

Master en Bioinformática y Biología Computacional

Organización institucional y programa académico

Madrid - 29/08/2016



**PYTHON &
BIGDATA**



**NETWORK &
SYSTEMS BIOLOGY**



**R & DATA
ANALYSIS**



**GENOMICS
NGS**



**PRECISION
MEDICINE**

Índice de contenido

Organización.....	3
Estructura y metodología.....	4
Primera fase: Período lectivo.....	4
Segunda fase: Prácticas.....	5
Evaluación final:.....	5
Programa.....	6
Módulo 0 – Cursos de Nivel Cero.....	6
0.1 - Programación nivel cero.....	6
0.2 - Introducción a la Biología Molecular.....	6
Módulo 1 - Análisis de Secuencias y Bioinformática Estructural.....	8
1.1 - Análisis de secuencias.....	8
1.2 - Filogenia.....	9
1.3 - Estructura de Proteínas.....	9
Módulo 2 - Informática y Programación.....	11
2.1 - Informática: Linux y <i>Cloud Computing</i>	11
2.2 - Programación en Python (24h).....	11
2.3 - BigData Handling and Processing.....	12
Módulo 3 - Estadística y Análisis de Datos.....	14
3.1 - Programación y Estadística en R.....	14
3.2 - Fundamentos de análisis de datos.....	15
3.3 - Text Mining.....	15
Módulo 4 – Genómica y NGS.....	17
4.1 - Introducción a la NGS.....	17
4.2 - Metagenómica y genómica de procariontes.....	17
4.3 - Genómica Traslacional.....	18
4.4 - Transcriptómica.....	19
4.5 - Regulación Genómica e Epigenómica.....	20
Módulo 5 – Biología de Sistemas y Redes.....	21
5.1 - Redes Biológicas y Biología de Sistemas.....	21
5.2 - Moléculas pequeñas y farmacogenómica.....	23
Módulo 6 – Actividad Profesional e Investigación en Bioinformática.....	24

Organización

El máster está organizado por la Escuela Nacional de Sanidad (ENS) del Instituto de Salud Carlos III (ISCIII), la Fundación Española para la Cooperación Internacional Salud y Política Social (FCSAI) en colaboración con el Centro Nacional de Investigaciones Oncológicas (CNIO).

Director:

- Alfonso Valencia (CNIO)

Directores honorarios:

- Federico Moran (UCM)
- Luis Vázquez (UCM)
- Francisco Montero (UCM)

Coordinación académica y secretaría:

- Miguel Ponce de León – coordinacion@masterbioinformatica.com
- Laura Cáceres – secretaria@masterbioinformatica.com

Profesorado

Las clases han sido impartidas en las anteriores ediciones por una plantilla de más de 40 profesores, procedentes de centros de investigación, universidades y empresas nacionales. Además, varios investigadores de centros internacionales son invitados para dar seminarios sobre temas específicos en varias áreas de la Bioinformática

Instituciones organizadoras:



Estructura y metodología

El máster se encuentra dividido en dos fases de trabajo y una fase final de evaluación. La primera fase corresponde a la parte lectiva y abarca un total de ~400h de clase presenciales distribuidas en 7 módulos conceptuales.

Los módulos son los siguiente:

- Módulo 0 – Cursos de Nivel Cero (Biología/Programación)
- Módulo 1 – Análisis de Secuencias y Bioinformática Estructural
- Módulo 2 – Informática y Programación
- Módulo 3 – Estadística y Análisis de Datos
- Módulo 4 – Genómica y NGS
- Módulo 5 – Biología de Sistemas y Redes
- Módulo 6 – Actividad Profesional e Investigación en Bioinformática



La segunda fase del curso corresponde a la realización de las prácticas de fin de máster, las cuales consisten en el desarrollo, por parte del estudiante, de un proyecto de investigación. Estos proyectos tienen una duración aproximada de entre 200 y 300 horas y se realizan en centros de investigación, universidades y empresas, tanto nacionales como de otros países de la UE.

Primera fase: Período lectivo

La primera fase consiste en las clases de teoría y práctica impartidas por los profesores. Algunas de sus características son:

- Programa de unas 400 horas
- 20-25 alumnos
- un ordenador por alumno
- más del 60% en trabajos prácticos con ordenador
- evaluaciones al final de cada módulo, mediante exámenes o trabajos, que se calificarán en una escala de 0 a 10. Los alumnos suspendidos tendrán una oportunidad para recuperar.

