



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

PROGRAMA ANALÍTICO DE ASIGNATURA

DEPARTAMENTO: Ingeniería en Sistemas de Información

CARRERA: Ingeniería en Sistemas de Información

NOMBRE DE LA ACTIVIDAD CURRICULAR: Bioinformática

Año Académico: 2023

Área: Sistemas Inteligentes

Bloque: Electivas

Nivel: 5º

Tipo: Electiva

Modalidad: Cuatrimestral

Cargas horarias totales:

<i>Horas reloj</i>	<i>Horas cátedra</i>	<i>Horas cátedra semanales</i>
72	96	6

COMPOSICIÓN DEL EQUIPO DE CÁTEDRA:

Profesor Adjunto: Lic. Mg. Doc. Patricio Yankilevich

FUNDAMENTACIÓN

La expansión biotecnológica y de las ciencias biomédicas ocurrida hacia fines del siglo pasado (y comienzos de este) ha generado una nueva revolución en el escenario de la investigación biológica. Las tecnologías y métodos experimentales utilizados actualmente en las áreas de genómica, proteómica y metabolómica, han dado lugar a nuevos procesos de diagnóstico e investigación molecular que están cambiando el escenario actual de la investigación científica (sólo comparable a la transformación provocada por el proyecto genoma humano). Los secuenciadores de nueva generación, los *microarrays* de ADN o los espectrómetros de masas de alta resolución permiten medir la actividad celular a nivel molecular y a escala de genomas y proteomas completos. La inmensa cantidad de información obtenida con estas tecnologías sobre los cambios moleculares asociados a distintas condiciones de un sistema biológico, hace posible encarar algunas preguntas fundamentales sobre los procesos fisiológicos y patológicos de la célula, o conocer cómo se construye un organismo. A medida que estas nuevas tecnologías van generando volúmenes de datos cada vez mayores, paralelamente también va creciendo la complejidad de las tareas de procesamiento y análisis de datos. En la actualidad,



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

esto supone un cuello de botella en las tareas de investigación bioinformática, ya que el ritmo de producción de datos es mucho mayor que los desarrollos de funcionalidad de las técnicas de análisis existentes. Consecuentemente, se ha creado la necesidad inmediata de desarrollos computacionales orientados a la creación de herramientas eficientes especializadas en el análisis y la comprensión de datos biológicos provenientes de estas nuevas tecnologías.

La bioinformática está teniendo un profundo impacto en la biología molecular, la biomedicina, la neurociencia y otras áreas relacionadas con la biotecnología. El perfil del bioinformático es hoy altamente demandado por todos los grupos de investigación en estas áreas, hecho que se ve reflejado en la incorporación de maestrías y/o carreras de grado en bioinformática, y en la designación de la bioinformática como materia obligatoria en todas las carreras de biotecnología. Los grandes centros de investigación y empresas farmacéuticas cuentan hoy con grupos de investigación bioinformática, los cuales llevan estrechas colaboraciones con otros de los grupos de investigación, en la comprensión de enfermedades, el desarrollo de nuevos y mejores métodos de diagnóstico y tratamiento de enfermedades, en el descubrimiento y desarrollo de nuevos fármacos, y en la implementación de estrategias de medicina personalizada.

Si bien la bioinformática es una ciencia interdisciplinaria que puede ser abordada por biólogos con interés computacional, el grado de complejidad de análisis de datos biológicos y el volumen de información a procesar requieren también de computadores científicos con comprensión de la biología para poder ser resueltos eficientemente.

OBJETIVOS

- Reconocer los conceptos teóricos y prácticos de los métodos y herramientas de la bioinformática utilizados en Genómica y Proteómica.
- Identificar ventajas y limitaciones de los principales algoritmos bioinformáticos para el análisis de secuencias génicas, análisis de expresión génica, análisis filogenético, análisis de genomas, identificación de proteínas, análisis estructural y funcional de proteínas.
- Reconocer los distintos métodos y herramientas disponibles para el análisis de secuencias génicas, análisis de expresión génica, análisis filogenético, análisis de genomas, identificación de proteínas, análisis estructural y funcional de proteínas.
- Distinguir lenguajes de programación bioinformática para utilizar en problemas biológicos concretos.
- Reconocer las principales bases de datos biológicos y herramientas computacionales de la bioinformática para evaluar y utilizar nuevas técnicas que puedan surgir en el futuro.

CONTENIDOS



*Universidad Tecnológica Nacional
Facultad Regional Buenos Aires*

Contenidos analíticos

Unidad Temática 1: Introducción a la Biología Celular y Molecular.

