

## DATOS BÁSICOS DE LA GUÍA DOCENTE:

<b>Materia:</b>	INTRODUCCIÓN A LA BIOINFORMÁTICA		
<b>Identificador:</b>	33664		
<b>Titulación:</b>	DOBLE GRADO EN FARMACIA Y BIOINFORMÁTICA. PLAN 2019		
<b>Módulo:</b>	BIOINFORMÁTICA		
<b>Tipo:</b>	OBLIGATORIA		
<b>Curso:</b>	1	<b>Periodo lectivo:</b>	Segundo Cuatrimestre
<b>Créditos:</b>	6	<b>Horas totales:</b>	150
<b>Actividades Presenciales:</b>	64	<b>Trabajo Autónomo:</b>	86
<b>Idioma Principal:</b>	Castellano	<b>Idioma Secundario:</b>	Inglés
<b>Profesor:</b>	RANERA BELTRAN, BEATRIZ (T)	<b>Correo electrónico:</b>	branera@usj.es

## PRESENTACIÓN:

Se realiza una introducción a la disciplina de la bioinformática estableciendo los conceptos básicos necesarios para poder ser aplicados en próximos cursos. Se repasan los fundamentos de biología molecular, dogma central, interiorizando los conceptos de ADN/ ARN/ proteínas (bases, codones, amino ácidos, etc) y su relación entre ellos a través de las diferentes ómicas para poder entender las principales técnicas HTS que generan datos que requieren de bioinformática. Se exploran los fundamentos de los análisis de secuencias, los tipos, algoritmos principales y usos, cuantificación de la similitud de las secuencias basándose en coincidencias, inserciones, deleciones y número de copias. Establecer el origen de estas posibles alteraciones: polimorfismos, mutaciones y evolución.

Principales bases de datos bioinformáticas, compresión de su organización, el tipo de mantenimiento que tienen, utilización y relación entre ellas.

Iniciación a la programación en los lenguajes R y Python. Explicación de la utilidad de cada uno de ellos y trabajo práctico para interiorizar la lógica de funcionamiento aplicando conocimientos adquiridos en asignaturas de carácter informático. Se trabajarán ejemplos prácticos bioinformáticos para manipular secuencias, interpretar datos obtenidos e integrarlos en un contexto biológico.

## COMPETENCIAS PROFESIONALES A DESARROLLAR EN LA MATERIA:

<b>Competencias Generales de la titulación</b>	G01	Utilizar estrategias de aprendizaje de forma autónoma para su aplicación en la mejora continua del ejercicio profesional.
	G02	Realizar el análisis y la síntesis de problemas propios de su actividad profesional y aplicarlos en entornos similares.
	G03	Cooperar para la consecución de resultados comunes mediante el trabajo en equipo en un contexto de integración, colaboración y potenciación de la discusión crítica.
	G04	Razonar de manera crítica basándose en la información, datos y líneas de actuación y su aplicación en temas relevantes de índole social, científico o ético.
	G05	Comunicar en lengua castellana y/ o inglesa temas profesionales en forma oral y escrita.
	G06	Resolver los problemas o imprevistos complejos que surgen durante la actividad profesional dentro de cualquier tipo de organización y la adaptación a las necesidades y exigencias de su entorno profesional.
	G07	Elegir entre diferentes modelos complejos de conocimiento para su aplicación a la resolución de problemas.
	G09	Aplicar las tecnologías de la información y comunicación en el ámbito profesional.
	G10	Aplicar la creatividad, independencia de pensamiento, autocrítica y autonomía en el ejercicio profesional.
	<b>Competencias Específicas de la titulación</b>	E02
E03		Aplicar los conceptos fundamentales de matemáticas, lógica, algorítmica y complejidad computacional para la resolución de problemas propios de la bioinformática.
E04		Programar aplicaciones de forma robusta, correcta, y eficiente, eligiendo el paradigma y los lenguajes de programación más adecuados, aplicando los conocimientos sobre procedimientos algorítmicos básicos y usando los tipos y estructuras de datos más apropiados.