Para superar la parte lectiva, el alumno deberá:

- haber asistido a al menos el 85% de las horas de cada módulo
- haber aprobado todos los módulos, o haber suspendido un módulo como máximo, y con una nota superior a 5.
- haber abonado la totalidad de las tasas

Segunda fase: Prácticas

Durante la segunda fase (de unas 300 horas) el alumno deberá realizar un proyecto fin de máster en una empresa del sector o en un centro de investigación público o privado.

Habrà un proceso de selección para elegir las prácticas de acuerdo a los intereses de los estudiantes y los de las empresas o centros de investigación, que podrán tener en cuenta las calificación finales de los estudiantes.

En caso de incompatibilidad absoluta de horario o calendario, también existirá la posibilidad de desarrollar un proyecto en casa, bajo la tutoría de un profesor del máster.

Los alumnos deberán escribir una memoria sobre el proyecto realizado con un formato y extensión predefinidos.

Los proyectos de prácticas son desarrollado en centros de investigación del sector público y privado, universidades y empresas, tanto nacionales como de otros países europeos ([ver enlace](#)).

Evaluación final:

Para obtener el título del máster el alumno deberá haber aprobado la fase lectiva, y el proyecto de fin de máster.

La calificación final se hará según no apto, apto y apto mención especial, y se determinará teniendo en cuenta la nota media obtenida durante la fase lectiva en los módulos, y la calificación de la memoria del proyecto "fin de máster".

Programa

Módulo 0 – Cursos de Nivel Cero

Coordinador: Miguel Ponce de León (UCM)

Asignaturas:

- Programación Nivel Cero
- Biología Nivel Cero

Total de horas: 36h

0.1 - Programación nivel cero

Descripción: Esta asignatura pretende introducir los conceptos más elementales sobre programación. Incluye conceptos generales sobre cómo se organizan los ordenadores modernos, la diferencia entre software y hardware, los distintos paradigmas de programación y las diferencias entre lenguajes compilados e interpretados. Empleando el lenguaje de programación Python se introducen los conceptos y elementos básicos del lenguaje, como variables, tipos de datos primitivos y las estructuras de control. El objetivo es que los estudiantes que no han tenido experiencia previa en programación y adquieren los fundamentos necesarios para ser capaces de escribir y ejecutar scripts muy simples.

Coordinador: Miguel Ponce de León

Docentes

Nombre	Afiliación	email
Miguel Ponce de León	Universidad Complutense de Madrid	coordinación@masterbioinformatica.com

Total de horas: 18h

Programa:

1. Introducción a los ordenadores y los lenguajes de programación
2. Primeros pasos en Python: intérprete y tipos de datos
3. Variables, Tipos y Operadores
4. Estructuras de control: If, While, For
5. Manejo básico de ficheros
6. Prácticas y Repaso

0.2 - Introducción a la Biología Molecular

Descripción: Esta asignatura es una introducción (y/o repaso) a los conceptos fundamentales de la bioquímica y la biología molecular. Como tal, el programa incluye el estudio de las propiedades de las macromoléculas biológicas así como las técnicas experimentales empleadas en su estudio. El temario también incluye temas relacionados con el metabolismo intermediario y el metabolismo de la información, así como el sistema inmune.

Coordinador: Mayte Villalba

Docentes

Nombre	Afiliación	email
Mayte Villalba	UCM	mvillalb@quim.ucm.es
Javier Lacadena	UCM	jlacaden@quim.ucm.es

Total de horas: 14h

Programa:

