

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
COORDINACIÓN GENERAL DE FORMACIÓN BÁSICA
COORDINACIÓN GENERAL DE FORMACIÓN PROFESIONAL Y VINCULACIÓN UNIVERSITARIA
PROGRAMA DE UNIDAD DE APRENDIZAJE

I. DATOS DE IDENTIFICACIÓN

1. **Unidad Académica:** Facultad de Ingeniería, Mexicali; Facultad de Ingeniería, Arquitectura y Diseño, Ensenada; y Escuela de Ciencias de la Ingeniería y Tecnología, Valle de las Palmas.
2. **Programa Educativo:** Bioingeniero
3. **Plan de Estudios:** 2020-1
4. **Nombre de la Unidad de Aprendizaje:** Bioinformática
5. **Clave:** 36273
6. **HC:** 02 **HL:** 00 **HT:** 02 **HPC:** 00 **HCL:** 00 **HE:** 02 **CR:** 06
7. **Etapas de Formación a la que Pertenece:** Terminal
8. **Carácter de la Unidad de Aprendizaje:** Optativa
9. **Requisitos para Cursar la Unidad de Aprendizaje:** Ninguno



Equipo de diseño de PUA

Dante Alberto Magdaleno Moncayo
Rafael Villa Angulo
José Luis Becerra Buenrostro
Claudia Mariana Gómez Gutiérrez

Firma

José Luis

**Vo.Bo. de Subdirectores de
Unidades Académicas**

Alejandro Mungaray Moctezuma
Humberto Cervantes de Ávila
María Cristina Castañón Bautista

Firma

M. Cristina Castañón B.

Fecha: 30 de octubre de 2018

II. PROPÓSITO DE LA UNIDAD DE APRENDIZAJE

En este curso el alumno desarrollará la habilidad para buscar y analizar secuencias biológicas de ADN, ARN y Proteínas, a través de herramientas bioinformáticas para la búsqueda en bases de datos públicas, identificación y visualización de material genético asociado con fenotipos específicos, en un marco ético científico que resguarde la integridad de la información analizada.

La importancia de este curso está centrada en la capacidad que desarrollará el estudiante para aportar soluciones en problemas prácticos de asociación de genes con enfermedades, de análisis filogenéticos, genómica de poblaciones y metagenómica a través del desarrollo de algoritmos e implementación de herramientas bioinformáticas de libre acceso.

Esta asignatura pertenece a la etapa terminal con carácter de optativa, y corresponde al área de conocimiento de Ingeniería Aplicada y Diseño.

III. COMPETENCIA DE LA UNIDAD DE APRENDIZAJE

Aplicar métodos de análisis bioinformático a secuencias de ácidos nucleicos y proteínas, utilizando herramientas bioinformáticas, para identificar genes relacionados con fenotipos, caracterizar y clasificar secuencias biológicas, análisis de genomas y metagenomas, con responsabilidad, pensamiento crítico y empatía.

IV. EVIDENCIA(S) DE DESEMPEÑO

Elabora y entrega un reporte con el análisis comparativo de secuencias de estructuras de ácidos nucleicos y proteínas, el cual debe integrar: análisis e interpretación de resultados. Cada ejercicio debe de contener: portada, marco teórico, planteamiento del problema, la metodología realizada, cálculos, resultados, conclusiones y bibliografía.

V. DESARROLLO POR UNIDADES

UNIDAD I. La bioinformática

Competencia:

Identificar la aplicación de la bioinformática, a través del estudio de sus orígenes y evolución, para reconocer su importancia en la industria, con actitud analítica y reflexiva.

Contenido:**Duración:** 6 horas

- 1.1 La Bioinformática
- 1.1 El campo de aplicación de la Bioinformática
- 1.2 La biología molecular y la bioinformática
- 1.3 Dogma central de la biología molecular
- 1.4 Estructura molecular de los genes y los cromosomas
- 1.2 Bases de datos públicas de información biológica
 - 1.2.1 Bases de datos primarias y secundarias.
 - 1.2.2 Base de datos NCBI (Centro Nacional de Información de Biotecnología, USA).
 - 1.2.3 Base de datos Genbank.
 - 1.2.4 Acceso a la Base de datos Protein Data Bank.
 - 1.2.5 Plataformas de secuenciación de alto rendimiento
- 1.3 Revisión general del problema de análisis de secuencias biológicas
 - 1.3.1 Alineación de cadenas por pares de nucleótidos
 - 1.3.2 Programación dinámica
 - 1.3.3 Alineación de secuencias múltiples
 - 1.3.4 Matrices de escores para alineación
 - 1.3.5 Software MEGA

UNIDAD II. El lenguaje de programación Python, Bash y la Bioinformática

Competencia:

Generar programas de análisis, a través de los lenguajes de programación Python y Bash, para generar programas de análisis de secuencias de información biológica, con actitud proactiva y colaborativa.

