



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

Asignatura: Introducción a la Bioinformática Molecular

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

Duración de la asignatura (en semanas)	16
Cuatrimestre(s) en que dicta (indicar cuatrimestre o verano):	1
Frecuencia en que se dicta (cuatrimestral, anual, bianual, etc.)	anual

ACTIVIDAD	Horas semanales	Número de semanas	Horas totales
Teóricas	4	16	64
Laboratorios	4	16	64
Problemas	2	16	32
Seminarios	-	-	-
Teórico- prácticos o Teórico- problemas	-	-	-
Si corresponde, especifique las horas de otras actividades (salidas de campo, etc.)			
Carga horaria semanal máxima	10		
Carga horaria semanal mínima	6		
Carga horaria total:	160		

Asignaturas correlativas:	QUIMICA BIOLOGICA GENETICA I
Forma de Evaluación:	2 Parciales Teóricos Virtuales Escritos (Con opción de 2 recuperatorios) Entrega digital de Trabajos Prácticos Promoción (nota promedio superior a 7) o Final (Virtual y/o Presencial).



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



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

OBJETIVOS I

Dar los conocimientos esenciales de los algoritmos y herramientas de la bioinformática Molecular para que el biólogo pueda desarrollar las tareas básicas de acceso y manejo de base de datos bioinformáticas, la comparación y estudio funcional de genes y proteínas, y el ensamblado y análisis de genomas.

El objetivo de las teóricas es que el alumnado comprenda los fundamentos básicos de la bioinformática moderna y los pueda aplicar en su trabajo futuro.

El objetivo de las clases de laboratorio es el manejo de herramientas bioinformáticas y poder sacar conclusiones de su uso.

CONTENIDOS MÍNIMOS (ya aprobados Anexo IV Plan 2019)

Introducción a la Bioinformática Molecular

Bases de datos Primarias: Definición de bases de datos primarias. Visión histórica de la creación de las mismas. Funcionamiento de las Bases de datos: índices, campos, métodos de búsqueda. Bases de datos de proteínas. Bases de datos de ADN. Ejemplos de bases de datos primarias: Genbank, EMBL, Swiss-Prot, TrEMBL, PDB. Análisis de secuencias: Introducción de probabilidad y estadística. Alineamiento global por pares. Alineamiento Múltiple. Generación de Matrices de score (BLOSUM, PAM). Dot-Plot. Programación dinámica.

Programas de alineamiento: BLAST. FASTA. Búsquedas en bases de datos por similitud de secuencia. Patrones de secuencias y perfiles. Filogenia molecular. PSI-BLAST, PHI-BLAST, Mega-Blast. Bases de datos Secundarias: Definición de bases de datos secundarias. Construcción de bases de secundarias. El problema de los falsos positivos/negativos. Modelos ocultos de Markov. Ejemplos de bases de Datos secundarias: Pfam, Gene-Ontology, UniProt, PRINTS, ProSite. Algoritmos, complejidad y heurísticas. Diseño y mantenimiento de bases de Datos secundarias.

Análisis Bioinformático de Genomas: Ensamblado y anotación de genomas, predicción de genes, Bidireccional best Hits y Iterative predictive Blast. Base de datos de Genomas. Mapeo físico de genes. Uso de Genome Browsers (NCBI), Ensembl y Galaxy. Comparación de Genomas.

Análisis Bioinformático de datos high-throughput de microarreglos (MicroArrays) Introducción a los MicroArrays, Análisis estadístico de significancia de los datos, Análisis de expresión por MicroArrays, definición de estado metabólico (expresoma, proteoma y metaboloma), MicroArrays específicos sobre splicing alternativo (exon arrays, splicing sensitive arrays), MicroArrays de Glicómica. Metagenómica y Metabolómica: Introducción a la metagenómica, secuenciación de próxima generación. Anotación y análisis de metagenomas. Introducción a la metabolómica, análisis de vías y estados metabólicos. Uso de Base de datos KEGG.

Bioinformática Estructural: Repaso de estructura de proteínas. Predicción de estructura secundaria, DSSP. Análisis bioinformático de estructuras, alineamiento estructural. Predicción de estructura terciaria. Threading. Modelado comparativo. Métodos ab-initio. Búsqueda de motivos estructurales. Interacciones entre biomoléculas: Bases moleculares de reconocimiento específico. Interacción proteína-proteína.



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

Interacción proteína-ADN. Interacción droga-proteína. Predicción de interacciones. Predicción de estructuras. Monte Carlo. Algoritmos genéticos. Docking. Clustering. Métodos evolutivos. Bases de datos de interacciones. BIND. DIP.CAPRI.

