



Programa de estudio Datos generales de la Unidad de Aprendizaje

Identificación	
Nombre: Bioinformática	Etapas: Metodológica
Clave:	Tipo de curso: Optativo
Modalidad educativa: Presencial	Modalidad de enseñanza-aprendizaje: Curso-Taller-Seminario
Número de horas: 128 al semestre	Créditos: 8
Secuencia anteriores: Ninguna Colaterales: Ninguna Posteriores: Ninguna	Requisitos de admisión: Ninguna
Fecha de elaboración: Abril del 2020	Fecha de aprobación:

1. Justificación y fundamentos

El estudiante del Doctorado en Recursos Naturales y Ecología de la opción terminal Ecología y Conservación es un posgraduado con alta personalidad científica, capaz de realizar investigaciones científicas con un conocimiento integral sobre el manejo sustentable de los recursos naturales. La bioinformática es una herramienta imprescindible para la caracterización, análisis e interpretación de bases de datos biológicas, las cuales han mostrado un crecimiento exponencial como resultado del desarrollo de las técnicas de secuenciación de nueva generación. Por ello, en esta unidad de aprendizaje se abordan los conceptos básicos de bioinformática y se capacita al alumno para la búsqueda y empleo de bases de datos biológicas disponibles en internet. Además, se abordan los fundamentos de las técnicas de secuenciación masiva y el análisis de datos genómicos y de transcriptomas.





2. Objetivo general

Que el alumno desarrolle competencias necesarias para el manejo básico de lenguaje de programación y bases de datos, además de que sea capaz de manipular las diferentes herramientas bioinformáticas para el análisis de secuencias genómicas y de transcriptomas.

Objetivos particulares

- Que sea capaz de encontrar información de los sistemas biológicos en las bases de datos internacionales.
- Que sea capaz de realizar alineamientos múltiples de secuencias para la búsqueda de motivos y dominios en biomoléculas
- Que conozca las diferentes plataformas de secuenciación masiva y características principales en la elaboración de las librerías.
- Que el alumno se familiarice con las principales técnicas bioinformáticas para el análisis de genomas.
- Que el alumno se familiarice con las principales técnicas bioinformáticas para el análisis de transcriptomas.

3. Competencias a desarrollar

Conocimientos	Habilidades y destrezas	Valores
Bioinformática, concepto y alcances. Conocimiento básico de bases de datos biológicas	Saber cómo buscar la información en las bases de datos disponibles en la red	Generar interés por la bioinformática y la enorme disponibilidad de recursos disponibles en la red
Alineamiento de secuencias y búsqueda de motivos	Entender los fundamentos matemáticos de los algoritmos para alineamiento de secuencias	Disposición para trabajar en equipo y compartir sus conocimientos





Tecnologías de secuenciación masiva y construcción de librerías	Comprender las distintas metodologías de secuenciación masiva y construcción de librerías y formatos de archivos de lecturas	Promover interés y por el creciente desarrollo de las técnicas de NGS Curiosidad por el desarrollo de librerías y formatos de archivos de lectura
Análisis de datos genómicos	Manejo de las herramientas bioinformáticas básicas para la manipulación, control de calidad y análisis de datos genómicos	Trabajo en equipo Promover el sentido de investigación sobre genómica
Análisis de transcriptomas	Manejo de las herramientas bioinformáticas básicas para la manipulación, control de calidad y análisis de transcriptomas	Trabajo en equipo Promover el sentido de investigación sobre transcriptómica

4. Contenidos

Unidad 1. Introducción a la bioinformática y bases de datos

- Introducción a la bioinformática: definición y aplicaciones
- Tipos de bases de datos: primarias, secundarias, especializadas
- Bases de datos de nucleótidos
- Bases de datos de genomas
- Bases de datos de proteínas
- Recuperación de información de bases de datos

Unidad 2. Similitud de secuencias: alineamientos y búsquedas

- Concepto de homología y similitud
- Alineamientos locales y globales
- Métodos de alineamientos pareados
- Métodos de alineamientos múltiples
- Búsqueda de motivos de ADN
- Búsqueda de dominios en proteínas





Unidad 3. Plataformas de secuenciación masiva y formatos de datos

- La bioinformática en la era de las ciencias genómicas
- Tecnologías de secuenciación masiva y aplicaciones en ecología
- Construcción y características de librerías
- Formatos de los archivos de lecturas de secuenciación masiva
- Funcionalidades del sistema operativo Unix

Unidad 4. Análisis de polimorfismos de un solo nucleótido

- Pre-procesado de lecturas: calidad, demultiplexado, eliminación de adaptadores
- Algoritmos para ensamble de genomas *de novo*
- Comparación con genomas de referencia
- Determinación de polimorfismos
- Anotación funcional de loci presentes en regiones codificantes
- Generación de archivos de salida en formatos utilizables para análisis de genética poblacional, filogenética, o mapeo

