



Ref. Expte. N° 7679/2019

Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Ciudad Autónoma de Buenos Aires, 25 NOV 2019

VISTO

La nota a fojas 1 presentada por el Secretario Académico del Departamento de Química Biológica, mediante la cual eleva la información del curso de posgrado **Bioinformática Avanzada** para el año 2020,

CONSIDERANDO

- Lo actuado por la Comisión de Doctorado,
- Lo actuado por la Comisión de Posgrado,
- Lo actuado por la Comisión de Presupuesto y Administración,
- Lo actuado por este Cuerpo en la sesión realizada en el día de la fecha,
- En uso de las atribuciones que le confiere el Artículo 113° del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD
DE CIENCIAS EXACTAS Y NATURALES
RESUELVE:**

- ARTÍCULO 1°:** Aprobar el nuevo curso de posgrado **Bioinformática Avanzada** de 50 horas de duración, que será dictado por el Dr. Marcelo Martí.
- ARTÍCULO 2°:** Aprobar el programa del curso de posgrado **Bioinformática Avanzada** obrante a fs. 11/12 para su dictado entre marzo y abril de 2020.
- ARTÍCULO 3°:** Aprobar un puntaje máximo de tres (3) puntos para la Carrera del Doctorado.
- ARTÍCULO 4°:** Aprobar un arancel de \$500 (pesos quinientos), estableciendo que dicho arancel estará sujeto a los descuentos y exenciones estipulados mediante la Resolución CD N° 1072/19. Disponer que los fondos recaudados ingresen en la cuenta presupuestaria habilitada para tal fin, y sean utilizados de acuerdo a la Resolución 072/03.
- ARTÍCULO 5°:** Disponer que de no mediar modificaciones en el programa, la carga horaria y el arancel, el presente Curso de Posgrado tendrá una vigencia de cinco (5) años a partir de la fecha de la presente Resolución.
- ARTÍCULO 6°:** Comuníquese a todos los Departamentos Docentes, a la Dirección de Estudiantes y Graduados, a la Dirección de Movimiento de Fondos, a la Dirección de Presupuesto y Contabilidad, a la Biblioteca de la FCEyN y a la Secretaría de Posgrado con copia del programa incluida. Cumplido, archívese.

RESOLUCIÓN CD N° 2877
SP-GA- 05/11-2019

2877

Dr. PABLO J. GROISMAN
Secretario Adjunto de Posgrado
FCEyN - UBA

Dr. JUAN CARLOS RESOREDA
DECANO

Afluencia a quien está dirigido el curso:
Estudiantes de Doctorado del Area Bio y/o Informática

Necesidades materiales del curso:
No posee

1-b-

Programa analítico del curso con Bibliografía (puede adjuntarse en hojas separadas):
Contenidos: Python y Biopython. Estructura de un programa de python. Manejo de secuencias. Alineamiento de a pares y múltiple de secuencias (ClustalW, Muscle, BLAST). Manejo y acceso a las base de datos bioinformáticas. Manejo y acceso de estructuras de biomoléculas. Lectura/escritura de archivos PDB. Representación de la estructura en Biopython. Modelos Ocultos de Markov. Utilización de HMMER. Filogenia con Bio.Phylo. Análisis de grafos y clusterización. Manejo de datos de NGS. Módulos de GATK.

Unidad 1) Python y Biopython.
Estructura de un programa de python, variables, listas, loops, condicionales, operaciones lógico/algebraicas, entrada/salida de datos módulos, funciones. Bibliotecas (numpy, pandas, etc) Entorno de programación Jupyter notebook. Que es Biopython? Instalación de Biopython. Quick Start.

Unidad 2) Manejo de secuencias.
Secuencias y alfabetos, transcripción, traducción, Creación de registros de tipo Seq (formato FASTA, genebank). Alineamiento de a pares y múltiple de secuencias (ClustalW, Muscle, BLAST).

