



PRESENTACIÓN

Breve descripción:

En los últimos años, las *tecnologías de alto rendimiento* aplicadas en el ámbito de la biología y la medicina nos han permitido detectar y cuantificar de forma simultánea miles de variantes génicas, transcritos o proteínas. Esto ha hecho posible la consolidación de las denominadas **-ómicas**. Por ejemplo, en el ámbito de la genómica y la transcriptómica los **microarrays de ADN** y la **secuenciación masiva (NGS)** han permitido el estudio de las bases moleculares de diversas enfermedades, la búsqueda de nuevos biomarcadores clínicos y la detección de nuevas dianas terapéuticas. Estos resultados pueden ser integrados con el objetivo de obtener el perfil ómico de un individuo, lo que convierte en una realidad el concepto de **medicina personalizada**. La gran cantidad de datos experimentales que se obtienen de estos experimentos y la complejidad de los métodos de análisis necesarios para obtener resultados de interés biológico han hecho necesaria la aparición de un área de conocimiento denominada **biología computacional o bioinformática**. Esta asignatura siempre tiene una base estadística, imprescindible para la toma de decisiones objetivas y científicas. Desde ella se afrontan los problemas informáticos, metodológicos y matemáticos derivados de estos estudios.

- **Titulación:** DATA SCIENCE
- **Módulo/Materia:** Módulo I: Investigación de medicamentos. Materia: Descubrimiento de fármacos
- **ECTS:** 3 ECTS
- **Curso, semestre:** 1er semestre
- **Carácter:** Obligatoria
- **Profesorado:** Ibon Tamayo Uria y Marta Garcia-Granero Marquez
- **Idioma:** Inglés
- **Aula, Horario:** [VER CALENDARIO](#)

COMPETENCIAS

CB6. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB9. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

CB10. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

CG4. Identificar y saber crear estrategias y acciones dirigidas a la consecución de los objetivos planteados y concretar los recursos necesarios para llevarlas a cabo, en el ámbito de la empresa farmacéutica.

CG5. Conocer las técnicas y tendencias actuales relacionadas con la investigación, el desarrollo y la innovación de medicamentos.

CG6. Poseer capacidad crítica para tomar las decisiones necesarias y adaptarse a las nuevas situaciones que puedan surgir en el ámbito de la empresa farmacéutica y afines.



CE1. Diseñar nuevos fármacos y terapias innovadoras para la industria farmacéutica aplicando técnicas especializadas de quimioinformática y bioinformática.

PROGRAMA

INTRODUCCIÓN A LA INFORMÁTICA (3 horas)

- Linux
- Línea de comandos
- Cluster
- Github
- Slack
- Markdown

INTRODUCCIÓN A LA ESTADÍSTICA (12 horas)

- Conceptos básicos de estadística
- Estadística descriptiva
- Distribución de probabilidad
- Inferencia estadística paramétrica
- Inferencia estadística no paramétrica
- Análisis de frecuencias
- Correlación y regresión
- Análisis de supervivencia

INTRODUCCIÓN AL LENGUAJE DE PROGRAMACIÓN R (12 horas)

- Fundamentos
- Objetos en R
- Importación y exportación de datos
- Gráficos en R
- Iniciación a la programación en R
- Prácticas de iniciación a R

INTRODUCCIÓN A LOS EXPERIMENTOS DE ALTO RENDIMIENTO (3 horas)

- Introducción a las tecnologías ómicas
- Secuenciación masiva (NGS)
- Métodos de análisis de datos en bioinformática
- Flujos de trabajo y estructuras básicas
- Sesiones prácticas

ACTIVIDADES FORMATIVAS

(3 ECTS x 25 = 75 h). El curso se compone de actividades presenciales (1,20 ECTS; 30 horas) y actividades no presenciales (1,80 ECTS; 45 horas), como sigue:

- Clases teórico-prácticas de estadística y programación: 28 horas



Universidad
de Navarra

EVALUACIÓN

CONVOCATORIA ORDINARIA

La evaluación de la asignatura se realizará teniendo en cuenta los siguientes criterios:

- Examen escrito consistente en la resolución de un cuestionario que incluye contenidos teóricos y prácticos : 80% (40% estadística y 40% programación).
- Presentación de las prácticas realizadas a lo largo de la asignatura, de un trabajo personal y participación en clase: 20 % (todos los informes de las prácticas deberán ser enviados antes de la fecha indicada por el profesor).

Para aprobar la asignatura será necesario obtener una puntuación superior a 4 sobre 10 en cada una de las dos partes del examen.

La nota final será el promedio ponderado de cada una de ellas.