- Tema 1. Estructura de proteínas. Niveles estructurales de proteínas: Dominios, secuencias consenso, proteínas intrínsecamente desestructuradas. Clasificación de proteínas. Desnaturalización. Plegamiento. Chaperonas y chaperoninas. Flexibilidad conformacional.
- Tema 2. Técnicas de separación y caracterización de proteínas. Criterios de separación. Cromatografía de intercambio iónico, penetrabilidad, hidrofobicidad y afinidad. Electroforesis convencional y capilar. Técnicas espectroscópicas. Digestión y secuenciación de proteínas. Espectrometría de masas. Huella peptídica de una proteína. Resolución caso práctico.
- Tema 3. Enzimas. Especificidad enzimática. Composición de centros activos: Aminoácidos claves en la funcionalidad. Cinética enzimática. Mecanismos catalíticos. Diseño de Inhibidores, sustratos, fármacos.
- Tema 4. Sistema Inmune. Estructura de los componentes del sistema inmune. Respuesta inmune humoral: Inmunoglobulinas. Tipos y función. Sistema inmune innato, citoquinas, células reguladoras. Anticuerpos monoclonales y policlonales. Tipos de vacunas.
- Tema 5. Introducción al metabolismo. Conceptos generales. Aspectos metabólicos de los diferentes tejidos. Estudio de una ruta como modelo metabólico. Transducción de señales. Metabolómica. Herramientas para el estudio de perfiles metabólicos. Regulación de la actividad enzimática. Base molecular de patologías genéticas. Ejemplo de patología.
- Tema 6. Estructura de ácidos nucleicos. Nucleótidos y sus propiedades. Estructura del ADN. Cromatina. Desnaturalización y renaturalización. Tipos y estructura del ARN. siRNA.
- Tema 7. Metabolismo de la información en organismos eucariotas. Replicación. Modificaciones del DNA: Reparación, recombinación, transposición, restricción.
- Tema 8. Procesos I. Transcripción. Maduración del RNA. Regulación de la expresión génica en procariontes y eucariotas. Metilación del DNA, silenciamiento de los genes y epigenética. RNA de interferencia. Edición de RNA.
- Tema 9. Procesos II. Traducción y biosíntesis de proteínas. Antibióticos. Modificaciones postraduccionales. Direccionamiento de las proteínas en eucariotas. Destrucción programada.
- Tema 10. Técnicas de Ingeniería Genética. Enzimas de restricción. Electroforesis. Hibridación de ácidos nucleicos: Southern y Northern blot. PCR. Clonaje de ADN. Expresión de ADN recombinante. Mutagénesis dirigida.
- Tema 11. Aplicaciones de la Ingeniería genética. Producción de fármacos. Hibridomas. Medicina forense. Técnicas de diagnóstico genético. Localización de desórdenes genéticos. Determinación del sexo. Expresión diferencial. Producción de transgénicos.

Módulo 1 - Análisis de Secuencias y Bioinformática Estructural

Coordinador del Módulo: Luis del Peso (UAM)

Asignaturas:

- Análisis de Secuencias
- Filogenia
- Estructura de Proteínas y Proteómica

Total de horas: ~72h

1.1 - Análisis de secuencias

Descripción: El análisis de secuencias está en el epicentro de la Biología Computacional. Si bien en muchos campos de la bioinformática se trata con secuencias (ejemplo paradigmático es el análisis de datos de secuenciación de ADN), en este módulo nos centramos en la comparación de secuencias que tienen un origen común. Aunque las secuencias contienen la información para determinar qué estructura y función tiene una proteína, aún no sabemos descifrar dicha información. Pero gracias a la comparación de secuencias, podemos identificar piezas claves, piezas variables, huellas evolutivas en definitiva, y tratar de desentrañar el mensaje cifrado que portan. Además de comprender las bases teóricas (tanto conceptuales como algorítmicas) de la comparación de secuencias, el objetivo de la asignatura es enseñar a utilizar herramientas de búsqueda de homólogos (BLAST, PSI-BLAST, HMMs) y alineamiento múltiple (Muscle, Mafft, Jalview). Las clases consistirán en 6 sesiones que combinarán teoría y casos prácticos para resolver en el aula y 2 sesiones de seminarios que impartiran expertos internacionales en análisis de secuencias.

Coordinador: Luis del Peso

Docentes

Nombre	Afiliación	email
Luis del Peso	UAM, Madrid	lpeso@iib.uam.es
Cedric Notredame	CRG, Barcelona	cedric.notredame@crg.es
Luis Sánchez-Pulido	FGU-MRC DPAG-University of Oxford	sanchezpulido@gmail.com

Total de horas: 27h

Programa:

- Tema 1: Bases de datos de secuencias y modelos estadísticos de secuencias biológicas.
- Tema 2: Alineamiento de pares de secuencias
- Tema 3: Representación de alineamientos múltiples (MSA), bases de datos asociadas y búsquedas con MSA
- Seminario 1: Dominios de proteínas y homología remota. [Luis Sánchez-Pulido]. Seminario por parte del Dr. Luis Sánchez-Pulido experto en la aplicación del análisis de secuencias a problemas concretos de la biología. Su trabajo ha permitido caracterizar la función y estructura de muchas proteínas y nos presentará algunos de sus ejemplos más pedagógicos e inspiradores.
- Seminario 2: Algoritmos para el alineamiento de múltiple secuencias [Cedric Notredame]..

1.2 - Filogenia

Descripción: Esta asignatura presenta los fundamentos generales sobre inferencia filogenética, la metodología para la reconstrucción de árboles filogenéticos y su interpretación. En particular, se hace especial énfasis en las diferentes aplicaciones prácticas del análisis filogenético al estudio de familias de proteínas y a la predicción de regiones funcionalmente importantes de una proteína. La asignatura posee clases prácticas orientadas a enseñar al estudiante cómo construir de manera autónoma árboles filogenéticos e interpretarlos, adquiriendo en ellas las habilidades necesarias para el manejo de herramientas bioinformáticas de uso común en filogenia molecular.