E05	Implementar aplicaciones fundamentadas, previamente diseñadas y analizadas, en las características de las bases de datos.
E06	Aplicar los principios fundamentales y técnicas básicas de los sistemas inteligentes y su aplicación práctica en el ámbito de la bioinformática.
E07	Aplicar los principios, metodologías y ciclos de vida de la ingeniería de software al desarrollo de un proyecto en el ámbito de la bioinformática.
E12	Aplicar los principios y técnicas del modelado computacional de proteínas para predecir su función biológica, su actividad o nuevas dianas terapéuticas (Bioinformática estructural, Toxicología Computacional).
E13	Aplicar las tecnologías ómicas para la extracción de información estadísticamente significativa y para la creación de bases relacionales de biodatos que puedan ser actualizables y accesibles públicamente a la comunidad científica.
E14	Utilizar los lenguajes de programación, de mayor uso en el ámbito de las Ciencias de la Vida, para desarrollar y evaluar técnicas y/o herramientas computacionales.
E15	Inferir la historia evolutiva de genes y proteínas mediante la creación e interpretación de árboles filogenéticos.
E16	Planificar estudios de ligamiento y asociación con fines médicos y medioambientales.
E17	Inducir relaciones complejas entre muestras mediante la aplicación de técnicas estadísticas y de clasificación.
E18	Aplicar métodos estadísticos y computacionales para resolver problemas en los campos de la biología molecular, la genómica, la investigación médica y de la genética de poblaciones.
E21	Aplicar técnicas computacionales y de procesamiento de datos para la integración de conceptos y datos físicos, químicos y biológicos para la descripción y/o predicción de la actividad de una sustancia en un determinado contexto.

### REQUISITOS PREVIOS:

Se recomienda que el alumno tenga claros conceptos básicos de bioquímica, biología molecular, biología celular, así como una base de programación informática, como qué son las estructuras de control y poder aplicarlas.

### PROGRAMACIÓN DE LA MATERIA:

#### Contenidos de la materia:

<b>1 - Bases de datos biológicos</b>
1.1 - Introducción. Tipos.
1.2 - Bases de datos de ácidos nucleicos. Tipos de datos almacenados.
1.3 - Bases de datos de proteínas. Tipos de datos almacenados
1.4 - Otras bases de datos. Metiloma. Metaboloma. Cáncer. Compuestos químicos.
<b>2 - Alineamiento de secuencias</b>
2.1 - Definición y tipos.
2.2 - Alineamiento de pares de secuencias. Algoritmos.
2.3 - BLAST
2.4 - Alineamiento de múltiples secuencias.
<b>3 - Principales ómicas y técnicas asociadas a ellas</b>
3.1 - Descripción de las ómicas existentes y relación entre ellas. Dogma central de la biología molecular.
3.2 - Next Generation Sequencing. Microarray. RNAseq. Espectometría.
<b>4 - R y Bioconductor</b>
4.1 - Instalación R y RStudio. Introducción R. Conceptos básicos del lenguaje en R.
4.2 - Entrada y salida de datos. Operaciones con vectores. Paquetes en R.
4.3 - Manipulación de datos. Bucles y funciones condicionales. Gráficos en R.
4.4 - Estadística básica. Tipos de variables. Análisis descriptivos y distribuciones para exploración de datos.
4.5 - Bioconductor. Análisis básico de secuencias. Paquete ape.
4.6 - Caso de uso: microarrays. Anotaciones.
<b>5 - Lenguaje Python</b>
5.1 - Entorno de programación. Versiones e intérpretes disponibles.
5.2 - Sintaxis y tipos de variables.
5.3 - Iteración y operaciones lógicas.

5.4 - Librerías específicas para la programación en bioinformática.

La planificación de la asignatura podrá verse modificada por motivos imprevistos (rendimiento del grupo, disponibilidad de recursos, modificaciones en el calendario académico, etc.) y por tanto no deberá considerarse como definitiva y cerrada.

## METODOLOGÍAS Y ACTIVIDADES DE ENSEÑANZA Y APRENDIZAJE:

### Metodologías de enseñanza-aprendizaje a desarrollar:

La asignatura exige un esfuerzo importante por parte del alumno para aplicar los conceptos de cada tema en los sucesivos. Por tanto, se recomienda un seguimiento continuo de la asignatura. Para ello la principal actividad que se propone consiste en el estudio individual y la realización de ejercicios por parte del alumno.