PROGRAMA ANALÍTICO

Bases de datos Primarias

El estudiante debe adquirir conocimientos básicos del manejo de bases de datos bioinformáticas y los conocimientos básicos informáticos de cómo funcionan las bases de datos relacionales y no relacionales. En particular con respecto a las bases de datos bioinformáticas se aprenderá el concepto de bases de datos primarias, secundarias, anotadas, curadas y redundancia. Con respecto a los conceptos informáticos se deberá comprender no como programar sino conceptos de: índices, campos, métodos de búsqueda y algoritmos de llenado. Ejemplos de bases de datos primarias: TrEMBL, Genbank, Swiss-Prot,

Análisis de secuencias

Aprender a Alinear secuencias y los algoritmos básicos. Introducción a los algoritmos de Smith Waterman y Needleman-Wunsch. Se realizará una descripción detallada del algoritmo de programación dinámica y los conceptos claves de comparación de secuencias. El concepto evolutivo de las matrices de scoring tanto Blosum como PAM. Diferenciar alineamientos locales y globales.

Algoritmos de alineamiento para bases de datos. Aproximaciones heurísticas y como ejemplo central se verá BLAST introduciendo los conceptos de WORD, E-VALUE, scoring pairs. Los programas que se verán en la parte práctica incluyen BLAST, FASTA, PSI-BLAST, PHI-BLAST, Mega-Blast.

El modulo de comparación de secuencias incluye los conceptos básicos de armado de patrones de secuencias y perfiles. Una introducción a la complejidad de los algoritmos y al concepto de heurísticas contra solución óptima.

Bases de datos Secundarias

El concepto de Modelos Oculto de Markov y el armado de algoritmos para comparar secuencias a partir de alineamientos múltiples. Armado de PFAM y división de proteínas en familias. El concepto de falsos y verdaderos positivos y negativos. Para profundizar este contenido se verán ejemplos de bases de Datos secundarias como, Pfam, Gene-Ontology y UniProt,

Análisis Bioinformático de Genomas

Contenido mínimo de análisis bioinformáticos de genomas completos. Conceptos básicos de técnicas de secuenciación de próxima generación, generación de lecturas y algoritmos de ensamblado. Algoritmos ocultos de Markov para identificación de genes en genomas. Concepto de Pipeline bioinformático aplicado a la anotación de genomas. Algoritmo de bidireccional best Hits y Iterative predictive Blast como concepto de anotación de genes por comparación en bases de datos. Aprender sobre las bases de datos de Genomas más usadas y comprender el concepto de mapeo físico y el uso de analizadores de genomas como Genome Browser.

Análisis Bioinformático de datos high-throughput

Se pretende tener una comprensión básica del análisis bioinformático de datos high-throughput en



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

biología molecular y las herramientas bioinformáticas para su análisis. Se espera que el alumno adquiera capacidad de análisis estadístico de datos high-throughput como ser MicroArrays, Rna-seq.

Metagenómica y Metabolómica

Nuevas herramientas en genómica y el análisis de los datos. Una breve introducción a la metagenómica, y los procesos de anotación y análisis de los metagenomas como extensión del caso de un genoma. La aplicación de la metagenómica a la metabolómica y el consecuente análisis de vías y estados metabólicos.

Bioinformática Estructural

El alumno deberá adquirir conceptos básicos de la estructura de proteínas incluyendo los 4 niveles de organización y el manejo de programas de visualización. Identificación de interacciones y algoritmos para clasificarlas y determinarlas. En particular a partir de la secuencia se podrá realizar la predicción de estructura secundaria, terciaria y cuaternaria. En este módulo los alumnos aprenderán a comparar estructuras en la computadora y el concepto básico de los algoritmos de alineamiento y medición de distancias. Con respecto a la predicción de estructura de proteínas se aprenderá en particular Modelado Comparativo y conceptos básicos de Threading. Introducción a los potenciales de interacción atómicos y de grano grueso y aplicarlos a los métodos de predicción estructural ab-initio o de primeros principios.

Interacciones entre biomoléculas

Conceptos básicos de las interacciones de macromoléculas con pequeñas moléculas y los algoritmos de predicción. Algoritmos de muestreo de Montecarlo y algoritmos genéticos aplicados a la generación de conformaciones de moléculas. Aplicación de algoritmos de scoring basados en potenciales a métodos de búsqueda de drogas.

BIBLIOGRAFIA II

- a) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Andreas D. Baxevanis, B. F. Francis Ouellette.
- b) Bioinformatics: Sequence and Genome Analysis. David W. Mount.
- c) Structural Bioinformatics (Methods of Biochemical Analysis, V. 44) Philip E. Bourne, Helge Weissig
- d) Developing Bioinformatics Computer Skills. Cynthia Gibas, Per Jambeck.
- e) Statistical Methods in Bioinformatics. Warren J. Ewens, Gregory R. Grant
- f) Protein Structure Prediction - A Practical Approach, M. J. E. Sternberg editor, Oxford University Press, 1996.
- g) Introduction to Protein Structure, C. Branden and J. Tooze Garland Publishing, Inc. New York and London, 1999.

