



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

GUÍA DOCENTE DE FUNDAMENTOS DE SECUENCIACIÓN DE ÚLTIMA GENERACION Y GENÓMICA TRASLACIONAL

La presente guía docente corresponde a la asignatura “Fundamentos de Secuenciación de Última Generación y Genómica Traslacional” (SGETR), 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 SGETR aprobada y publicada antes del periodo de matrícula tiene el carácter de contrato con el estudiante.



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

ASIGNATURA

Fundamentos de Secuenciación de Última Generación y Genómica Traslacional (SGETR)

1.1. Código

33088 del Master de Bioinformática y Biología Computacional

1.2. Materia

Módulo IV: Genómica y Análisis de datos de Secuenciación de Última Generación.

1.3. Tipo

Obligatoria

1.4. Nivel

Máster

1.5. Curso

1º

1.6. Semestre

1º

1.7. Número de créditos

6 ECTS

1.8. Requisitos previos

Se indican los requisitos para cursar con aprovechamiento la asignatura “Fundamentos de Secuenciación de Última Generación y Genómica Traslacional” así como las asignaturas en las que se cubren dichos requisitos

- Conocimientos de Genética Molecular (nivel grado).
- Técnicas de Biología Molecular (nivel grado).
- Matemáticas y estadística (nivel de un grado de ciencias).
- Inglés escrito y hablado (nivel B2).



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

- Conocimientos de UNIX/Linux y programación científica (nivel impartido en la asignatura “Programación, Linux y Bases de Datos”)

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 que se estará disponible para los alumnos
- la revisión de material complementario en la red.

Es imprescindible disponer de un dominio de inglés (nivel B2) para leer la bibliografía de consulta.

Asimismo, se recomienda

- 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.
- Trabajo en grupo
 - Realización de las prácticas de la asignatura.
 - Participación en el aula.

1.9. Requisitos mínimos de asistencia a las sesiones presenciales

Asistencia no obligatoria pero muy recomendada si se quieren adquirir las competencias ofertadas por la asignatura.



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

1.10. Datos del equipo docente

Profesores de teoría:

Dra. Fátima Al-Shahrour Núñez (Coordinadora)

Unidad de Bioinformática

Centro Nacional de Investigaciones Oncológicas (CNIO)

Despacho: EA.P0.018

Teléfono: +34 91 732 80 00 ext 2940

Correo electrónico: falshahrour@cnio.es

Página web: <http://www.cnio.es>

Twitter: @BU_CNIO

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

Dr. Gonzalo Gómez López (Coordinador)

Departamento de Bioquímica

Facultad de Medicina, UAM.

Despacho: Unidad de Bioinformática, Centro Nacional de Investigaciones Oncológicas (CNIO)

Teléfono: +34 91 732 80 00 ext 3062

Correo electrónico: ggomez@cnio.es, gonzalo.gomez@uam.es

Página web: <https://www.bq.uam.es/>

Twitter: @BU_CNIO

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

Dra. Nuria Malats

Laboratorio de Epidemiología Genética y Molecular

Centro Nacional de Investigaciones Oncológicas (CNIO)

Despacho: EA.P0.011A

Teléfono: +34 91 732 80 00 ext 3330

Correo electrónico: nmalats@cnio.es

Página web: <http://www.cnio.es>

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

Dr. Hafid Laayouni El Alaoui

Departament de Ciències Experimentals i de la Salut

Biologia Evolutiva

Universitat Pompeu Fabra (UPF)

Teléfono: +34 93 316 08 45

Correo electrónico: hafid.laayouni@upf.edu

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

Dr. José Luis García López

Centro Investigaciones Biológicas (CIB)

