de Colab de prueba enviado por correo.
- Verificar que la laptop tiene al menos 2 GB de RAM libres y una versión reciente del navegador.
- Completar una encuesta breve de expectativas y nivel (Google Forms, 3 minutos).

9.2 Lecturas recomendadas

- Rosenblum LS et al. (2025). The Emergence of Artificial Intelligence-Guided Karyotyping: A Review and Reflection. Genes, 16(6):685. Review accesible, buen punto de partida para el bloque de IA.

- Riggs ER et al. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: ACMG and ClinGen joint consensus. *Genet Med*, 22(2):245–257. Marco del Bloque B del miércoles.
- Russell WMS, Burch RL (1959). *The Principles of Humane Experimental Technique*. Capítulos introductorios sobre las 3R. Lectura opcional pero recomendada para aterrizar el marco ético del curso.

10. Evaluación

La evaluación tiene un peso formativo predominante. El objetivo no es filtrar sino asegurar que cada participante se lleve aprendizajes accionables. La calificación final se emite como acreditado / no acreditado, con retroalimentación individualizada al reporte del caso integrador.

10.1 Rúbrica

Criterio	Ponderación	Descripción
Asistencia y participación	20 %	Asistencia a las dos sesiones de bioinformática. Participación activa en demos y discusiones.
Ejercicios en sesión	30 %	Finalización en vivo de los cuatro ejercicios del jueves (biomaRt, GenomicRanges, CNN en Colab, ortólogos/IMPC). Evidencia: capturas guardadas en el proyecto de Posit Cloud.
Caso integrador	50 %	Reporte breve (máximo 2 páginas) con la respuesta al caso presentado el jueves: flujo seguido, evidencia encontrada, decisión tomada respecto al modelo murino, y justificación explícita en términos de las 3R. Entrega hasta el viernes por la mañana.

10.2 Criterios mínimos de acreditación

- Asistencia ≥ 80 % a las dos sesiones de bioinformática.
- Finalización de al menos 3 de los 4 ejercicios en sesión.
- Entrega del reporte del caso integrador con decisión argumentada y mención explícita del marco 3R.

11. Entregables y material de apoyo

El material de apoyo se distribuye antes del curso a través del proyecto compartido en Posit Cloud y de un repositorio GitHub con licencia abierta (MIT o CC BY 4.0). La licencia abierta es una elección

deliberada: permite a los participantes reutilizar el material en sus propias clases, grupos de estudio e investigación, alineado con el principio de acceso abierto que enmarca el curso.

11.1 Material preparado previamente

- Presentación en PDF de la sesión del miércoles (bloques A, B, C, D).
- Proyecto de R/Posit Cloud con cuatro scripts comentados paso a paso (uno por ejercicio del jueves).
- *Notebook* de Google Colab con el ejercicio de IA para cariotipo, pre-configurado con el *dataset* abierto.
- *Dataset* de ejemplo: diez regiones cromosómicas de interés clínico (CNVs reportadas), en formato BED y como tabla CSV con anotaciones.
- Hoja de referencia rápida (cheat sheet) de URLs de las bases de datos y funciones clave de biomaRt y GenomicRanges (PDF).
- Plantilla del reporte del caso integrador (Word o Markdown).

11.2 Material que se entrega a los participantes

- Acceso al proyecto compartido en Posit Cloud y al *notebook* de Colab.
- Presentación y *cheat sheet* en PDF.
- Lista de referencias con DOI clicable para consulta posterior.
- Retroalimentación personalizada al reporte del caso integrador (dentro de las dos semanas posteriores al curso).

12. Bibliografía

12.1 Bases de datos y herramientas bioinformáticas

1. Cunningham F et al. (2025). Ensembl 2025. Nucleic Acids Research.
2. Firth HV et al. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet, 84(4):524–533.
3. Durinck S et al. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics, 21(16):3439–3440.
4. Lawrence M et al. (2013). Software for Computing and Annotating Genomic Ranges. PLoS Comput Biol, 9(8):e1003118.
5. Gel B, Serra E (2017). karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics, 33(19):3088–3090.
6. Blake JA et al. (2021). Mouse Genome Database (MGD): Knowledgebase for mouse–human comparative biology. Nucleic Acids Research, 49(D1):D981–D987.

7. Groza T et al. (2023). The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping. *Nucleic Acids Research*, 51(D1):D1038–D1045.

