



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Comisión de Carrera de Ciencias Biológicas

<http://cccbfcen.wixsite.com/cccb>

Int. Güiraldes 2620

Ciudad Universitaria - Pab. II, 4º Piso

CPA: C1428EHA, Ciudad Autónoma de Buenos Aires
 ARGENTINA.

☎: +54 11 4576-3349 / 5285-8665

I

Asignatura: GENÓMICA APLICADA

Carrera: Licenciatura en Ciencias Biológicas	Código de la carrera: 05
	Código de la asignatura:
CARÁCTER:	Tache lo que no corresponde
Curso obligatorio de licenciatura (plan 2019)	NO/SI
Curso electivo/optativo de licenciatura (plan 2019)	Electivo/ Optativo

Duración de la asignatura (en semanas)	6
Cuatrimestre(s) en que dicta (indicar cuatrimestre o verano):	verano
Frecuencia en que se dicta (cuatrimestral, anual, bianual, etc.)	anual

ACTIVIDAD	Horas semanales	Número de semanas	Horas totales
Teóricas	12	6	72
Problemas	0	0	0
Laboratorios	12	6	72
Seminarios	8	2	16
Teórico- prácticos o Teórico- problemas			
Si corresponde, especifique las horas de otras actividades (salidas de campo, etc.)			
Carga horaria semanal máxima	32		
Carga horaria semanal mínima	24		
Carga horaria total:	160		

Asignaturas correlativas:	<u>Biología Molecular o Genética Molecular</u>
Forma de Evaluación:	Examen Domiciliario o Regular/ Resolución de un problema bioinformático basado en datos propios o provistos por los docentes/ Presentación de Seminario

OBJETIVOS II

Módulo de Genómica Estructural:

La era de la genómica no terminó aún. Surgieron nuevas tecnologías de secuenciación genómica que aportan nuevas ventajas en la profundidad y cobertura de las lecturas de fragmentos genómicos (por ej. Illumina), mientras otras avanzan sobre la longitud de los fragmentos secuenciados (por ej. Minlon y PacBio que permiten secuenciar un genoma bacteriano en un día) a expensas de las ventajas de los primeros métodos. Cada una de ellas requiere de programas de análisis bioinformático específico para realizar el ensamblado genómico y se requiere que los alumnos se familiaricen con su uso y con el manejo de las variables específicas de cada uno. Esto se complica cuando la secuenciación comprende a varios organismos de distintas especies como ocurre en la metagenómica. Por ello, el módulo de genómica estructural profundiza en estos aspectos específico de análisis de secuencias y su interpretación. Por otro lado, otro impacto importante lo constituye el uso de la secuenciación genómica para la genotipificación (o genotipado) de individuos en un contexto poblacional. La genotipificación por secuenciación (GBS genotyping by sequencing) tiene aplicaciones como: descubrimiento de marcadores moleculares de tipo SNP y microsatélites de uso en genética de poblaciones, aislamiento e identificación de polimorfismos en genes funcionales representativos del genoma, identificación de individuos, poblaciones y especies, código de barras genético (barcoding), GWAS, etc permitiendo realizar estudios genómicos en especies que han sido secuenciadas y poseen un genoma de referencia y, más importante, en especies que no han sido secuenciadas hasta ahora. Todas estas aplicaciones tienen que analizarse con programas bioinformáticos específicos.

Módulo Genómica Funcional:

Por otro lado, ingresamos en la era postgenómica, con nuevas expresiones lingüísticas como perfiles de expresión "de escala o amplitud genómica" (genome-wide) a niveles de transcriptoma, proteoma, metaboloma y, el más reciente y complejo, fenoma que debería ser el resultado de la acción de los anteriores. Cada una de estas disciplinas postgenómicas tienen sus programas bioinformáticos específicos. Más importante aún, la interpretación de las bases de datos generadas requiere integración para poder arribar a interpretaciones biológicas significativas. La biología o genética de sistemas provee otros programas bioinformáticos que integran datos de diferente origen y característica (cuantitativos, cualitativos, discretos) que requieren de conocimientos en bioestadística genómica además de informática. El módulo de Genómica Funcional tiene como objetivo familiarizar al alumno con estas aproximaciones experimentales y su interpretación biológica.