Contenido:

Duración: 8 horas

2.1 Introducción al lenguaje de programación Python y Bash.

2.1.1 Aplicaciones de Python y Bash en la bioinformática.

2.2 Análisis de secuencias biológicas usando Python

2.2.1 Diseño e implementación de un programa para lectura de secuencias biológicas (ADN, ARN y proteínas).

2.2.2 Diseño e implementación de un programa para deducir la secuencia complementaria de una hebra de ADN.

2.2.3 Diseño e implementación de un programa para traducir una hebra de ADN.

2.2.4 Diseño e implementación del algoritmo de programación dinámica para alineamiento de dos cadenas de ADN.

2.3 Definición del problema de modelado familias de secuencias biológicas.

2.4 Tipos de modelos para familias biológicas.

2.4.1 Secuencias de consenso.

2.4.2 Expresiones regulares.

2.4.3 Logos de secuencias.

2.4.4 Matrices de scores de posición específicas (PSSM).

2.4.5 Modelos de Markov

UNIDAD III. El material genético codificante

Competencia:

Comparar secuencias de ácidos nucleicos y proteínas, por medio de algoritmos de alineamiento, para realizar análisis filogenéticos, con actitud crítica, creativa y de respeto al medio ambiente.

Contenido:

Duración: 6 horas

- 3.1 Estudio de las partes estructurales básicas del ADN codificante.
 - 3.1.1 Codones de inicio y paro.
 - 3.1.2 Definición de Reading Frame.
 - 3.1.3 Definición de UTR (Untranslated region).
 - 3.1.4 Definición de promotor.
 - 3.1.5 Intrones y Exones.
- 3.2 Introducción a las proteínas.
 - 3.2.1 Análisis del código genético.
 - 3.2.2 Estructura de las proteínas.
 - 3.2.3 Diagramas de Ramachandran.
 - 3.2.4 Herramienta bioinformática Swiss PDB-Viewer.
- 3.3 Introducción al análisis filogenético y sus aplicaciones.
 - 3.3.1 Procedimiento general para realizar análisis filogenéticos.
- 3.4 Paradigmas para generar arboles filogenéticos.
 - 3.4.1 Métodos basados en Distancia.
 - 3.4.2 Métodos basados en Máxima Parsimonia.
 - 3.4.3 Métodos basados en Máxima Verosimilitud
- 3.5 Implementación del método UPGMA (Unweighted Pair Group Method with Arithmetic Mean).
- 3.6 Implementación del método Neighbor Joining.

UNIDAD IV. Tecnologías de secuenciación de ADN

Competencia:

Aplicar tecnologías de secuenciación masiva, para el ensamble de genomas y metagenomas y conocer la arquitectura cromosómica, por medio de algoritmos de ensamblados de novo y referencia, con orden, disciplina y trabajo colaborativo.

Contenido:

Duración: 6 horas

- 4.1 Introducción a la genómica.
- 4.2 Estrategias de secuenciación de genomas completos.
 - 4.2.1 Secuenciación Jerárquica.
 - 4.2.1 Secuenciación por Shotgun.
- 4.3 Técnicas de ensamblado de genes y genomas.
 - 4.3.1 Algoritmos de corrección de errores de secuenciación.
 - 4.3.2 Conteo de K-mers.
 - 4.3.3 Obtención del valor de K y estimación del tamaño de genomas.
 - 4.3.4 Ensamble de novo
 - 4.3.5 Ensamble por referencia
- 4.4 Metagenómica de amplicones y shotgun
- 4.5 Técnicas de anotación de genes y genomas.
- 4.6 Tecnologías de secuenciación de ADN
 - 4.5.1 Método de secuenciación de Sanger y Maxam-Gilbert.
 - 4.5.2 Método de secuenciación de la segunda generación.
 - 4.5.3 Método de secuenciación de la tercera generación.
 - 4.5.4 Método de secuenciación de la cuarta generación.

UNIDAD V. Análisis bioinformáticos de genomas y metagenomas

Competencia:

Aplicar análisis bioinformáticos, a través de las herramientas R y MEGAN, para realizar estudios de asociación, de estructura genética y caracterización de genomas y metagenomas, con objetividad y disposición al trabajo equipo.