CONVOCATORIA EXTRAORDINARIA

- Los alumnos que no aprueben en la convocatoria ordinaria (puntuación mínima de 5 sobre 10), deberán volver a realizar un examen con las mismas características anteriormente citadas en la convocatoria ordinaria.

HORARIOS DE ATENCIÓN

Ibon Tamayo Uría: itamayou@unav.es

Marta García Granero: mggranero@unav.es

Horarios de atención

- Sin horario definido. Enviar email a los profesores para concertar cita

BIBLIOGRAFÍA

<https://cran.r-project.org/>

Recursos principales de la asignatura

El material obligatorio para el seguimiento de la asignatura será:

- Presentaciones realizadas por el profesor para la impartición de las clases (este material se dejará en ADI antes del comienzo de las clases).
- Informes para la realización de las prácticas en la sala de ordenadores. Se dejarán en ADI antes de la realización de las mismas.

Lecturas recomendadas



Artículos relacionados con las ómicas, la tecnología de microarrays y la secuenciación masiva (NGS):

- Oltvai ZN, Barabási AL. **Life's complexity pyramid**. Science. 2002 Oct 25;298(5594):763-4.
- Kahvejian A, Quackenbush J, Thompson JF. **What would you do if you could sequence everything?**. Nat Biotechnol. 2008 Oct;26(10):1125-33.
- Green ED, Guyer MS; National Human Genome Research Institute. **Charting a course for genomic medicine from base pairs to bedside**. Nature. 2011 Feb 10;470(7333):204-13.
- Benjamin M. Bolstad. **Pre-Processing DNA Microarray Data**. **Fundamentals of Data Mining in Genomics and Proteomics** (Chapter 3). Springer.
- Holloway AJ, van Laar RK, Tothill RW, Bowtell DD. **Options available--from start to finish--for obtaining data from DNA microarrays II**. Nat Genet. 2002 Dec;32 Suppl:481-9
- Abdullah-Sayani A, Bueno-de-Mesquita JM, van de Vijver MJ. **Technology Insight: tuning into the genetic orchestra using microarrays--limitations of DNA microarrays in clinical practice**. Nat Clin Pract Oncol. 2006 Sep;3(9):501-16.
- Martin JA, Wang Z. **Next-generation transcriptome assembly**. Nat Rev Genet. 2011 Sep 7;12(10):671-82. doi: 10.1038/nrg3068. Review.
- Mardis ER. **Next-generation DNA sequencing methods**. Annu Rev Genomics Hum Genet. 2008;9:387-402. doi: 10.1146/annurev.genom.9.081307.164359. Review.
- Matthew D. Young , Davis J. McCarthy , Matthew J. Wakefield, Gordon K. Smyth , Alicia Oshlack and Mark D. Robinson. **Differential Expression for RNA Sequencing (RNA-Seq) Data: Mapping, Summarization, Statistical Analysis, and Experimental Design**. **Bioinformatics for high throughput sequencing (Chapter 10)**. Naiara Rodríguez-Ezpeleta, Michael Hackenberg, Ana M. Aransay Editors. Springer.
- Park PJ. **ChIP-seq: advantages and challenges of a maturing technology**. Nat Rev Genet. 2009 Oct;10(10):669-80. doi: 10.1038/nrg2641.
- Quackenbush J. **Computational analysis of microarray data**. Nat Rev Genet. 2001 Jun;2(6):418-27.
- Tarca AL, Romero R, Draghici S. **Analysis of microarray experiments of gene expression profiling**. Am J Obstet Gynecol. 2006 Aug;195(2):373-88.
- Normand R, Yanai I. **An introduction to high throughput sequencing experiments: design and bioinformatics analysis**. Methods Mol Biol. 2013;1038:1-26. doi: 10.1007/978-1-62703-514-9_1.

Material para el aprendizaje de R/Bioconductor

- Manual gratuito: W. N. Venables, D. M. Smith and the R Core Team. **An Introduction to R**. (Notes on R: A Programming Environment for Data Analysis and Graphics). (<https://cran.r-project.org/doc/manuals/R-intro.pdf>)
- Manual gratuito: John Verzani. **simpleR: Using R for Introductory Statistics**. (<http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf>)
- Libro: **Dalgaard, Peter. Introductory Statistics with R** (Springer) 2008. [Localízalo en la Biblioteca](#)
- Libro: Gentleman, R., Carey, V., Huber, W., Irizarry, R., Dudoit, S. (Eds.) **Bioinformatics and Computational Biology Solutions Using R and Bioconductor** (Springer) 2005. [Acceso electrónico](#) ; [Localízalo en la Biblioteca](#)

Páginas web

- The Comprehensive R Archive Network (<https://cran.r-project.org/>)
- Bioconductor: Open source for bioinformatics (<https://www.bioconductor.org/>)



Universidad
de Navarra