General:

Enseñar los métodos y modos de razonamiento propios de la investigación genómica, aplicada tanto a resolver problemas de investigación científica como también tecnológica.

Establecer la discusión crítica de resultados de las aplicaciones más relevantes en este campo.

Brindar ejemplos de hechos recientes relacionados con la Genómica que tengan impacto económico-social.

Estimular el pensamiento reflexivo acerca del estado del conocimiento en los temas de la materia.

Las teóricas ilustrarán cómo desde el conocimiento biológico se pueden diseñar experimentos para abordar la resolución de problemas o hipótesis e ilustrar diferentes formas de analizar estos datos. Las prácticas consistirán en ejercicios de computación que permitan a los participantes a aplicar métodos estadísticos para el análisis de los datos bajo la guía de los profesores y docentes auxiliares.

CONTENIDOS MÍNIMOS (ya aprobados Anexo IV Plan 2019)

Genómica estructural: mapeo físico de genomas, genotificado por secuenciación (GBS) y mapeo por asociación (GWAS), citogenómica (FISH y GISH). Secuenciación y resecuenciación de genomas (Sanger, NGS y posteriores). Bioinformática genómica. Metagenómica funcional y caracterización taxonómica de biodiversidad en ecosistemas. Genómica funcional (del genoma al fenoma): Transcriptómica, RNAseq y micromatrices. Metabolómica, interactómica y otras X-ómicas. Interpretación y relacionamiento de datos. Fenómica y genética cuantitativa relacionada. Biología de Sistemas. Aplicaciones.

PROGRAMA ANALÍTICO

I. Módulo de Genómica Estructural:

1. SECUENCIACION GENOMICA

Aspectos metodológicos de la secuenciación a gran escala, principios de funcionamiento de los secuenciadores automáticos, lectura e interpretación de los resultados, creación de bases de datos y análisis informático de los mismos, principales programas informáticos.

Secuenciación de primera, segunda (NGS) y tercera generación (nanopore). Whole-Genome Shotgun (WGS). Envío de información a bancos de secuencias: obtención de número de entrada (*accession number*) y registro de la secuencia. Estadística y programas informáticos involucrados. GeneBank y utilización de recursos *via* INTERNET (Blast, Phred, Phrap) para validación, curado de secuencias y búsqueda de homologías y solapamientos con otras secuencias nucleotídicas ya ingresadas al banco.

2. GENOTIPIFICACIÓN POR SECUENCIACIÓN

Marcadores moleculares clásicos. SNPs y microsatélites ¿Qué son? Usos y aplicaciones más frecuentes. Uso de NGS para identificación de SNPs y microsatélites. Uso de GBS/RAD para análisis de datos (Stack). La cuestión de disponibilidad (o no) de un genoma de referencia. ¿Cómo se trabaja en cada caso? Genotipificación de un cruzamiento genético. Genómica de poblaciones incluidos el mapeo genético y de QTL, asociación y selección genómica y estudios filogeográficos. Aspectos metodológicos básicos. Análisis de datos.

3. METAGENÓMICA (DIVERSIDAD)

Ecología de comunidades microbianas (abundancia, riqueza, equitatividad, diversidad, composición, estructura). Métodos moleculares dependientes e independientes de cultivo y ciencias ómicas. Análisis de un único gen, genes marcadores, fingerprinting, análisis de genes totales. Métodos basados en genes ribosomales: 16S. ITS. Biblioteca de amplicones vs. secuenciación a gran escala. Diseño experimental. Asignación de OTUs. Parámetros y estimadores ecológicos. Bases de datos de referencia (Greengenes, Silva, etc.). Plataformas de análisis (Mothur, QIIME2). El ejemplo de la caracterización de microorganismos del suelo. Taxonomía molecular. Confección de catálogos de diversidad genética.

4. METAGENÓMICA II (ENSAMBLADO, BINNING Y ANOTACIÓN DE METAGENOMAS)

Secuenciación directa de ADN en muestras complejas. Whole-Genome Shotgun (WGS). Next-Generation Sequencing (NGS). Ensamblado de secuencias. Estadísticas (cobertura, N50). Anotación. Ensamblado de un genoma a partir de un metagenoma. Cobertura

diferencial. Binning independiente de la composición. Programas y software (GroopM, MaxBin, MetaBat). Análisis de calidad de bins (BUSCO, CheckM).

5. METAGENÓMICA III (BIOPROSPECCIÓN DE ENZIMAS)

Búsqueda de recursos genéticos y bioquímicos de valor comercial. Enzimas microbianas. Asociación de función enzimática y ambiente microbiano. Usos en la industria. Características deseadas. Estrategias de bioprospección. Cultivos bacterianos. Microorganismos no cultivables (ADN total), metagenómica. Secuencia vs. función. Desafíos del descubrimiento a la aplicación. Ejemplos: usos de metilhaluros, biomasa lignocelulósica como sustrato para biocombustibles. Análisis de resultados de secuenciación metagenómica.

6. BIOINFORMÁTICA GENÓMICA

Introducción teórico-práctica a Linux y R. Utilización práctica de herramientas informáticas disponibles para el análisis de problemas clásicos, análisis de datos de secuenciación por NGS (secuenciación de nueva generación) y Nanopore; con y sin genoma de referencia, genotipificación y mapeo con marcadores moleculares SSR, SNP y GBS (genotyping by sequencing), metagenómica/diversidad genómico-taxonómica de microbiota. Ensamblado de secuencias. Anotación. Ensamblado. Cobertura. Programas y software.

II. Módulo de Genómica Funcional

7. TRANSCRIPTÓMICA

Contrucción de librerías de cDNA para su secuenciación por NGS y curado bioinformático. RNAseq: alineación, manipulación y visualización de datos, evaluación de la calidad de los datos, la expresión diferencial y el análisis estadístico, utilizando los programas Bowtie, Tophat y R/Bioconductor. ¿Cómo se trabaja con datos crudos? Mapeo de lecturas (“reads”) usando un genoma de referencia. Manejo de los datos alineados: Uso de archivos SAM/BAM. Manejo de los datos alineados: Normalización de datos, expresión diferencial, uso de codones. Almacenamiento y visualización de datos con la herramienta: ATGC

8. PROTEÓMICA

Generación de patrones de electroforesis bidimensional. Análisis de datos. Microsecuenciación de aminoácidos. Estudio de estructura de proteínas y su predicción bioinformática. Búsqueda de los genes correspondientes en el genoma e integración de datos.

9. METABOLÓMICA

Estudio y caracterización de metabolitos en muestras biológicas. Principios de cromatografía gaseosa, líquida aplicada y espectrometría de masa a la generación de perfiles metabólicos. Tipos de metabolitos detectados. Su informatización para la creación y uso de una base de datos. Aplicaciones prácticas.

10. FENÓMICA

Caracteres de distribución continua (“cuantitativos”) y su variación genética heredable. Aproximaciones experimentales genotipo-fenotipo como “common garden experiments” y otras. Informatización para la creación y uso de una base de datos. Aplicaciones en el mejoramiento.


11. BIOLOGÍA DE SISTEMAS

¿Qué es la biología de sistemas? El enfoque sistémico. Propiedades emergentes de los sistemas. Su integración con las aproximaciones genómicas y moleculares. Informática relacionada: construcción de modelos y constatación práctica de dichos modelos. Aproximaciones bioinformáticas.