Coordinador: David de Juan

Docentes

Nombre	Afiliación	email
David de Juan	CNIO	dadejuan@cnio.es
Salvador Capella	CNIO/INB	scapella@cnio.es

Total de horas: 18h

Programa :

- Tema 1: Introducción
- Tema 2: Inferencia Filogenética
- Tema 3: Aplicación de árboles filogenéticos al estudio de familias de proteínas
- Tema 4: Predicción de regiones importantes de proteínas basada en información evolutiva
- Examen

1.3 - Estructura de Proteínas

Descripción: Esta asignatura trata de predicción de estructuras 3D de proteínas mediante técnicas computacionales que incluyen el modelado por homología así como métodos más complejos. El temario también incluye la introducción al uso de herramientas para la visualización de estructuras así como la evaluación de modelos predichos.

Coordinador: Michael Tress

Docentes

Nombre	Afiliación	email
Michael Tress	CNIO	mtress@cnio.es
Iakes Ezkurdia	CNIC	iezkurdia@cnic.es
Juan Rodriguez	CNIO	jrodriguezr@cnio.es
Juan Fernandez Recio	BSC	
Paolo Maietta	NimGenetics	pmaietta@nimgenetics.com

Total de horas: 27h

Programa:

1. Bases de datos de estructuras

2. Visualización de estructuras 3D
3. Predicción de estructura secundaria y otras características 1D
4. Predicción de estructura 3D mediante modelado por homología
5. Predicción de estructura 3D mediante métodos más complejos
6. Evaluación de modelos
7. Proteómica

Módulo 2 - Informática y Programación

Coordinador: Miguel Ponce de Leon (UCM)

Bloques:

1. Informática: Linux y Cloud Computing
2. Programación en Python
3. BigData Processing

Total de horas: ~72h

2.1 - Informática: Linux y Cloud Computing

Descripción: Esta asignatura ofrece al alumno unas herramientas básicas con las que enfrentarse al sistema operativo más empleado en Bioinformática, y así aprovechar al máximo el resto de asignaturas prácticas del Máster. Al alumno proveniente de Ingeniería Informática (y similares), esta asignatura sirve de repaso a los aspectos más básicos de Linux. Además, se cuenta con un módulo específico de *cloud computing*, un paradigma en el que, áreas de frontera como la Bioinformática, están encontrando solución a sus problemas computacionales más exigentes. Se ofrecen tutorías a lo largo del curso, esto es, incluso después de la finalización de la propia asignatura.

Coordinador: José Luis Vázquez Poletti

Docentes

Nombre	Afiliación	email
José Luis Vázquez Poletti	Universidad Complutense de Madrid	jlvazquez@fdi.ucm.es

Total de horas: 24h

Programa:

1. Introducción a la informática y al SO Unix/Linux
2. Gestión Básica del Sistema Operativo Linux
3. Gestión de procesos
4. Sistema de Fichero
5. Ordenes avanzadas
6. Intérprete de órdenes y configuración del usuario
7. Cloud Computing

2.2 - Programación en Python (24h)

Descripción: Esta asignatura presenta a los alumnos los conocimientos básicos de programación. Enseñará a los alumnos cómo pasar de las ideas para solucionar un problema a la implementación de un algoritmo que lo soluciona. Para ello utilizaremos como herramienta el lenguaje Python. Se estudiarán las estructuras de datos y control fundamentales, los conceptos básicos de orientación a objetos y se estudiarán módulos básicos útiles en el entorno de la bioinformática.

Coordinador: Daniel Mozos Muñoz

Docentes

Nombre	Afiliación	email
Daniel Mozos Muñoz	Universidad Complutense de Madrid	
Carlos González Calvo	Universidad Complutense de Madrid	carlosgonzalez@fdi.ucm.es

Total de horas: 24h

Programa:

Tema 1: Python básico: Estructuras de datos y de control de flujo. Ficheros

Tema 2: Funciones, orientación a objetos y módulos

Tema 3: Identificación de patrones

Tema 4: Módulos y biopython

2.3 - BigData Handling and Processing

Descripción: Los pasos previos al análisis de datos son la adquisición, extracción, transformación y normalización de los mismos. Estas tareas, además de consumir gran parte del tiempo, son críticas a la hora de minimizar a nivel sintáctico y semántico las incongruencias e inconsistencias de la información a procesar. En esta asignatura se enseñan los principios prácticos programáticos y de línea de comandos básicos a la hora de procesar y manipular grandes volúmenes de datos, haciendo especial énfasis en su aplicación al campo de la bioinformática y la biología computacional. Durante la asignatura se enseñará la forma de afrontar el parseo de formatos de datos comunes en la bioinformática, así como patrones de línea de comandos y técnicas de programación básicos que minimicen las necesidades de memoria y de tiempo de ejecución a la hora de realizar ese parseo.

Coordinador: José María Fernández González

Docentes

Nombre	Afiliación	email
José María Fernández González	INB-GN2, CNIO	jmfernandez@cnio.es
Miguel Vazquez	CNIO	mvazquez@cnio.es
David G. Pisano	Sanitas	

Total de horas: 24h

Programa:

1. El día a día de un Data Scientist.
2. Uso práctico de la línea de comandos para la extracción y correlación de información.
 - Repaso de conceptos: pipes y filtros en la shell
 - Expresiones regulares en acción
 - Otras herramientas
3. Estrategias programáticas de parseo y extracción de datos.
 - Familias de formatos de datos más comunes. Estrategias de parseo de los mismos.
 - Estructura de datos avanzadas relacionadas con el parseo.

- Algunas librerías existentes: BioPython, Pandas.
4. Minería de datos básica mediante bases de datos SQL.
- Operaciones de inserción, actualización, borrado y consulta.
 - Consultas avanzadas: joins, consultas de agregación, etc...
 - Uso programático de bases de datos SQL.

Módulo 3 - Estadística y Análisis de Datos

Coordinador:

Bloques:

- Programación y estadística en R
- Inteligencia Computacional
- Textmining

Total de horas: 72h

3.1 - Programación y Estadística en R

Descripción: Esta asignatura cubre el sistema de programación y análisis estadísticos R y lo utiliza para presentar algunos elementos fundamentales de estadística de uso habitual en el análisis de datos "ómicos" (modelos lineales, ajuste por *multiple testing*, clasificación, etc). En la asignatura se combinan estos dos elementos: utilizamos R para resolver o contestar preguntas típicas en el análisis de datos ómicos (o utilizamos ejemplos de datos ómicos para ver cómo lo resolveríamos con R). Empezamos con una introducción a R (que nos ocupa unos tres días) que nos permite hacernos una idea general del lenguaje. A partir de ahí, profundizamos más (tanto estadísticamente como en el uso de R) en ciertos temas elegidos (por ej., modelos lineales).

Coordinador: Ramón Díaz Uriarte

Docentes

Nombre	Afiliación	email
Ramón Díaz Uriarte	Departamento de Bioquímica, UAM	ramon.diaz@iib.uam.es
Anotni Picornell	Gregorio Marañón Health Research Institute	bio.acpc@gmail.com
Enrique Carrillo	CNIO	ecarrillo@cnio.es

Total de horas: 33h

Programa:

1. Introducción a R
 - a) Uso interactivo y gráficos: dos ejemplos
 - b) Scripts, lectura y escritura de ficheros
 - c) Manejo de Paquetes
 - d) Tipos de objetos (vectores, factores, matrices, listas, data frames)
 - e) Manipulando datos y la familia apply
 - f) Fundamentos de programación en R
2. Modelos lineales
3. Multiple testing
4. Clasificación y predicción
5. Programación en R (II)
6. Más sobre gráficos en R
7. Más modelos lineales

3.2 - Fundamentos de análisis de datos

Descripción: Esta asignatura tiene como objetivo capacitar en la aplicación de técnicas de preprocesado, auditoría y limpieza a conjuntos de datos así como en la utilización de métodos de aprendizaje automático para el análisis de datos.

Coordinador: Gonzalo Martínez Muñoz

Docentes

Nombre	Afiliación	email
David Arroyo	UAM	david.arroyo@uam.es
Ana M ^a González	UAM	ana.marcos@uam.es
Luis Lago	UAM	luis.lago@uam.es
Gonzalo Martínez	UAM	gonzalo.martinez@uam.es
Manuel Sánchez Montañés	UAM	manuel.smontanes@uam.es

Total de horas: 24

Programa:

- Introducción al aprendizaje automático (Ana M^a González)
- Preprocesado de datos: Reducción de la dimensionalidad (Ana M^a González)
- Clasificación y regresión con máquinas de vectores soporte (Luis Lago y Manuel Sánchez Montañés)
- Conjuntos de clasificadores y árboles de decisión (Gonzalo Martínez)
- Clustering (David Arroyo)

3.3 - Text Mining

Descripción: Esta asignatura presenta los aspectos básicos de las técnicas de procesamiento de lenguaje natural (NLP) aplicadas a la literatura biomédica y a la minería de texto. El programa incluye una breve introducción a los principales temas de la NLP, así como los aspectos básicos que caracterizan la estructura de la literatura biomédica. La asignatura también introduce el uso de algunas de las herramientas de software más importantes, desarrolladas para manejar textos de biología molecular. La sesión práctica se dedica principalmente a proporcionar una visión general sobre el desarrollo y uso de la minería de texto y aplicaciones biomédicas .