Las sesiones presenciales podrán ser clases teóricas utilizando la clase magistral a través de presentaciones o mostrando directamente el workflow necesario para la resolución de problemas bioinformáticos. Las clases prácticas podrán ser clases con problemas a resolver de cada una de las unidades o talleres en los que se demuestran casos reales de utilización de herramientas bioinformáticas. En todas ellas se hará participe en la medida de lo posible al alumno, procurando modular los niveles de dificultad en función de la comprensión del grupo de alumnos.

Se realizará trabajo de laboratorio utilizando las herramientas bioinformáticas aprendidas en las dos primeras unidades para el análisis en profundidad de los aspectos relacionados con una proteína dada, atendiendo a las variaciones polimórficas, resultado o no de mutaciones, y las consecuencias desde nivel molecular a nivel fisiológico de las mismas.

Además, se potenciará el trabajo individual a través de la resolución de un proyecto de análisis bioinformático utilizando el entorno de R, para que los alumnos pongan en práctica la exploración de datos, la posibilidad de establecer contrastes de hipótesis, resolver el problema planteado por la hipótesis y establecer conclusiones.

Finalmente, como apoyo al desarrollo de trabajo autónomo del alumno, necesario para asimilación de conceptos de programación, tanto en R como Python, se plantearán ejercicios propuestos como casos a los alumnos que deberán resolver y entregar.

El alumno presentará en todo momento apoyo por parte de la docente de la materia a través del correo electrónico y tutorías presenciales. Además, si fuera necesario podrían plantearse foros presenciales de ayuda, así como foros online a través de la PDU, en los que toda la clase sea participe y se contribuyan a resolver los problemas de forma conjunta.

### Volumen de trabajo del alumno:

Modalidad organizativa	Métodos de enseñanza	Horas estimadas
<b>Actividades Presenciales</b>	Clase magistral	25
	Casos prácticos	4
	Resolución de prácticas, problemas, ejercicios etc.	20
	Exposiciones de trabajos de los alumnos	2
	Talleres	5
	Actividades de evaluación	6
	Asistencia a actividades externas (visitas, conferencias, etc.)	2
<b>Trabajo Autónomo</b>	Asistencia a tutorías	4
	Estudio individual	30
	Preparación de trabajos individuales	26
	Tareas de investigación y búsqueda de información	9
	Lecturas obligatorias	8
	Lectura libre	5
	Asistencia a actividades externas (visitas, conferencias, etc.)	4
<b>Horas totales:</b>		<b>150</b>

## SISTEMA DE EVALUACIÓN:

### Obtención de la nota final:

Pruebas escritas:	20	%
Trabajos individuales:	45	%
Prueba final:	35	%
<b>TOTAL</b>	<b>100</b>	<b>%</b>

\*Las observaciones específicas sobre el sistema de evaluación serán comunicadas por escrito a los alumnos al inicio de la materia.

## **BIBLIOGRAFÍA Y DOCUMENTACIÓN:**

### **Bibliografía básica:**

Pevsner, Jonathan. Bioinformatics and Functional Genomics 3rd Edition. New York, Wiley Blackwell, 2015
Model, Mitchell L. Bioinformatics Programming Using Python. Sebastopol, CA: O'Reilly, 2010
Crawley, Michael J. The R Book. New York: Wiley, 2012.

### **Bibliografía recomendada:**

Xiong, Jin. Essential Bioinformatics. Cambridge, UK: Cambridge Press, 2006
Braun WJ, Murdoch DJ. A First Course in Statistical Programming with R. New York, USA: Cambridge University Press, 2016
Antao T. Bioinformatics with Python Cookbook. PacktLib, 2015

### **Páginas web recomendadas:**

The Pevsner laboratory	<a href="http://www.bioinfbook.org/">http://www.bioinfbook.org/</a>
Bioconductor	<a href="http://bioconductor.org/">http://bioconductor.org/</a>
BLAST	<a href="https://blast.ncbi.nlm.nih.gov/Blast.cgi">https://blast.ncbi.nlm.nih.gov/Blast.cgi</a>
Clustal	<a href="https://www.ebi.ac.uk/Tools/msa/clustalo/">https://www.ebi.ac.uk/Tools/msa/clustalo/</a>
Documentación R	<a href="https://www.rdocumentation.org/">https://www.rdocumentation.org/</a>
Problemas Python	<a href="http://rosalind.info/problems/locations/">http://rosalind.info/problems/locations/</a>
Documentación Python	<a href="https://docs.python.org/2.7/">https://docs.python.org/2.7/</a>