BIBLIOGRAFIA III

Textos (la cátedra dispone de copias electrónicas en pdf para la libre consulta por parte de los cursantes)

- *Analysing Gene Expression: A Handbook of Methods: Possibilities and Pitfalls* 2003 S. Lorkowski y P. Cullen, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- *Basic Applied Bioinformatics*. Chandra Sekhar Mukhopadhyay, Ratan Kumar Choudhary, Mir Asif Iquebal. ISBN: 978-1-119-24441-7 September 2017 Wiley-Blackwell 472 Pages
- *Bioinformatics and Functional Genomics*. Jonathan Pevsner. 3rd Edition. ISBN: 978-1-118-58176-6 August 2015 Wiley-Blackwell 1168 Pages
- *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, Third Edition, Andreas D. Baxeavanis, B. F. Francis Ouellette, 2015. ISBN: 0-471-47878-4.
- *Bioinformatique. Génomique et post-génomique* 2006 Frédéric Dardel y François Képès Translated into English by Noah Hardy John Wiley & Sons Ltd. The Atrium, Southern Gate, Chichester, England.
- *Computational and Statistical Approaches To Genomics*, 2 Ed 2006 Wei Zhang e Ilya Shmulevich, Springer Science+Business Media, Inc. New York.
- *Computational Methods for Next Generation Sequencing Data Analysis* Ion Mandoiu, Alexander Zelikovsky. ISBN: 978-1-119-27217-5 September 2016 464 Pages
- *DNA Microarrays* 2007 Ulrike A Nuber, Taylor & Francis Group, LLC, New York.
- *Essentials of Genomics and Bioinformatics*. Christoph W. Sensen ISBN: 978-3-527-61265-9 September 2008 Wiley-Blackwell 442 Pages
- *Functional Genomics: Methods and Protocols* en *Methods in Molecular Biology* vol. 224 M. J. Brownstein y A. Khodursky , Humana Press Inc., Totowa, NJ.
- *Genomics, Proteomics and Vaccines* 2004 Guido Grandi, John Wiley & Sons, Ltd. Hoboken, New Jersey.
- *Handbook of Statistical Genomics*, 4th Edition. David J. Balding, Ida Moltke, John Marioni ISBN: 978-1-119-42914-2 August 2019 1224 Pages
- *Metabolomics* 2007 W. Weckwerth, en *Methods and Protocols Methods in Molecular Biology* vol. 358, Humana Press Inc., Totowa, NJ.
- *Microarray Data Analysis: Methods and Applications* 2007 M. J. Korenberg , en *Methods in Molecular Biology*, vol. 377, Humana Press Inc., Totowa, NJ.

Profesores/as a cargo:	Hopp, Horacio Esteban	
Firmas		Fecha: 28/5/2018
Aclaraciones		

ANEXO I

CONTENIDOS DESGLOSADOS **IV**

a) Clases de Problemas

No contemplados

b) Prácticos de Laboratorio

- Secuenciación por Sanger y NGS con visita al laboratorio UGB. Uso del robot de colonias y ABI, generación de resultados. Uso del secuenciador para análisis de fragmentos. Tipos de datos generados (interpretación, procesamiento x software).
- Uso de NGS para identificación de SNPs y microsatélites.
- Uso de GBS/RAD para análisis de datos (Stack) con o sin un genoma de referencia.
- Genotipificado de un cruzamiento genético con GBS/RAD.
- Uso del MinION® para secuenciación y práctica de ensamblaje de un borrador de genoma bacteriano
- Metagenómica: Diversidad – QIIME2. Evaluación de la diversidad de un ecosistema microbiológico y clasificación taxonómica de sus componentes.
- Ensamblado, binning y anotación de metagenomas.
- Metagenómica aplicada: Bioprospección de enzimas.
- RNAseq: alineación, manipulación y visualización de datos, evaluación de la calidad de los datos, la expresión diferencial y el análisis estadístico, utilizando los programas Bowtie, Tophat y R/Bioconductor
- RNAseq: visualización de datos con la herramienta: ATGC. Manejo bioinformático de los datos: normalización, expresión diferencial, uso de codones.
- Fenómica: taller de análisis de imágenes y cuantificación de datos con Image J. Análisis computacional de datos mediante scripts en R. Morfometría.
- Biología de sistemas: Integración de datos ómicos. Utilización del programa MapMan: integración de datos de transcriptómica y metabolómica.

