

FICHA DE ASIGNATURA

Título: Secuenciación genómica y análisis de variantes

Descripción: Esta asignatura se centra en profundizar en uno de los pipelines más utilizados en bioinformática, el de DNAseq. Desde que las muestras salen del secuenciador, hasta la obtención de resultados, veremos todo el procesamiento bioinformático necesario. Para ello nos centraremos en la secuenciación genómica y análisis de variantes, mediante técnicas de resecuenciación y de mapeado "de novo". Utilizaremos para ello las herramientas más actuales y de mayor distribución, basándonos en software libre. Además veremos el diseño inicial de experimentos y herramientas de análisis de resultados.

Carácter: Obligatorio

Créditos ECTS: 6

Contextualización: Esta asignatura forma parte de la materia del Máster Bioinformática Genómica (18 ECTS). Durante esta materia se conocerán con detalle tres flujos de trabajo muy utilizados en bioinformática (DNAseq, RNAseq y metagenómica). El objetivo de cada asignatura en esta materia será el de mostrar desde el diseño experimental hasta el análisis de resultados pasando por el pipeline bioinformática para cada uno de los casos.

Modalidad: Online

Temario: Los contenidos que trataremos en la asignatura son:

- Diseños experimentales para la secuenciación genómica y análisis de variantes.
- Procesado, limpieza y filtrado de secuencias: uso de las herramientas bioinformáticas de procesado, limpieza y filtrado de secuencias: FastQC, Trimmomatic.
- Resecuenciación: uso de las herramientas bioinformáticas de mapeado y alineamiento de secuencias BWA, Bowtie2.
- Visualización de alineamientos: uso de las herramientas de visualización de alineamientos Integrative Genome Viewer (IGV).
- Bioinformática para la medicina personalizada: biomarcadores, búsqueda y filtrado de variantes, polimorfismo puntual de un solo nucleótido (SNPs), variación en el número de copias (CNVs). Uso de las herramientas de búsqueda y filtrado de SNPs, InDels y reordenamientos cromosómicos: Samtools, VarScan, VCFtools, coverage across the reference, GATK, CNVKit.
- Secuenciación "de novo": uso de las herramientas de secuenciación "de novo" y ensamblaje de secuencias SOAPdenovo, Velvet, SPAdes.
- Análisis e interpretación biológica de resultados, principales herramientas de análisis bioestadístico en R aplicadas a la secuenciación genómica y análisis de variantes.

Competencias:

Competencias específicas

- CE1: Ejecutar programas de Python en el entorno de la bioinformática.
- CE2: Analizar ficheros biológicos mediante el lenguaje de programación Python.
- CE3: Interpretar los resultados de los análisis bioinformáticos en el lenguaje de programación Python.
- CE4: Ejecutar programas de R en el entorno de la bioinformática.
- CE5: Analizar ficheros biológicos mediante el lenguaje de programación R.
- CE6: Interpretar los resultados de los análisis bioinformáticos en el lenguaje de programación R.
- CE7: Utilizar los principales lenguajes de programación en cluster en bioinformática.
- CE9: Aplicar herramientas bioinformáticas de R en el análisis bioestadístico de datos.
- CE10: Evaluar los resultados de los análisis bioestadísticos de datos ómicos.
- CE11: Analizar los principales formatos de secuencias en la aplicación de datos ómicos.
- CE12: Extraer la información necesaria de las principales bases de datos de deposito de información biológica en la resolución de problemas.
- CE15: Conocer los principales métodos, herramientas y técnicas en la secuenciación genómica, en la valoración de la correcta metodología para cada flujo de trabajo en bioinformática.
- CE16: Establecer los distintos parámetros que definen la calidad de las secuencias que se obtienen de los secuenciadores.
- CE17: Aplicar los principales métodos de selección y mejora de calidad de secuencias en la bioinformática.
- CE18: Diseñar el flujo de trabajo aplicando los principios generales del diseño de experimentos ómicos.
- CE19: Aplicar los principales algoritmos de alineamiento de secuencias de datos ómicos.
- CE20: Identificar las principales herramientas de análisis de datos para la medicina personalizada.
- CE21: Analizar grandes volúmenes de datos mediante aplicaciones bioinformáticas en medicina personalizada.
- CE22: Analizar resultados de datos ómicos generados por las principales herramientas bioestadísticas.

Actividades Formativas:

Actividad Formativa	Horas	Presencialidad
---------------------	-------	----------------

Clases expositivas	12	0%
Clases prácticas (Estudio de casos, resolución de problemas)	12	0%
Tutorías	10	0%
Trabajo autónomo	114	0%
Prueba final	2	100%

Metodologías docentes:

Metodologías docentes	
Lección magistral	El profesor expone los contenidos de la asignatura sin intervención del estudiante.
Estudio de casos	El profesor facilita al estudiante herramientas para facilitar el aprendizaje activo y que este adquiera las competencias asignadas a la materia.
Resolución de problemas	La finalidad de esta metodología es favorecer la consecución de un grado elevado de autonomía intelectual mediante un planteamiento concreto formulado por el profesor.

Sistema de Evaluación:

Sistemas de evaluación	Ponderación mínima	Ponderación máxima
Evaluación del portafolio. Estudio de casos	20	50
Evaluación del portafolio. Resolución de problemas	20	50
Prueba Final	40	60

Normativa específica: *(En el caso de que haya prerrequisitos)*

Bibliografía: *Autor. (Año de publicación.) Título en itálicas (edición). Lugar de publicación: Casa publicadora.*

Ejemplo:

- Elgar, R. (1965). Introduction to the double bass (2ª Edición) Lussex: Raymond Elgar