Coordinador: Martin Krallinger

Docentes

Nombre	Afiliación	email
Martin Krallinger	CNIO	mkrallinger@cnio.es

Total de horas: 15h

Programa:

1. Biomedical literature
 - Importance of free textual data for bioinformatics and biomedicine

- Basic characteristics of biomedical literature
 - Literature repositories for life science
 - Accessibility of literature data
 - Database annotations and literature curation (biocuration)
 - Text-mining assisted literature curation
 - Controlled vocabularies, ontologies and hierarchical terminologies
2. Natural language processing
- Information retrieval
 - Text clustering
 - Text classification
 - Information extraction
 - Named entity recognition (NER)
 - Question answering
 - Knowledge discovery
3. Applications of text mining systems in biology (Bio-NLP)
- Biomedical information retrieval systems
 - Bio-entity recognition: genes, proteins, DNA, RNA, cell lines, species, compounds
 - Protein/gene normalization (grounding)
 - Relation extraction systems: protein interactions
 - Annotation retrieval: associations of proteins with terms (GO, anatomy terms, diseases)
 - Indirect relations
 - Gene ranking and text classification
 - Other tools: mutation detection, methylation, localization, acronyms

Módulo 4 – Genómica y NGS

Coordinador: Gonzalo Gomez

Bloques:

- *Introducción a la NGS*
- *Metagenómica*
- *Genómica Traslacional*
- *Transcriptómica*
- *Epigenómica y regulación*

Total de horas: 105h

4.1 - Introducción a la NGS

Descripción: Esta asignatura presenta una introducción a las técnicas de secuenciación y a las herramientas y técnicas esenciales para el análisis de datos de secuenciación de la nueva generación.

Coordinador: Gonzalo Gómez López

Docentes

Nombre	Afiliación	email
Gonzalo Gómez López	CNIO	ggomez@cnio.es
Jose Luis García	CIB	jlgarcia@cib.csic.es
Miriam Rubio	CNIO	mrubioc@cnio.es

Total de horas: 12h

Programa:

1. Métodos de Secuenciación
2. Introducción a la genómica
3. Alineadores y Formatos para datos de NGS
4. Ensamblaje de genomas *de novo* con datos de NGS, estrategias y recursos de software

4.2 - Metagenómica y genómica de procariontas

Descripción: Esta asignatura busca introducir al estudiante al campo de la Metagenómica, así como su aplicación en áreas como la ecología, las ciencias ambientales y la biomedicina. A lo largo de las clases se presentarán los distintos métodos bioinformáticos empleados en el tratamiento, análisis e interpretación de datos obtenidos en estudios metagenómicos. Realizaremos una primera introducción al análisis de genomas procariontas, para pasar después al trabajo con datos metagenómicos reales. Se utilizarán las aproximaciones más novedosas al análisis metagenómico, basadas en coensamblaje y binning por abundancia, lo que permitirá obtener y separar genomas individuales, que también anotaremos. Por último, veremos el modo de trabajar con datos metagenómicos, realizando la cuantificación de la expresión de los genes encontrados en la muestra.

Coordinador: Javier Tamames

Docentes

Nombre	Afiliación	email
Javier Tamámes	CNB	tamames@cnb.csic.es

Total de horas: 15h

Programa:

1. Genómica
 - Anotación de genomas
 - Validación y refinado de las anotaciones
2. Metagenómica I
 - Ensamblaje y coensamblaje de metagenomas
 - Validación del ensamblaje
 - Predicción de genes y búsqueda de homología
3. Metagenómica II
 - Anotación taxonómica de genes y contigs
 - Anotación funcional de genes
 - Mapeo de lecturas a contigs para obtener valores de abundancia
4. Metagenómica III
 - Mapeo de lecturas
 - Binning: Obtención de genomas a partir de los metagenomas
5. Metatranscriptómica IV
 - Limpieza de rRNAs
 - Cuantificación de la expresión mediante mapeo a referencia
 - Comparación de metagenomas/metatranscriptomas