Unidad 5. Análisis de transcriptomas

- Pre-procesado de lecturas: calidad, eliminación de adaptadores y datos redundantes
- Algoritmos para ensamble de genomas/transcriptoma *de novo*
- Comparación con genomas/transcriptomas de referencia
- Anotación funcional de transcritos
- Cuantificación de niveles de expresión
- Análisis de expresión diferencial de genes

5. Orientaciones didácticas

- Presentar al inicio del curso el objetivo de la asignatura y su relación con otras del plan de estudios, así como el contenido y las actividades de aprendizaje.
- Presentar clases teóricas frente al grupo en cada una de las unidades
- Ejercicios sobre el análisis de datos biológicos en cada clase
- Resumen de la clase, desarrollo de preguntas y ejercicios, en cada tema, para reforzar los puntos más importantes de cada unidad.
- Lectura y discusión de artículos científicos para reforzar técnicas aprendidas en clase





- Realización de exposiciones por parte de los estudiantes, mesa de análisis y discusión de acuerdo con los temas y lecturas.

6. Actividades de aprendizaje

Bajo la conducción del docente	Trabajo independiente del alumno
Exposición del profesor	<p>En el aula</p> <p>Revisión de lectura Resolución de ejercicios Exposiciones Exámenes</p> <p>Fuera del aula</p> <p>Trabajos de Investigación Realización de tareas individuales Síntesis de lecturas Lectura de libros de texto, de consulta o artículos</p>
Trabajo en equipo	
Exposición de los estudiantes	
Discusiones sobre la información teórica en las sesiones de trabajo	
Ejercicios en clase	

7. Evaluación

Este curso debe ser evaluado atendiendo al logro del objetivo general propuesto. Por tanto, para evaluar este logro se plantea que la evaluación se haga sobre la base de dos criterios: del dominio teórico y el dominio de la aplicación práctica. Las formas de evaluación que se utilizarán son:

- Asistencia 10%
- Tareas, ejercicios y participación en clase. 20%
- Exposiciones 30%
- Control de lecturas 20%
- Laboratorio 20%





8. Bibliografía básica y complementaria

Bibliografía básica

Augen J. (2004). Bioinformatics in the post-genomic era: Genome, transcriptome, proteome, and information-based medicine. Addison-Wesley Professional.

Bailey T.L., Baker M.E. y Elkan C.P. (1997). An artificial intelligence approach to motif discovery in protein sequences: application to steroid dehydrogenases. The Journal of steroid biochemistry and molecular biology, 62(1): 29-44.

Hertz G.Z., Hartzell III G.W. y Stormo G.D. (1990). Identification of consensus patterns in unaligned DNA sequences known to be functionally related. Bioinformatics, 6(2): 81-92.

Ouellette B.F.F. y Baxevanis A. (2001). Bioinformatics: A practical guide to the analysis of genes and proteins. John Wiley. 457 pp.

Shaik N.A., Hakeem K.R., Banaganapalli B. y Elango R. (2019). Essentials of Bioinformatics, Volume I: Understanding Bioinformatics: Genes to Proteins. Springer.

Simpson J.T. y Pop M. (2015). The theory and practice of genome sequence assembly. Annual review of genomics and human genetics. 16: 153-172.

Bibliografía complementaria

Andrews K.R., Good J.M., Miller M.R., Luikart G. y Hohenlohe P.A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics. 17(2): 81.

Ansorge W.J. (2009). "Next-generation DNA sequencing techniques. New biotechnology 25(4): 195-203.

Catchen J., Hohenlohe P.A., Bassham S., Amores A. y Cresko W.A. (2013). Stacks: an analysis tool set for population genomics. Molecular ecology, 22(11): 3124-3140.

De Heredia U.L. (2016). Las técnicas de secuenciación masiva en el estudio de la diversidad biológica. Munibe Ciencias Naturales. Natur zientziak, (64), pp.7-31.

Escobar C.A.M., Murillo L.V.R. y Soto J.F. (2011). Tecnologías bioinformáticas para el análisis de secuencias de ADN. Scientia et technica, 3(49): 116-121.

Claverie J.M. y Notredame C. (2011). Bioinformatics for dummies. John Wiley & Sons.





Mastretta-Yanes A., Arrigo N., Alvarez N., Jorgensen T.H., Piñero D. y Emerson B.C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular ecology resources*, 15(1): 28-41.

9. Perfil del profesor

El docente que imparta esta Unidad de Aprendizaje deberá contar con nivel de doctorado con experiencia probada en bioinformática, biología computacional o ciencias bioquímicas