12.2 Criterios clínicos

8. Riggs ER et al. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the ACMG and ClinGen. *Genet Med*, 22(2):245–257.

12.3 IA aplicada a citogenética

9. Rosenblum LS, Holmes J, Taghiyev AF (2025). The Emergence of Artificial Intelligence-Guided Karyotyping: A Review and Reflection. *Genes*, 16(6):685.
10. Tabassum S et al. (2025). Automated karyogram analysis for early detection of genetic and neurodegenerative disorders. *Front Comput Neurosci*, 18:1525895.
11. Kuo CE et al. (2023). An Open Dataset of Annotated Metaphase Cell Images for Chromosome Identification. *Sci Data*, 10:104.
12. Shamsi Z et al. (2025). Karyotype AI for Precision Oncology. *arXiv:2211.14312v5*.

12.4 Marco ético y las 3R

13. Russell WMS, Burch RL (1959). *The Principles of Humane Experimental Technique*. Methuen, London. (Referencia fundacional de las 3R.)
14. Tannenbaum J, Bennett BT (2015). Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose. *J Am Assoc Lab Anim Sci*, 54(2):120–132.
15. Graham ML, Prescott MJ (2015). The multifactorial role of the 3Rs in shifting the harm–benefit analysis in animal models of disease. *Eur J Pharmacol*, 759:19–29.

13. Anexos

Anexo A · Cuerpo docente

Dra. Yalbi Itzel Balderas-Martínez. Investigadora en Ciencias Médicas “C”; Titular del Laboratorio de Biología Computacional, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER). SNI Nivel

Ing. Luis Alberto Meza Cova. Miembro del Laboratorio de Biología Computacional del INER. Experiencia en análisis bioinformático con R/Bioconductor (Seurat, DESeq2, scater). Especialista en el modelo murino aplicado a enfermedades respiratorias.

Anexo B · Responsabilidades y coordinación

La Dra. Balderas-Martínez es responsable del diseño pedagógico de las dos sesiones de bioinformática, de la coordinación con la Dra. Sandra Cabrera, de la preparación del material de apoyo (*scripts* de R, *notebook* de Python, *cheat sheet*, *dataset*) y de los bloques de navegadores genómicos, variantes cromosómicas e IA aplicada. El Ing. Meza Cova es responsable de los bloques sobre modelo murino y genómica comparativa (MGI, IMPC, ortólogos) y co-imparte el caso integrador del jueves.

Anexo C · Seguridad, integridad de datos y buenas prácticas

En coherencia con los estándares de seguridad y rigor que rigen el Laboratorio de Biología Computacional del INER, el curso adopta las siguientes prácticas explícitas:

- Todo el código distribuido es de lectura pública y está alojado en un repositorio con licencia abierta, revisable por los participantes. No se distribuyen binarios opacos.
- Las cuentas en Posit Cloud y Google son de los propios participantes; no se comparten credenciales ni se usan cuentas institucionales compartidas.
- Los *datasets* utilizados son públicos, libres de información identificable. No se manejan datos clínicos reales de pacientes; el caso integrador es un caso hipotético construido a partir de variantes ya publicadas.
- Los navegadores se usan con conexión HTTPS en todo momento; las URLs se verifican contra las publicadas en la documentación oficial de cada base de datos.
- Las notas, scripts y notebooks se distribuyen por canales verificables (correo institucional, repositorio GitHub oficial del curso).

Anexo D · Glosario mínimo

CNV: copy number variant. Variante de número de copias; segmento genómico ganado o perdido respecto al genoma de referencia.

CNN: convolutional neural network. Red neuronal convolucional; arquitectura de aprendizaje profundo particularmente adecuada para el análisis de imágenes.

ISCN: International System for Human Cytogenomic Nomenclature. Sistema internacional de nomenclatura citogenética humana.

ACMG/ClinGen: American College of Medical Genetics and Genomics / Clinical Genome Resource. Instituciones que emiten los estándares para clasificación de variantes.

MGI: Mouse Genome Informatics. Base de datos integral del genoma del ratón.

IMPC: International Mouse Phenotyping Consortium. Consorcio que caracteriza fenotipos de knockouts murinos a gran escala.

3R: Replace, Reduce, Refine. Marco ético propuesto por Russell y Burch (1959) para la experimentación con animales.