Unidad 3) Manejo y acceso a las base de datos bioinformáticas
Acceso y bajada de NCBI. Swiss-Prot y Expasy, Prosite. PubMed

Unidad 4) Manejo y acceso de estructuras.
Modulo PDB. Lectura/escritura de archivos PDB. Representación de la estructura en Biopython. Navegando una estructura. Análisis estructurales (distancias, angulos, dihedros) Superposición estructural. Acceso al PDB.

Unidad 5) Modelos Ocultos de Markov.
Utilización de HMMER. Problema de entrenamiento (hmmbuild, hmmlcalibrate, hmmit). Problema de alineamiento (hmmalign). Búsqueda en bases de datos (hmmsearch, hmfetch, hmmpfam).

Unidad 6) Filogenia con Bio.Phylo
Construcción de árboles filogenéticos. Colorear ramas de los árboles. Búsqueda y modificación de árboles

Unidad 7) Análisis de grafos y clusterización.
Construcción de grafos. Bases de datos mapeadas sobre grafos (KEGG, String). Análisis de grafos (centralidad, choke point). Funciones de distancia. Clusterización Jerárquica. Análisis de componentes principales.

Unidad 8) Manejo de datos de NGS.

Módulos de GATK. Mapeo, Alineamiento y llamado de Variantes. VQSR. Anotación de VCF.

Bibliografía:

- 1) Biopython Tutorial and Cookbook. Jeff Chang, Brad Chapman, Iddo Friedberg, Thomas Hamelryck, Michiel de Hoon, Peter Cock, Tiago Antao, Eric Talevich, Bartek Wilczyński **2019** (Biopython 1.74)
- 2) Applied Computational Genomics, Springer Nature Singapore Pte Ltd. **2018** DOI <https://doi.org/10.1007/978-981-13-1071-3>
- 3) Bioinformatics Algorithms: An Active Learning Approach. Philp Compeau. Editor: Active Learning Publishers (**2014**)
- 4) STRUCTURAL BIOINFORMATICS : AN ALGORITHMIC APPROACH CRC press 2017.

Bibliografía complementaria.

- 1) Translational Bioinformatics, PLOS Computational Biology, December 2012 | Volume 8 | Issue 12 | e1002796
- 2) Bioinformatics: Sequence and Genome Analysis, by David Mount (Author), Cold Spring Harbor Laboratory Press. 2013
- 3) Essentials of Genomic and Personalized Medicine, by Geoffrey S. Ginsburg (Editor), Huntington F Willard PhD (Editor). Academic Press; 1 edition (2009)
- 4) Jason R. Miller, Assembly algorithms for next-generation sequencing data, Genomics 95 (2010) 315–327

1-c-

Actividades prácticas propuestas (puede adjuntarse en hojas separadas):

Trabajo Práctico 1) Manejo De secuencias

Los alumnos deberán programar alguno(s) de los siguientes ejercicios:

- i) Programar buscador de ORFs
- ii) Programar todas las mutaciones generadas por el cambio de 1 base de cada codon (programar todas las mutaciones de 1 aminoácido posibles)
- iii) Hacer estadística de probabilidad de mutaciones en base a codones
- iv) Generar secuencia de ADN/Proteína al azar

Trabajo Práctico 2) Manejo y acceso de estructuras.

Los alumnos deberán programar alguno(s) de los siguientes ejercicios:

Ejercicios Cargar PDB calcular

- i) calcular radio de giro
- ii) Mapa de contactos
- iii) Residuos de superficie vs interior
- iv) Estructura secundaria

Trabajo Práctico 3) Manejo y acceso de estructuras.

Los alumnos deberán programar alguno(s) de los siguientes ejercicios:

- i) Mapear features de Uniprot sobre el PDB
- ii) Buscador de variantes de uniprot en ExAC

Trabajo Práctico 4) Minería de datos y HMMs

Los alumnos deberán programar alguno(s) de los siguientes ejercicios:

- i) Hacer minería de secuencias con un HMM calculado propósito específico
- ii) Dada una secuencia determinar dominios, analizar features asociadas.

Trabajo Práctico 5) Variantes Anotacion

Los alumnos deberán programar alguno(s) de los siguientes ejercicios: