

Aspectos generales

Título:	Aproximaciones Computacionales para el Análisis Ómico: Fundamentos y Aplicaciones
Programas de posgrado o planes de estudio en donde se ofertará adicionalmente:	
Posgrado de ciencias bioquímicas y ciencias biológicas	
Área del conocimiento:	Genética, genómica y bioinformática
Semestre:	2026-1
Modalidad:	Tópico selecto
Horario:	miércoles de 10:00 a 14:00
No. sesiones:	13
Horas por sesión:	4.0
Total alumnos PDCB:	15
Total alumnos:	15
Videoconferencia:	Si
Lugar donde se imparte:	Instituto de investigaciones Biomédicas, sede exterior, UNAM CU
Informes:	sromero@iibiomedicas.unam.mx

Métodos de evaluación

MÉTODO	PORCENTAJE	NOTAS
Participación en clase	50%	
Presentación de trabajo final	50%	

Contribución de este curso/tópico en la formación del alumnado del PDCB:

Este curso proporciona al alumno del PDCB una base sólida en los principios, estrategias y herramientas fundamentales para el análisis computacional de datos ómicos, un área clave en la investigación biomédica contemporánea. A través de un enfoque teórico-práctico, el alumno desarrollará habilidades para el diseño de experimentos, la manipulación y análisis de datos de alto volumen, y la interpretación crítica de resultados derivados de tecnologías masivas. Además, adquirirá competencias en el uso de plataformas públicas, estrategias bioinformáticas y herramientas de inteligencia artificial aplicadas al análisis ómico, fortaleciendo su capacidad de integrar información multiescala en proyectos de investigación. Este curso busca formar profesionales capaces de enfrentar los retos del análisis de datos biomédicos, con una visión actualizada, crítica y ética del impacto de la genómica y otras disciplinas ómicas en el avance científico.

Profesor (a) responsable

Nombre:	Romero Córdoba Sandra Lorena
Teléfono:	(55) 54 87 09 00 Ext. 2006
Email:	sromero@iibiomedicas.unam.mx

Profesores (as) participantes

PARTICIPANTE	ENTIDAD O ADSCRIPCIÓN	SESIONES
ROMERO CÓRDOBA SANDRA LORENA Responsable	Instituto de Investigaciones Biomédicas	análisis de microarreglos con herramientas públicas Análisis dedicados Análisis downstream de transcriptómica Análisis downstream de variantes Introducción Presentación de proyectos finales

SALIDO GUADARRMA ALBERTO IVAN
Integrante

Instituto de Perinatología

Introducción al análisis de datos transcriptómicos usando RNA-seq
Introducción al análisis de datos transcriptómicos usando RNA-seq II (Aplicaciones de IA y modelos LLM en el análisis multi-ómico)
Descubrimiento de variantes de un solo nucleótido (SNP) y pequeñas indels
Descubrimiento de variantes de un solo nucleótido (SNP) y pequeñas indels II
Manejo de datos de NGS
Uso de la plataforma bioinformática Galaxy

Introducción

El análisis multi-ómico se ha consolidado como una estrategia esencial para caracterizar de manera integral los procesos biológicos, ofreciendo una comprensión más profunda de fenómenos como el desarrollo tumoral y otros estados patológicos y fisiológicos. En este curso teórico-práctico, adquirirás una visión actualizada de los conceptos fundamentales, las estrategias de análisis computacional y el estado del arte en la generación e interpretación de datos ómicos. A lo largo de las sesiones, aprenderás sobre el diseño de experimentos, el uso de tecnologías masivas de análisis de datos, y el aprovechamiento de herramientas públicas ampliamente empleadas en la comunidad científica.

Trabajarás con los principales fundamentos para la construcción de pipelines computacionales, comprendiendo su lógica interna, los criterios de calidad de los datos y los puntos críticos donde pueden surgir sesgos o errores en el análisis. Para fortalecer una perspectiva integral, utilizarás herramientas de uso libre, basadas en entornos gráficos, que facilitarán el procesamiento y visualización de los datos. Al finalizar el curso, contarás con una perspectiva crítica y sólida sobre el análisis multi-ómico, sus alcances, limitaciones, y las herramientas de biología computacional que sustentan el estudio contemporáneo de los sistemas biológicos.

Temario

Módulo 1: Introducción al análisis ómico y uso de bases públicas

En este módulo conocerás los fundamentos de la genómica aplicada a la investigación biomédica, el impacto de las tecnologías masivas y el acceso a repositorios públicos de datos ómicos.

1. Introducción (4 h, Sandra Romero 5h)

- ¿Qué es un genoma y cuál es su relevancia en la investigación biomédica?
- Aplicaciones de los genomas para entender procesos normales y patológicos.
- Predicción de enfermedades a partir de datos genómicos.
- Tecnologías masivas: principios de secuenciación de nueva generación, microarreglos, Nanostring, PCR digital y análisis de célula única.
- Introducción a repositorios públicos: GEO, TCGA (GDAC, cBioPortal, UCSC Xena), ICGC, Genome Browser.
- Tipos de datos, niveles de accesibilidad y requisitos para su obtención.
- Exploración de bases de datos especializadas.

2. Análisis de microarreglos con herramientas públicas (5h, Sandra Romero)

- Principios de normalización de datos de expresión.
- Análisis de expresión diferencial.
- Visualización de resultados.
- Introducción práctica a GenePattern y BART para el procesamiento de datos.

Módulo 2: Análisis computacional básico de datos ómicos

Este módulo introduce plataformas bioinformáticas accesibles y procedimientos fundamentales para el manejo y procesamiento de datos de secuenciación masiva.

3. Uso de la plataforma bioinformática Galaxy (5 h, Ivan Salido)

- Introducción a la interfaz de Galaxy.
- Preparación, carga y manejo de archivos de datos ómicos.
- Ejecución de herramientas y módulos de análisis bioinformático.
- Gestión de historias y flujos de trabajo.
- Visualización y exportación de resultados.

4. Manejo de datos de NGS (5 h, Ivan Salido 3h, Sandra Romero 2h)

- Exploración de archivos FASTQ y evaluación de calidad de secuencias.
- Mapeo de lecturas genómicas contra un genoma de referencia.
- Análisis de profundidad de secuenciación.
- Remoción de duplicados, recalibración y filtrado de lecturas.
- Manejo de archivos BAM y visualización de alineamientos.
- Procesamiento de datos de secuenciación Pair-end y Single-end.

5. Introducción al análisis de datos transcriptómicos usando RNA-seq I (4 h, Ivan Salido)

- Fundamentos de la transcriptómica.
- Preparación de archivos para análisis.
- Evaluación de la calidad de secuenciación mediante archivos FASTQ.

- Preprocesamiento de lecturas (trimming) y mapeo de lecturas.

6. Introducción al análisis de datos transcriptómicos usando RNA-seq II (4 h, Ivan Salido 3h, Sandra Romero 2h)

- Detección y cuantificación de transcritos.
- Análisis de expresión diferencial.
- Introducción al análisis de enriquecimiento génico.

7. Análisis downstream de transcriptómica (5 h, Sandra Romero)

- Análisis de enriquecimiento de vías moleculares.
- Implementación de análisis GSEA.
- Visualización de resultados con Volcano plots y Heatmaps.

Módulo 3: Análisis de variantes genéticas

Este módulo profundiza en la detección e interpretación de variantes genómicas a partir de datos de NGS.

8. Descubrimiento de variantes de un solo nucleótido (SNP) y pequeñas indels I (5 h, Ivan Salido)

- Principios básicos del análisis de SNPs e indels.
- Llamado de variantes en genomas diploides.

9. Descubrimiento de variantes de un solo nucleótido (SNP) y pequeñas indels II (5 h, Ivan Salido)

- Identificación de variantes germinales y somáticas.
- Exploración y manipulación de archivos VCF.
- Anotación funcional de variantes.

10. Análisis downstream de variantes (5 h, Sandra Romero)

- Visualización de variantes en IGV.
- Identificación de mutaciones significativas.
- Análisis de firmas mutacionales.
- Introducción al análisis de alteraciones genómicas mediante GISTIC.

Módulo 4: Análisis avanzado de datos ómicos

Este módulo ofrece una aproximación a metodologías especializadas para el análisis de datos de interacción proteína-DNA y de infiltración inmune en tumores.

11. Métodos bioinformáticos para el análisis de datos de ChIP-seq (5 h, Ivan Salido)

- Principios y estrategias de análisis de datos ChIP-seq.

12. Análisis dedicados (5 h, Sandra Romero)

- Aplicación de métodos de clustering en datos transcriptómicos.
- Estudio de infiltración inmune en tumores mediante herramientas de deconvolución.
- Visualización dedicada de datos ómicos.
- Análisis práctico en plataformas cBioPortal y Xena.

Módulo 5: Integración de inteligencia artificial en el análisis ómico

En este módulo se explorarán herramientas de inteligencia artificial y modelos de lenguaje (LLM) aplicados a la optimización de diseños experimentales y estrategias computacionales en el análisis de datos ómicos.

13. Aplicaciones de IA y modelos LLM en el análisis multi-ómico (5 h, Sandra Romero 2.5 h, Ivan Salido Guadarrama 2.5 h)

- Introducción al uso de IA en la bioinformática ómica.
- Modelos de lenguaje (LLM) para apoyo en el diseño experimental y análisis.
- Automatización y mejora de flujos de trabajo bioinformáticos.
- Potencial, limitaciones y buenas prácticas en la integración de IA en proyectos de análisis de datos.

Módulo 6: Presentación de proyectos finales

En este módulo, los participantes integrarán los conocimientos adquiridos a través de la elaboración y presentación de un proyecto de análisis de datos ómicos.

Bibliografía

Afgan, E.; Baker, D.; van den Beek, M.; Blankenberg, D.; Bouvier, D.; ?ech, M.; Chilton, J.; Clements, D.; Coraor, N.; Eberhard, C.; Gru?ning, B.; Guerler, A.; HillmanJackson, J.; Von Kuster, G.; Rasche, E.; Soranzo, N.; Turaga, N.; Taylor, J.; Nekrutenko, A.; Goecks, J. (8 July 2016). "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update". Nucleic Acids Research. 44 (W1): W3–W10. doi:10.1093/nar/gkw343. Afgan, E.; Baker, D.; vanden Beek, M.; Blankenberg, D.; Bouvier, D.; ?ech, M.; Chilton, J.; Clements, D.; Coraor, N.; Eberhard, C.; Gru?ning, B.; Guerler, A.; Hillman-Jackson, J.; Von Kuster, G.; Rasche, E.; Soranzo, N.; Turaga, N.; Taylor, J.; Nekrutenko, A.; Goecks, J. (8 July 2016). "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update". Nucleic Acids Research. 44 (W1): W3–W10. doi:10.1093/nar/gkw343.

Observaciones

Los alumnos deberán traer computadora ala clase para realizar los ejercicios. Solo se permitirá formato ZOOM para aquellos alumnos que están fuera de Cd Mx