Se estudia el origen de la vida. Teoría de la evolución. Trabajos realizados por Gregor Mendel. Origen de la genética.

Unidad Temática 2: Dogma Central de la Biología Molecular. Métodos y Modelos Experimentales. Paradigma de Darwin.

Dogma central de la Biología. Componentes moleculares de la vida. Estructura y Función. Célula en Eucariotas, Procariotas, Organismos multicelulares. Biología Molecular del Gen. Expresión y Regulación Génica. Instrumentación, Modelos y Métodos experimentales.

Unidad Temática 3: Definición de Bioinformática. Introducción al Análisis de Secuencias. Formatos de datos de secuencias. Bases de datos de secuencias.

Se estudian los distintos formatos de secuencias de ADN, ARN y PROTEÍNAS. Se analizan los formatos en que se anotan los Genomas. Se introduce el concepto de búsqueda de secuencias en bases de datos públicas.

Unidad Temática 4: Alineamiento de Secuencias. Matrices de Sustitución. Algoritmo de Needleman-Wunsch.

Se estudian los principales algoritmos de alineamiento de pares de secuencias. Alineamiento global y local. Algoritmo de Needleman-Wunsch (global).

Unidad Temática 5: Algoritmo de Smith-Waterman. Algoritmo de Ends-Free. Métodos heurísticos de alineamiento. Búsqueda de secuencias en bases de datos.

Se estudia el algoritmo de Smith-Waterman (local). Se analizan otros algoritmos de alineamiento de secuencias. Se explica el funcionamiento de BLAST.

Unidad Temática 6: ENCODE. BLAST. Bases de datos de secuencias de ADN y ARN.

Se estudia el proyecto ENCODE. Se ven variante del algoritmo BLAST. Se estudia el algoritmo BLAT. Se estudian las matrices de puntuación y penalidades Blosum y PAM. Bases de datos ARNcentral, Genbank. Predicción de estructura secundaria de ARN.

Unidad Temática 7: Bases de datos de Secuencias de Proteínas. Bases de datos secundarias.

Se estudian las bases de datos de proteínas Uniprot, PDB, entre otras. Predicción de estructura de proteínas y su clasificación.

Unidad Temática 8: Alineamiento Múltiple de Secuencias. Alineamiento Múltiple de Secuencias de ADN. Alineamiento Múltiple de Secuencias de Proteínas.

Se estudia el alineamiento de secuencias múltiples (MSA). Se estudia el funcionamiento del algoritmo Clustal.



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

Unidad Temática 9: Linux. Perl, BioPerl. local BLAST.

Se estudia como armar una workstation para bioinformática. Se ven distintos lenguajes de programación para analizar secuencias como Perl, BioPerl y otros.

Unidad Temática 10: Perfiles, Dominios, HMMs.

Definición y explicación de dominios funcionales, de motivos y Hidden Markov Models para análisis de secuencias.

Unidad Temática 11: Predicción de Genes.

Algoritmos de predicción de genes en genomas.

Unidad Temática 12: Introducción al análisis filogenético.

Estudio de distintas técnicas de predicción de árboles filogenéticos de secuencias.

Unidad Temática 13: Introducción al Modelado de Estructura de Proteínas.

Modelado ab initio, modelado por homología. Google Alphafold.

Unidad Temática 14: Introducción al Análisis de Genomas. Ensamblado, análisis, visualización y anotación de Genomas. Genome browsers.

BIBLIOGRAFÍA OBLIGATORIA

- BUFFALO, V. (2015) Bioinformatics Data Skills: Reproducible and Robust Research with Open Source Tools. Ed. O'Reilly & Associates.
- DURBIN, R., S. R. EDDY, A. KROGH Y G. MITCHISON. (1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Ed. Cambridge University Press.
- GIBAS, C. Y P. JAMBECK. (2001) Developing Bioinformatics Computer Skills. Ed. O'Reilly & Associates.
- MOUNT, D. W. (2004) Bioinformatics: Sequence and Genome Analysis. Ed. Cold Spring Harbor Laboratory.
- TISDALL, J. (2001) Beginning Perl for Bioinformatics. Ed. O'Reilly & Associates.

PÁGINAS WEB DE INTERÉS

- European Bioinformatics Institute - <https://www.ebi.ac.uk>
- National Center for Biotechnology Information - <https://www.ncbi.nlm.nih.gov>

CORRELATIVAS

Para cursar y rendir

- Cursadas:
 - Redes de Datos



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

- Administración de Sistemas de Información
 - Simulación
 - Ingeniería y Calidad de Software
- Aprobadas:
 - Diseño de Sistemas de Información
 - Sistemas Operativos
 - Bases de Datos