4.3 - Genómica Traslacional

Descripción: Este módulo se centra en el estudio de variantes genómicas puntuales y estructurales que determinan buena parte de la variabilidad poblacional así como su asociación a enfermedades humanas. Los alumnos recibirán lecciones teórico-prácticas sobre la detección, análisis e interpretación de mutaciones (SNVs, indels) y variantes asociadas al número de copias génicas (CNVs) obtenidos por técnicas de secuenciación masiva así como polimorfismos (SNPs) a través de estudios de asociación (GWAS). Se abordarán además los aspectos relacionados con el estudio epidemiológico de los posibles factores de riesgo asociados a dichas variantes genómicas. Finalmente, se revisará el estado actual de la aplicación de datos genómicos de pacientes en el ámbito clínico y se describirán las técnicas y metodologías existentes para el tratamiento personalizado de enfermedades.

Coordinadora: Fátima Al-Shahrour

Docentes

Nombre	Afiliación	email
--------	------------	-------

Fatima Al-Shahrour	CNIO	falshahrour@cnio.es
Hafid Laayouni	UPF	hafid.laayouni@upf.edu
Nuria Malats	CNIO	nmalats@cnio.es
Elena Piñeiro	CNIO	epineiro@cnio.es
Miriam Rubio Camarillo	CNIO	mrubio@cnio.es
María Rigau	CNIO	mrigau@cnio.es
Javier Suela	NIM Genetics	jsuela@nimgenetics.com

Total de horas: 30h

Programa:

1. Introducción a las variantes genómicas y sus aplicaciones traslacionales.
2. Análisis de polimorfismos en poblaciones mediante SNPs y GWAS.
3. Epidemiología Molecular.
4. Mutaciones I: NGS y detección.
5. Mutaciones II: Anotación.
6. Mutaciones III: Interpretación y consecuencias clínicas.
7. Detección de variantes estructurales (CNV) en poblaciones humanas, bases de la variabilidad. Principales fuentes de datos y estudios, y herramientas para el análisis e interpretación de los resultados.

4.4 - Transcriptómica

Descripción: Esta asignatura pretende mostrar a los alumnos las aplicaciones de la bioinformática en el ámbito de la expresión de los genomas. Para ello, se estudiarán de modo teórico-práctico los métodos de análisis de transcripción génica basados en microarrays y secuenciación masiva (RNA-seq). Durante la asignatura se adquirirán, además, conocimientos teóricos para el análisis funcional de listas de genes (métodos de sobrerepresentación y análisis de enriquecimiento de grupos de genes) que se aplicarán de forma práctica sobre casos reales. Finalmente, los alumnos recibirán nociones básicas sobre métodos de clasificación de muestras en función de sus perfiles de expresión génica.

Coordinador: Gonzalo Gómez López

Docentes

Nombre	Afiliación	email
Gonzalo Gómez López	CNIO	ggomez@cnio.es
Osvaldo Graña Castro	CNIO	ograna@cnio.es
Hector Tejero	CNIO	htejero@cnio.es
Joaquín Dopazo	CIPF	jdopazo@cipf.es
Francisco García	CIPF	fgarcia@cipf.es

Total de horas: 24h

Programa:

1. Introducción al análisis de datos de transcriptómica y microarrays.
2. Bases de datos transcriptómicos y herramientas de software para procesamiento y análisis

- (Bioconductor).
3. Análisis de arrays. Normalización, expresión diferencial, multiple testing. Practicas con Bioconductor.
 4. Análisis de RNA-seq, formatos y métodos computacionales.
 5. Análisis funcional, métodos de bloques de genes (Practicas con GSEA)
 6. Anotación funcional de listas de genes y los métodos de sobrerrepresentación
 7. Métodos clasificación supervisada y no supervisada (clustering) y su aplicación a la interpretación de genes diferencialmente expresados.

4.5 - Regulación Genómica e Epigenómica

Descripción: En esta asignatura tratará sobre como la secuencia de ADN y la configuración de la cromatina regula la expresión génica por la unión de factores de transcripción y cambios epigenéticos. Los alumnos aprenderán la base de los análisis a escala genómica de metilación de ADN, modificaciones de histonas y otras proteínas que se unen al ADN y/o la cromatina mediante técnicas basadas en microarrays y next-generation sequencing (NGS). El alumno se familiarizará con el acceso y utilización este tipo de datos que han sido producidos por el proyecto ENCODE (Encyclopedia of DNA elements).