Contenido:

Duración: 6 horas

- 5.1 Introducción a la herramienta R project.
 - 5.1.1 Ambiente y lenguaje de programación en R.
 - 5.1.2 Paqueterías de R para realizar análisis bioinformático.
 - 5.1.3 Análisis bioinformáticos en R.
- 5.2 Introducción a la genómica de poblaciones.
- 5.3 Tecnologías de salida masiva para genómica de poblaciones
 - 5.3.1 SNPchips para genotipificación de salida masiva.
 - 5.3.2 Genchips para expresión genética.
- 5.4 Técnicas de Metagenómica.
 - 5.4.1 Software MEGAN
- 5.5 Proyecto final de bioinformática

VI. ESTRUCTURA DE LAS PRÁCTICAS DE TALLER

No. de Práctica	Competencia	Descripción	Material de Apoyo	Duración
UNIDAD I				
1	<p>identificar las características de las bases de datos biológicas, mediante la exploración de las bases de datos, para la búsqueda de secuencias de ácidos nucleicos y proteínas, con pensamiento analíticos y organizado.</p>	<p>Realiza búsqueda de ácidos nucleicos y proteínas ingresa a distintas bases de datos y explora los elementos que la integran (índices, subíndices, formatos de descarga, opciones de alineamiento).</p> <p>identifica elementos de para el diseño de plásmidos de expresión de proteínas recombinantes e identificación de elementos GTRACK para la edición de genomas.</p> <p>entrega mapa del plásmido, alineamientos de secuencias, con la información recabada.</p>	Equipo de cómputo con acceso a internet. bases de datos, software	8 horas
UNIDAD II				
2	<p>instalar lenguajes de programación, por medio de la línea de comandos, para generar programas y realizar el análisis de secuenciación, con actitud ordenada y analítica.</p>	<p>Descarga de internet las herramientas para programar en el lenguaje Python e implementar distintos programas para leer archivos con secuencias de ADN, deducir las cadenas</p>	Equipo de cómputo con acceso a internet	8 horas

		<p>complementarias, seleccionar codones y producir las proteínas correspondientes.</p> <p>Implementa distintas herramientas bioinformáticas para modelar múltiples secuencias. Las técnicas serán LOGOS, Modelos Escondidos de Markov, secuencias de consenso, y expresiones regulares.</p> <p>Entrega los scripts.</p>		
UNIDAD III y IV				
3	<p>identificar genes en genomas y metagenomas completos, y secuencias de proteínas y alineamientos de secuencias de ADN, por medio de lenguajes de Python y Bash, para conocer la sintenia y funcionalidad del genoma, y distancias evolutivas, con pensamiento crítico y proactivo.</p>	<p>Reconoce la sintenia y funcionalidad del genoma.</p> <p>El alumno descarga de una base de datos el genoma completo del virus del VIH. Posteriormente diseñará e implementará un programa en lenguaje Python para buscar todos los genes que tiene el genoma.</p> <p>Entrega el script del análisis de genomas.</p>	<p>Equipo de cómputo con acceso a Internet, así como las herramientas de programación para el lenguaje Python</p>	4 horas
4		<p>El alumno utiliza el software Swiss-Pdb-Viewer para modelar y analizar proteínas en forma tridimensional, así como validar su modelo con análisis de Ramachandran y entrega modelado tridimensional.</p>	<p>Equipo de cómputo con acceso a Internet, con el software Swiss-pdb-viewer instalado.</p>	3 horas
5		<p>Reconoce las distancias evolutivas por medio de alineamientos de secuencias de ácidos nucleicos y aminoácidos. Calcula distancias</p>	<p>Equipo de cómputo con una hoja de cálculo (como Excel) instalada.</p>	3 horas

		<p>evolutivas. Implementa los algoritmos UPGMS y Neighbour Joining generando distintos árboles filogenéticos. Entrega árbol filogenético con distancias evolutivas.</p>		
UNIDAD V				
6	<p>Identificar secuencias reguladoras cis y elementos trans en genomas procariones y eucariotes, por medio de las herramientas Bioconductor de lenguaje R, para conocer la diferencia transcripcional entre las células procariontas y eucariotas, con actitud crítica y ordenada.</p>	<p>realiza procesamiento de secuencias de ADN y proteínas, aplicando la herramienta R. identifica los elementos reguladores cis y trans. identifica las diferencias de regulación entre las células procariontas y eucariotas. entrega análisis comparativo.</p>	<p>Equipo de cómputo con el software R instalado.</p>	<p>6 horas</p>

VII. MÉTODO DE TRABAJO

Estrategia docente (enseñanza):

Exposición utilizando pizarrón y presentaciones electrónicas

Elaboración de los exámenes y cuestionarios de actividades para evaluar en el alumno la comprensión de los temas expuestos

Seleccionar las bases de datos públicas y elaborar los ejercicios prácticos comprendidos en las competencias

Realizar secciones de preguntas a los alumnos con las que se evaluara la comprensión de los temas expuestos

Formular un proyecto adecuado para los estudiantes, generar grupos de trabajo y ejecutar el proyecto.

