



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

1. GUÍA DOCENTE DE BIOINFORMÁTICA ESTRUCTURAL

La presente guía docente corresponde a la asignatura BIOINFORMÁTICA ESTRUCTURAL (STRBIO), aprobada para el curso lectivo 2017-2018 en Junta de Centro y publicada en su versión definitiva en la página web de la Escuela Politécnica Superior. La guía docente de STRBIO aprobada y publicada antes del periodo de matrícula tiene el carácter de contrato con el estudiante.



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

2. ASIGNATURA

BIOINFORMÁTICA ESTRUCTURAL (STRBIO)

1.1. Código

33082 del Máster en Bioinformática y Biología Computacional

1.2. Materia

Análisis de secuencias

1.3. Tipo

Obligatoria

1.4. Nivel

Máster

1.5. Curso

1º

1.6. Semestre

1º

1.7. Número de créditos

3 ECTS

1.8. Requisitos previos

Se indican los requisitos para cursar con aprovechamiento la asignatura “Filogenia Molecular” así como las asignaturas en las que se cubren dichos requisitos

- Conocimientos de estructura de ácidos nucleicos y proteínas
- Conocimientos de evolución molecular



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

Estos conocimientos se deben haber adquirido en cualquiera de los grados en ciencias de la vida/salud o en la asignatura “Complementos Docentes Bioquímica y Biología Molecular”

Para garantizar la asimilación de los contenidos y la adquisición de habilidades se recomienda

- la lectura crítica de los textos de la bibliografía
- el uso del material electrónico disponible en la plataforma Moodle (<http://moodle.uam.es>)
- la revisión de material complementario en la red.

Es recomendable disponer de un dominio de inglés para leer la bibliografía de consulta.

Asimismo, se requiere

- Trabajo individual antes de las sesiones teóricas y prácticas: Lectura del material a cubrir con el objetivo de familiarizarse con los conceptos que se van a manejar y los problemas a resolver.
- Trabajo individual posterior a las sesiones:
 - Revisión de las notas tomadas en clase.
 - Consulta de la bibliografía.
 - Realización de ejercicios.

1.9. Requisitos mínimos de asistencia a las sesiones presenciales

La asistencia a clase no es obligatoria

1.10. Datos del equipo docente

Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Dr. Michael Tress (Co-cordinador)

Centro Nacional de Investigaciones Oncológicas (CNIO)

Despacho: Structural Computational Biology Group

Teléfono: +34 91 372 8000

Correo electrónico: mtress@cnio.es

Página web: <http://ubio.bioinfo.cnio.es/people/mtress/>

Dr. Marco Trevisan (Co-cordinador)

Centro Nacional de Investigaciones Cardiovasculares (CNIC)

Despacho: Proteómica cardiovascular

Teléfono: +34 914 53 12 00 (ext 1200)

Correo electrónico: mtrevisan@cnic.es



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

Página web: <https://www.cnic.es/wiki/proteomica/index.php/User:Mtrevisan>

Dr. Luis del Peso (Coordinador)

Departamento de Bioquímica
Facultad de Medicina
Laboratorio: 252 (Instituto de Investigaciones Biomédicas)
Teléfono: +34 91 585 44 40
Correo electrónico: luis.peso@uam.es

Página web: <https://www.bq.uam.es/miembros/lpeso/index.php>

1.11. Objetivos del curso

En este curso se tratan las aproximaciones algorítmicas utilizadas para predecir la estructura primaria, secundaria y terciaria de la proteínas, así como cuantificar su expresión (proteómica). Además de entender las bases teóricas, se estudiará cómo aplicar estos métodos en la práctica para predicción de estructuras y análisis de expresión. Puesto que la bioinformática es un área en rápida evolución, más que centrarnos en el uso concreto de una herramienta específica de cada aplicación se hará énfasis a los formatos de entrada, parámetros e interpretación de resultados comunes a muchas de ellas.

Las **competencias específicas** que se adquieren en esta asignatura son:

- CE1 - Capacidad de aplicar los conocimientos de biología, matemáticas, física y estadística a la bioinformática.
- CE2 - Capacidad de aplicar métodos computacionales a la resolución de problemas en el ámbito de la investigación biomédica básica y traslacional.
- CE3 - Capacidad de analizar, explotar e investigar secuencias biológicas comparativa, evolutiva y funcionalmente.
- CE4 - Capacidad de predecir estructuras, funciones, y dinámica de biomoléculas.
- CE9 - Capacidad de explotación de la información de las bases de datos biológicas sobre secuencias, estructuras, transcriptomas, genomas, proteomas, etc.

Al final del semestre (objetivos generales), y de cada unidad (objetivos por tema) el estudiante deberá ser capaz de:

Comprender las bases teóricas, tanto conceptuales como algorítmicas, de la predicción y análisis de estructura de macromoléculas y su dinámica.
Aplicar herramientas bioinformáticas de visualización, predicción y análisis de estructura e interacciones de macromoléculas, e interpretar los resultados de estos programas.



Integrar los resultados de modelado de estructuras para resolver problemas biológicos.

OBJETIVOS GENERALES	
G1	Identificar las principales bases de datos de estructuras biológicas, así como acceder y extraer información de ellas.
G2	Comprender las bases teóricas, tanto conceptuales como algorítmicas, de la predicción y análisis de estructura de macromoléculas
G3	Aplicar herramientas bioinformáticas de visualización, predicción y análisis de estructura e interacciones de macromoléculas, e interpretar los resultados de estos programas.
G4	Integrar los resultados de modelado de estructuras para resolver problemas biológicos.
G5	Comprender la base experimental y computacional de la caracterización de proteomas.

OBJETIVOS ESPECIFICOS POR TEMA	
TEMA 1.- Bases de datos de estructuras	
1.1.	Conocer las principales bases de datos de estructuras de macromoléculas biológicas y de clasificación de proteínas
1.2.	Ser capaz de hacer búsquedas y extraer información de las bases de datos de estructuras
TEMA 2.- Visualización de estructuras 3D	
2.1.	Conocer y ser capaz de usar programas de visualización de estructuras
TEMA 3.- Predicción de estructura secundaria y 3D	
3.1	Comprender las principales estrategias algorítmicas para la predicción de estructura 2D y 3D de proteínas
3.2	Interpretar los resultados de los métodos de predicción y entender sus limitaciones
3.3	Ser capaz de emplear programas específicos para predecir estructuras 2D y 3D
TEMA 4.- Proteómica	
4.1	Comprender las bases del análisis de proteomas y la identificación de proteínas por huella peptídica
4.2	Entender y aplicar las técnicas de identificación de proteínas por fragmentación (MS/MS)
4.2	Conocer las bases de la proteómica cuantitativa

1.12. Contenidos del programa

Programa Sintético

TEMA 1. Bases de datos de estructuras



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

TEMA 2. Visualización de estructuras 2D y 3D
TEMA 3. Predicción de estructura secundaria y 3D
TEMA 4. Proteómica

Programa Detallado

1. Bases de datos de estructuras
 - 1.1. Protein Data Bank
 - 1.2. SCOP y CATH
2. Visualización de estructuras 2D y 3D
 - 1.1. Pymol y Rasmol
3. Predicción de estructura secundaria y 3D
 - 3.1. Predicción de estructura secundaria y otras características 1D
 - 3.2. Predicción de estructura 3D mediante modelado por homología
 - 3.3. Predicción de estructura 3D mediante métodos más complejos
 - 3.4. Evaluación de modelos
4. Proteómica
 - 4.1. Proteómica por espectrometría de masas
 - 4.2. Identificación de proteínas utilizando bases de datos, por huella peptídica (peptide mass fingerprinting), e identificación de péptidos por fragmentación (MS/MS, o MS2)
 - 4.3. Fundamentos de proteómica cuantitativa

1.13. Referencias de consulta

Bibliografía:

Libros

- Structural Bioinformatics 2 by Jenny Gu (Redactor), Philip E. Bourne (Redactor) Wiley-Blackwell 2nd edition 2011
- A basic introduction to protein structure and function: Protein Structure and Function Gregory A. Petsko (Autor), Dagmar Ringe (Autor) OUP Oxford 2nd edition (2008)

Artículos

- **Artículo donde se presenta el programa Comet (véase más abajo):**
Eng, J, Jahan, TA, and Hoopmann, MR, «Comet: An open-source MS/MS sequence database search tool», *Proteomics*, 2012
<http://onlinelibrary.wiley.com/doi/10.1002/pmic.201200439/full>



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

- **Artículo en el que se presenta el XCorr, el parámetro principal utilizado por Comet para identificar péptidos:** Eng, J, McCormack, AL, and Yates III, JR, «An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database», *Journal of the American Society for Mass Spectrometry*, 1994
<http://www.sciencedirect.com/science/article/pii/S1044030594800162>
- **Artículo donde se habla de la principal estrategia para identificar proteínas, utilizando bases de datos señuelo (*target-decoy approach*):** Elias, JE, and Gygi, SP, «Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry», *Nature Methods*, 2007
<http://www.nature.com/nmeth/journal/v4/n3/abs/nmeth1019.html>
- **Revisión sobre proteómica basada en espectrometría de masas:** Aebersold, R, and Mann, M, «Mass spectrometry-based proteomics», *Nature*, 2003
<http://www.nature.com/nature/journal/v422/n6928/abs/nature01511.html>

Recursos adicionales

- **Comet:** programa de código abierto para la identificación de secuencias peptídicas. Descargable gratuitamente desde aquí: <http://comet-ms.sourceforge.net>
- **Mascot:** programa (de código no abierto) para la identificación de secuencias peptídicas. Se puede utilizar en red sin necesidad de instalar nada. La versión gratuita tiene ciertas limitaciones, pero no afectan a la práctica individual. Se puede encontrar aquí: http://www.matrixscience.com/search_form_select.html
- «**Interpreting MS/MS Proteomics Results**», de Brian Searle: Una presentación con diapositivas amena y didáctica que cuenta bastantes detalles sobre cómo se identifican proteínas, http://www-bac.esi.umontreal.ca/~db-cm2003/proteomique/docs/Interpreting_MSMS_results_cartoon.ppt

Material electrónico de trabajo: los documentos electrónicos de trabajo (material adicional, ejercicios del curso y ejemplos de exámenes, etc.) se publican en la sección de IA en plataforma Moodle (<http://moodle.uam.es/>)

2. Métodos docentes

La metodología utilizada en el desarrollo de la actividad docente incluye los siguientes tipos de actividades:

*Clases de teoría:



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

Actividad del profesor

Clases expositivas en las que se realizarán programas y ejercicios ilustrativos. Se utilizará la pizarra, combinada con la explicación de programas en formato electrónico cuya ejecución será proyectada en clase.

En la parte de identificación de proteínas, se trabajarán los conceptos planteados, utilizando la pizarra y diapositivas. Se combinarán con cuestiones y preguntas a los alumnos para estimular el debate y asentamiento de los conceptos presentados.

Actividad del estudiante:

Actividad presencial: Toma de apuntes, participación en clase respondiendo a las cuestiones planteadas. Resolución de los ejercicios propuestos.

Actividad no presencial: lectura del material bibliográfico y de apoyo, estudio de la materia y realización de ejercicios propuestos y de los cuestionarios planteados en la plataforma Moodle.

*Tutorías en aula:

Actividad del profesor:

Tutorización a toda la clase o en grupo reducidos con el objetivo de resolver dudas comunes plantadas por los alumnos a nivel individual o en grupo, surgidas a partir de cuestiones/ejercicios señalados en clase para tal fin y orientarlos en la realización de los mismos.

Actividad del estudiante:

Actividad presencial: Planteamiento de dudas individuales o en grupo y enfoque de posibles soluciones a las tareas planteadas.

Actividad no presencial: Estudio de las tareas marcadas y debate de las soluciones planteadas en el seno del grupo.

*Prácticas:

Actividad del profesor:

Asignar una práctica/proyecto a cada grupo de trabajo y explicar la práctica asignada a cada grupo de trabajo al comienzo de la sesión de prácticas. Los medios utilizados son los entornos de programación, *software* específico de uso local y en servidores remotos y los ordenadores del propio laboratorio para la ejecución y análisis de los programas realizados.