c) Seminarios

Lectura, exposición y discusión por parte de alumnos y docentes de publicaciones originales recientes de revistas periódicas internacionales. Los trabajos serán seleccionados para ofrecer una visión actualizada de los últimos avances en la temática teórica del curso. Cada participante realiza una completa actualización de un tema específico sintetizada mediante al menos una exposición durante el desarrollo del curso y redactado como informe. Más específicamente, los cursantes se agruparán de a dos o de a tres para realizar una completa actualización de temas claves y específicos (*reviews* y las principales publicaciones que les sirven de base) del estado del arte en temas de genómica aplicada. Las revisiones serán seleccionadas por los docentes responsables y los cursantes podrán elegir, dentro de cierto margen, los temas de su interés para ponerse en contacto con el docente y preparar las exposiciones en forma acorde.

d) Teórico-Práctico o Teórico-Problemas

No contemplados

e) Salidas de campo/viajes^V.

No contemplados.

ANEXO II Adjuntar un ejemplo del cronograma de la Materia, o de los cronogramas en caso de que tenga distintas formas (cuatrimestrales, verano, etc.) ^{VI}

Cronograma

Salvo indicación expresa, las clases son de 8.30 a 17 hs con media hora de almuerzo.

Jueves 1 de febrero: 10 hs Teórica de **secuenciación** + TP de secuenciación y NGS con visita al laboratorio UGB (equipo de Andrea Puebla). El TP se baja de esta página Web. Métodos de secuenciación genómica y bioinformática relacionada. TP de secuenciación (Andrea Puebla-Pablo Vera-Vero Nishi-Natalia Aguirre). Demostrativo: uso del robot y ABI, generación de resultados. Uso del secuenciador para análisis de fragmentos. Tipos de datos generados (interpretación, procesamiento x software).

Viernes 2: Se deja parte el día libre para que los graduados procedentes del interior o exterior puedan completar trámites en Postgrado y pagar aranceles en tesorería (la cual está cerrada en diciembre durante el período de inscripción). Los demás pueden venir para la conexión de portátiles a la máquina virtual (para correr Putty y Xming preinstalados) y otras máquinas virtuales a utilizar durante el curso.

Lunes 5: teórica de introducción a Linux/Shell

(<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbXJz3JuYXNlcWlu dGF0Z2FjGd4OjRmYmNiYTUyMTgzY2NmNjM>) y de R

(<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbXJz3JuYXNlcWlu dGF0Z2FjGd4OjYzNGJjZmViNDZiMzZlY2Q>) y

(<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbXJz3JuYXNlcWlu dGF0Z2FjGd4OjFjZDU1NDVhZWY1ZmFhYWQ>). A cargo de Sergio González y ayuda de Carla Filippi y Natalia Aguirre

Martes 6: Teórica (Marisa Farber) y práctica de secuenciación con los sistemas de tercera generación (tipo Nanopore), demostrativo de uso del MinION para secuenciación y práctica de ensamblaje de un borrador de genoma bacteriano. A cargo de Natalia Pin Viso.