Estrategia alumno (aprendizaje)

Lectura de temas en los libros de textos y artículos para elaboración de resúmenes

Participación objetiva, clara y con fundamento en los temas expuestos al grupo

Resolver los exámenes teóricos de conocimiento de los temas expuestos

Ejecución de los ejercicios prácticos

Ejecución del proyecto formulado por el maestro, y preparación de una exposición, a ser presentada al grupo, sobre el trabajo y resultados obtenidos del proyecto.

VIII. CRITERIOS DE EVALUACIÓN

La evaluación será llevada a cabo de forma permanente durante el desarrollo de la unidad de aprendizaje de la siguiente manera:

Criterios de acreditación

- Para tener derecho a examen ordinario y extraordinario, el estudiante debe cumplir los porcentajes de asistencia que establece el Estatuto Escolar vigente.
- Calificación en escala del 0 al 100, con un mínimo aprobatorio de 60.

Criterios de evaluación

Evaluación continua:

2 exámenes parciales escritos30 %

Ejercicios prácticos..... 50 %

evidencia de desempeño.....20 %

(Elabora y entrega un reporte con el análisis comparativo de secuencias de estructuras de ácidos nucleicos y proteínas, el cual debe integrar: análisis e interpretación de resultados. Cada ejercicio debe de contener: portada, marco teórico, planteamiento del problema, la metodología realizada, cálculos, resultados, conclusiones y bibliografía.)

total.....100%

Criterios de evaluación:

1. Los exámenes teóricos serán sobre los temas de las dos primeras competencias. Estos deberán ser contestados de forma individual y de forma clara y concisa
2. Los ejercicios prácticos consistirán en el acceso a una base de datos pública sobre información genérica y/o genómica. La obtención de información sobre secuencias biológicas de interés, y la aplicación de las herramientas algorítmicas necesarias para obtener el análisis requerido por el maestro
3. Los reportes deberán ser entregado una semana después de la asignación del ejercicio práctico y se evaluará tanto el contenido de la información como la redacción y claridad de la misma
4. La exposición final deberá ser presentada con proyector y se podrá utilizar sistemas multimedios para enriquecer la presentación

Criterios actitudinales: Colaboración equitativa, ética y de responsabilidad hacia el interior como al exterior del programa de curso.

IX. REFERENCIAS

Básicas

- Bassi, S. (2017). *Python for Bioinformatics* (2ª ed.). Estados Unidos: CRC Press. ISBN: 9781138035263.
- Brown, S.M. (2015). *Next-Generation DNA Sequencing Informatics* (2ª ed.). Estados Unidos: CSHL Press. ISBN: 9781621821236.
- Model, M. (2010). *Bioinformatics Programming Using Python*. Estados Unidos: O'reilly. ISBN: 9780596154509. [clásica]
- Keith, J. M. (2017). *Bioinformatics Volume I: Data, Sequence Analysis, and Evolution* (2ª ed.) Estados Unidos: Springer. ISBN: 9781493966219.
- Keith, J. M. (2017). *Bioinformatics Volume II: Structure, Function, and Applications*. (2ª ed.) Estados Unidos: Springer. ISBN: 9781493966110.
- Naruya S. (2018). *Introduction to evolutionary genomics* (2ª ed.). Gran Bretaña: Springer. ISBN: 9783319926414.
- Xuhua X. (2013). *Comparative Genomics*. Estados Unidos: Springer. ISBN: 9783642371455. [clásica]

Complementarias

- Lodish, H., Arnold, B., Kaiser, C., Krieger, M., y Matthew, P. (2016). *Biología celular y molecular*. Editorial Médica Panamericana (7ª ed.).
- Watson, J. D.; Baker T.; Bell S., Gann A., Levine M. y Losick, R. (2016). *Biología Molecular del gen*. Editorial Médica Panamericana.

X. PERFIL DEL DOCENTE

El docente que imparta esta asignatura debe contar con título de biología, biología computacional, bioinformática o áreas afines, con posgrado en bioinformática, ciencias genómicas o biología computacional.