En la segunda parte identificaremos proteínas utilizando herramientas accesibles en red y archivos procedentes de espectrometría de masas proporcionados por el profesor. Identificaremos proteínas tanto por huella peptídica como por fragmentación. Veremos las diferencias entre ambas técnicas, y cómo mejoran o empeoran los resultados en función de los parámetros utilizados.

Actividad del estudiante:



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

Actividad presencial: Planteamiento inicial, previo al desarrollo de la práctica, sobre información contenida en el enunciado. Debate en el seno del grupo sobre el planteamiento de la solución óptima. Ejecución de la práctica y entrega de resultados. Los alumnos participarán activamente en la identificación de las proteínas con los espectros proporcionados.

Actividad no presencial: Profundizar en el enunciado de la práctica, plantear el diagrama de flujo óptimo para la resolución de la misma e implementar los ejercicios propuestos. Los alumnos dispondrán de archivos adicionales para hacer más búsquedas fuera de clase si lo desean, así como interpretar los resultados obtenidos.

3. Tiempo de trabajo del estudiante

		Nº de horas	Porcentaje
Presencial	Clases teóricas	14 h (52%)	27 h (36%)
	Clases prácticas/ejercicios	10 h (37%)	
	Realización de pruebas escritas parciales y final	3 h (11%)	
No presencial	Estudio semanal	11 h (22.9%)	48 h (64%)
	Realización de actividades prácticas	10 h (20.8%)	
	Preparación del examen (convocatoria ordinaria)	8 h (16.7%)	
	Preparación del examen (convocatoria extraordinaria)	19 h (39.6%)	
Carga total de horas de trabajo: 25 horas x 3 ECTS		75 h	

4. Métodos de evaluación y porcentaje en la calificación final

Todas las partes de la asignatura, teoría y prácticas puntúan sobre 10.

La calificación final de la asignatura en la convocatoria ordinaria se obtiene a partir de las notas de examen(es) de teoría (Teoría) y calificación obtenido en los ejercicios y/o prácticas y participación en clase (Prácticas) mediante la ecuación:

$$\text{Calificación: } 0.5 * \text{Prácticas} + 0.5 * \text{Teoría}$$

Para aprobar la asignatura es obligatorio obtener una nota mayor o igual a 5 puntos tanto en la parte de teoría como en las prácticas. En caso contrario, la nota final en actas será

$$\text{Calificación: } \text{Mín}(4.9, 0.5 * \text{Prácticas} + 0.5 * \text{Teoría})$$



Asignatura: Bioinformática Estructural
Código: 33082
Centro: Escuela Politécnica Superior
Titulación: Máster en Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3

La calificación de teoría será media aritmética ponderada (por cantidad contenido incluido en el examen) de los exámenes que se realicen. La pruebas no realizadas por causas sin justificar computan como cero en la media.

La calificación de prácticas será media aritmética ponderada (por cantidad contenido incluido en el examen) de los ejercicios que se realicen. Los retrasos en las entregas de ejercicios/prácticas serán penalizados con el 25% de la calificación inicial por día de retraso.

La calificación de la parte práctica tendrá en cuenta la calidad de los diseños realizados, los resultados obtenidos y de su presentación.

Para la convocatoria extraordinaria

- Los estudiantes deben revisar las entregas realizadas completándolas y corrigiendo sus errores sobre la versión entregada durante el curso.
- Realizarán un examen escrito o/y oral

ATENCIÓN: Cualquier copia descubierta que se haya realizado a lo largo del curso, tanto en cualquiera de las actividades de teoría desarrolladas, ejercicios, y debate como en cualquiera de los apartados de las prácticas, serán penalizadas con rigurosidad. La penalización por copia implica la aplicación de la normativa interna de la EPS, que supone suspender la convocatoria actual y no poder presentarse a la siguiente.

5. Cronograma

Semana	Contenido	Horas presenciales	Horas no presenciales
1	Introducción a las bases de datos de estructuras Visualización de estructuras	5	6
2	Predicción de estructuras	8	9.5
3	Predicción de estructuras Proteómica	8	9.5
3	Proteómica	3	4
	Examen Final convocatoria ordinaria	3	8
	Examen Final convocatoria extraordinaria	3	19