Miércoles 7: Genotipificado por Secuenciación (GBS) (teórica con algún problema biológico como ejemplo y práctica). Uso de NGS para identificación de SNPs. Uso de SNPs (Carla Filippi y Natalia Aguirre). Ver: <https://sites.google.com/site/gbsgenomaplicada2017/>

Jueves 8: Uso de GBS/RAD para análisis de datos (Stack). La cuestión de disponibilidad (o no) de un genoma de referencia. ¿Cómo se trabaja en cada caso? Genotipificado de un cruzamiento genético (Carla Filippi y Natalia Aguirre).

Viernes 9: Continuación y análisis de resultados

Miércoles 14: Metagenómica: Teóricas de análisis de biodiversidad basados en genes ribosomales y funcionales (Eva Figuerola y Paola Talia). Práctica: análisis bioinformático de secuenciación a gran escala de amplicones, interpretación de un problema biológico con el programa USearch U-Parc (Eva Figuerola - Natalia Pin Viso)

Jueves 15: Teórica de Ensamblado de Genomas y Práctica (Eva Figuerola)

Viernes 16: Teórica y Práctica de Bioprospección de enzimas (Eva Figuerola)

Lunes 19: Teóricas de Metabolómica (Fernando Carrari), Fenómica (Esteban Hopp) y **Biología de Sistemas** (Luis de Haro).

Martes 20: Metabolómica y Biología de Sistemas. TP de Integración de datos ómicos (Luis de Haro y Soledad Lucero) Utilización del programa MapMan: integración de datos de transcriptómica y metabolómica

Miércoles 21: Teórica de transcriptómica (incluye RNASeq) en la FIL (Pablo Cerdán) Hasta las 14 hs.

Jueves 22: Instalación de programas en las portátiles. Taller de **RNAseq**: alineación, manipulación y visualización de datos, evaluación de la calidad de los datos, la expresión diferencial y el análisis estadístico, utilizando los programas Bowtie, Tophat y R/Bioconductor (Sergio González). Se seguirán los siguientes lineamientos: <https://sites.google.com/site/rnaseqgenomaplicada2017/>

Viernes 23: ¿Cómo se trabaja con datos crudos, mapeo de lecturas ("reads") usando un genoma de referencia? Manejo de los datos alineados: Uso de archivos SAM/BAM. (Máximo Rivarola y Sergio González). Manejo de los datos alineados: Normalización de datos, expresión diferencial, uso de codones. Almacenamiento y visualización de datos con la herramienta: ATGC (Máximo Rivarola y Sergio González).

26, 27 y 28: Seminarios

Sábado 3 de marzo: Examen Final (hasta las 13 hs)

Notas:

^I El contenido de este documento se ratificará o rectificará bianualmente

^{II} Objetivos: redactados en función de los aprendizajes buscados (no en función de lo que los docentes hacen para alcanzar esa meta). Por ejemplo, la redacción de cada objetivo debería comenzar con alguna frase como “Que los/as estudiantes sean capaces de... conozcan... comprendan..., etc.”.

Por favor evitar frases *imprecisas* (ej.; “Se hará énfasis en las distintas estrategias y en las distintas metodologías de estudio”) o *incorrectas* (ej.; “El docente fomentará...”)

Si un el objetivo es que el/la estudiante priorice el espíritu crítico sobre dogmas, entonces, debería estar redactado de ese modo, en términos de lo que debe lograr el/la estudiante. Si se incluyen estos objetivos cognitivos de largo plazo como el anterior deben ser coherentes con las actividades y evaluaciones que permitan alcanzar los mismos. Para la elaboración y/o redacción de los objetivos puede consultar al CEFIEC a través de los emails: emeinardi@gmail.com o leomgalli@gmail.com

^{III} Bibliografía obligatoria. De manera optativa bibliografía sugerida para ampliar temas.

^{IV} De acuerdo a lo indicado en los ítems de “Actividad”: Títulos y muy breve descripción del tema a desarrollar, de 160 caracteres como máximo.

^V Máximo: 320 caracteres.

^{VI} Los cronogramas pueden ser enviado en cualquier formato.