Profesores/as a cargo: Marcelo A. Marti



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

Firmas
Aclaraciones

y

Fecha: 16 Junio 2020

Marcelo A. Marti



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

ANEXO I

CONTENIDOS DESGLOSADOS III

a) Clases de Problemas

Acompañan las clases teóricas con la resolución de casos modelo que explican los algoritmos discutidos en la teórica.

b) Prácticos de Laboratorio

Bases de datos Primarias : Bases de datos Primarias: Sitios Populares de Bioninformática, Búsquedas en bases de datos: el NCBI y entrez, Análisis de registros. Gene-Ontology.

Alineamiento I: Alineamiento de Secuencias: Dot-plot. Alineamiento, Búsqueda en base de datos por BLAST. Alineamiento Múltiple. Construcción de Árboles Filogenéticos (Caso real Kinasas). Búsqueda y uso de bases de datos secundarias: PFAM.

Alineamiento II: Programación de un algoritmo de alineamiento, análisis y generación de las matrices de scoring.

Bases de datos Secundarias: Búsquedas y navegación en PFAM, Prosite, construcción de patrones, interrelación entre las bases.

Predicción de genes: Algoritmos de predicción de genes, Glimmer, Blast-Extended Repraze. Determinación de la calidad y confianza de la predicción.

Análisis y anotación de genomas: Estudio de los mapas físicos de genomas (Genome Browser del NCBI), ensamblado y anotación de genomas.

Análisis de resultados de NGS: Análisis de resultados de microarreglos de expresión, significancia de los datos, asociación a estados metabólicos según KEGG

Estructura de proteínas: Manejo de programas de visualización. Visualización de estructuras de proteínas. Análisis de motivos. Análisis de interacciones entre biomoléculas. Uso del PDB, Entrez y PDBsum.

Predicción de estructura de proteínas: Modelado Comparativo. Uso de Modeller. Ejemplos puntuales. Análisis por What-Check.



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

Interacción proteína-ligando: Docking proteína ligando: (Caso real inhibición Kinasas).
 Comparación de algoritmos.

b) Seminarios

no contemplado en la presente propuesta.

c) Teórico-Práctico o Teórico-Problemas

no contemplado en la presente propuesta

d) Salidas de campo/viajes^{IV}.

No corresponde

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

Inicio 1er cuatrimestre (Lunes 19/3)	
miércoles 21/3	Introducción Bases de Datos primarias
viernes 23/3	NCBI ENtrez
miércoles 28/3	Metodología en Bases de datos
viernes 30/3	Introducción a las herramientas bioinformáticas
miércoles 4/4	Métodos de Alineamiento. Dot Plot. Programación dinámica.
viernes 6/4	Bases de datos
miércoles 11/4	Métodos de Alineamiento. Blast.Alineamientos Múltiples Búsqueda de motivos en secuencia. (Adrián)
viernes 13/4	Blast/dOT PLOT



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

miércoles 18/4	Bases de Datos Secundarias
viernes 20/4	Clustal PROSITE
miércoles 25/4	Secuenciación con técnicas de nueva generación (Marcelo Martí)
viernes 27/4	PFAM
miércoles 2/5	Repaso
viernes 4/5	1er Parcial
miércoles 9/5	Ensamblado, Mapeo y Alin Genomas (Marcelo Marti)
viernes 11/5	Análisis de Genomas. Anotacion (RAST)
miércoles 16/5	Anotación de genomas (Adri)
viernes 18/5	Análisis de Genomas. Anotacion (XomeQ)
miércoles 23/5	Genomas Humano.(Marcelo Marti)
viernes 25/5	Feriado revolcuion de mayo
miércoles 30/5	Estructura de proteínas y Bioinformática de las interacciones (Enfoques Estadísticos y Biofísicos) (Adrian)
viernes 1/6	VMD Comparación Estructural
miércoles 6/6	Modelado Comparativo o por Homología (Adrián) - Threading, Prediccion de estructura secundaria y Métodos Ab-initio (Adrián)
viernes 8/6	Modelado por Homologia -Usos de servidores online
miércoles 13/6	Análisis de redes metabólicas y biología de sistemas (Dario Doporto)
viernes 15/6	Práctica de vias metabólicas
miércoles 20/6	Docking droga proteína y proteína-proteína (Adrián)
viernes 22/6	Trabajo especial



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

miércoles 27/6	Repaso
viernes 29/6	SEGUNDO PARCIAL
miércoles 4/7	REPASO
viernes 6/7	Recuperatorio
Fin de 1er cuatrimestre (Sabado 7/7)	

Notas:

^I 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

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

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

^{IV} Máximo: 320 caracteres.

^V Los cronogramas pueden ser enviado en cualquier formato.