Coordinador: Enrique Carillo

Docentes

Nombre	Afiliación	email
Enrique Carrillo	CNIO	ecarrillo@cnio.es
Filipe N Were	CNIO	fnicolau@cnio.es

Total de horas: 24h

Programa (tentativo):

1. Regulación genómica, epigenómica y factores de transcripción
2. Análisis de datos de CHIP-seq
3. Detección de estados de la cromatina a partir de datos de CHIP-seq
4. Introducción a la metilación de ADN. Análisis de microarrays de metilación de ADN
5. Análisis de metilación de ADN mediante NGS
6. ENCODE

Módulo 5 – Biología de Sistemas y Redes

Coordinador: Alfonso Valencia y Florencio Pazos

Bloques:

- Farmacogenómica
- Biología de Sistemas

Total de horas: 45h

5.1 - Redes Biológicas y Biología de Sistemas

Descripción: Esta asignatura pretende familiarizar a los alumnos con las nuevas aproximaciones sistémicas al estudio de fenómenos biológicos y con las redes biológicas, que son el prototipo de estudio de estas métodos. Todas estas aproximaciones tienen un componente bioinformático muy grande ya que, en general, implican el procesamiento de grandes cantidades de datos. Los alumnos se familiarizarán con las metodologías y algoritmos usados para el tratamiento bioinformático de redes biológicas así como con el manejo del software que los implementan.

Coordinador: Florencio Pazos Cabaleiro

Docentes

Nombre	Afiliación	email
Florencio Pazos	CNB-CSIC	pazos@cnb.csic.es
Alfonso Valencia	CNIO	valencia@cnio.es
Carlos Aguirre	UAM	Carlos.Aguirre@uam.es
Mónica Chagoyen	CNB-CSIC	monica.chagoyen@cnb.csic.es
Francisco Montero	UCM	tramonte@quim.ucm.es
Ildefonso Cases	GBPA	icases@gmail.com

Total de horas: 24h

Programa:

1. Introducción
2. Redes de Interacciones entre Proteínas
3. Teoría de Grafos
4. Manejo, Visualización y Cálculos Topológicos con Redes Biológicas
5. Características Topológicas y Funcionales de las Redes Metabólicas
6. Análisis Estequiométrico de Redes Metabólicas
7. Redes y Enfermedades
8. Características Topológicas y Funcionales de las Redes de Regulación Génica
9. Otras Redes Biológicas
10. Seminario científico a cargo de un investigador invitado.

5.2 - Moléculas pequeñas y farmacogenómica

Descripción:

- Introducción a las moléculas que tienen afinidad por las macromoléculas biológicas, incluyendo los sustratos de las enzimas y los cofactores enzimáticos, los fármacos, etc, explicando la naturaleza de las fuerzas que determinan la asociación entre moléculas, las características de los sitios de unión, las bases de la selectividad del reconocimiento molecular y las consecuencias de la formación de complejos.
- Introducción a las principales herramientas de modelado molecular y simulación del comportamiento dinámico de los sistemas modelados con detalle atómico.
- Introducción al estudio de las relaciones existentes entre composición química y propiedades biológicas.
- Introducción a los conceptos de quimiotecas, cribado virtual y diseño de fármacos basado en la estructura tridimensional de la macromolécula diana.

Coordinador: Federico Gago Badenas

Docentes

Nombre	Afiliación	email
Federico Gago Badenas	UAH	federico.gago@uah.es
Héctor Tejero	CNIO	htejero@cnio.es

Total de horas: 21h

Programa:

1. Visualización en 3D de estructuras de macromoléculas y sus complejos con ligandos mediante gráficos moleculares interactivos.
2. Movimientos en proteínas y unión de ligandos.
3. Interacción ligando-proteína.
4. Reconocimiento de ADN por ligandos y proteínas.
5. Dinámica molecular.
6. Relaciones cuantitativas estructura-actividad (QSAR)
7. Acoplamiento (docking) ligando-proteína.
8. Cribado virtual y diseño de nuevos ligandos.
9. Farmacogenómica y reposicionamiento de fármacos

Módulo 6 – Actividad Profesional e Investigación en Bioinformática

Coordinador: Alfonso Valencia

Bloques:

- Seminarios de presentación de papers (6h)
- Charlas y conferencias científicas (9h)
- Clases de preparación de Cvs y entrevistas (3h)

Descripción: Este módulo engloba un conjunto de actividades que incluyen la serie de conferencias científicas que se presentan a lo largo del máster, así como los seminarios de presentación y discusión de publicaciones científicas en el área de la bioinformática y las clases de preparación de CV y entrevistas.

Total de horas: 18h

Docentes

Nombre	Afiliación	email
Alfonso Valencia	CNIO	valencia@cnio.es
Isabel Martinez		imartinezroman@yahoo.es