Laboratorio 370/344

Teléfono: +34 91 837 31 12 ext 4427/4263



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

Correo electrónico: jlgarcia@cib.csic.es
Página web: <https://www.cib.csic.es/es/miembros/jose-luis-garcia-lopez-0>
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Dr. Héctor Tejero Franco
Unidad de Bioinformática
Centro Nacional de Investigaciones Oncológicas (CNIO)
Teléfono: +34 91 732 80 00 ext 3062
Correo electrónico: htejero@cniio.es
Página web: <http://www.cniio.es>
Twitter: @BU_CNIO
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Javier Perales Patón
Unidad de Bioinformática
Centro Nacional de Investigaciones Oncológicas (CNIO)
Teléfono: +34 91 732 80 00 ext 3062
Correo electrónico: jperales@cniio.es
Página web: <http://www.cniio.es>
Twitter: @BU_CNIO
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Elena Piñeiro Yáñez
Unidad de Bioinformática
Centro Nacional de Investigaciones Oncológicas (CNIO)
Teléfono: +34 91 732 80 00 ext 3062
Correo electrónico: epineiro@cniio.es
Página web: <http://www.cniio.es>
Twitter: @BU_CNIO
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Dr. Joaquín Dopazo
Fundación Progreso y Salud
Hospital Virgen del Rocío, Sevilla.
Correo electrónico: joaquin.dopazo@juntadeandalucia.es
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Angela del Pozo Mate
INGEMM - Instituto de Genética Médica y Molecular
Instituto de Investigación IdIPAZ
Hospital Universitario La Paz
Correo electrónico: ingemm.adelpozo@gmail.com
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

Dr. Javier Suela
Director Técnico NIMGenetics
Correo electrónico: jsuela@nimgenetics.com



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

Página web: <https://www.nimgenetics.com/profesional/profesional-8/>
Horario de tutorías: Petición de cita previa en clase o por correo electrónico.

1.11. Objetivos del curso

A lo largo de esta asignatura, el estudiante se familiarizará con las diferentes estrategias y aproximaciones tecnológicas existentes en la actualidad para la generación de datos multi-ómicos a gran escala en biología y biomedicina. Además, adquirirá conocimientos avanzados para el análisis de los datos generados por dichas tecnologías y, muy especialmente, será capaz de adquirir, procesar e interpretar datos obtenidos desde técnicas de secuenciación masiva (NGS).

Por otro lado, esta asignatura 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 estudiantes recibirán lecciones teórico-prácticas sobre la detección, análisis e interpretación de mutaciones y variantes asociadas al número de copias génicas (CNVs) obtenidos por técnicas de secuenciación masiva. Se abordará además el análisis de polimorfismos (SNPs) a través de estudios de asociación (GWAS) y 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.

OBJETIVOS GENERALES

- Adquirir conocimientos teóricos de los bloques temáticos de la asignatura.
- Conocer las estructuras de datos y tipos de formatos empleados en los estudios de NGS.
- Aplicar técnicas de análisis de datos de NGS en experimentos con datos reales.
- Aplicar métodos computacionales para la resolución de problemas científicos reales.
- Disponer de un catálogo de recursos bioinformáticos y herramientas públicas útiles para el análisis de datos de NGS.
- Estar familiarizado con las principales bases de datos relacionados con los temas cubiertos en la asignatura.

Las **competencias específicas (CE)** asociadas al título de Máster en Bioinformática y Biología Computacional que se adquieren en esta asignatura son:

CE2: Capacidad de aplicar métodos computacionales a la resolución de problemas en el ámbito de la investigación biomédica básica y traslacional.



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

CE7: Capacidad de analizar computacionalmente los datos generados por las tecnologías "ómicas" de alto rendimiento en biología y biomedicina.

CE9: Capacidad de explotación de la información de las bases de datos biológicas sobre secuencias, estructuras, transcriptomas, genomas, proteomas, etc.

Las **competencias generales (CG) básicas (CB) y transversales (CT)** del Máster en Bioinformática y Biología Computacional que se adquieren en esta asignatura son:

CG1: Capacidad para comprender y aplicar métodos y técnicas de investigación en el ámbito de la Bioinformática.

CG2: Capacidad para proyectar, calcular y diseñar productos bioinformáticos.

CG3: Capacidad para trabajar en equipos multidisciplinares, comunicándose eficientemente y desarrollando su actividad de acuerdo con las buenas prácticas científicas.

CG4: Capacidad para la investigación, desarrollo e innovación, en empresas y centros tecnológicos, en el ámbito de la Bioinformática.

CG5: Capacidad para la aplicación de los conocimientos adquiridos y resolución de problemas en entornos nuevos o poco conocidos en el ámbito de la Bioinformática.

CG6: Capacidad de búsqueda, análisis y gestión de información; incluyendo la capacidad de interpretación y evaluación con un razonamiento crítico y autocrítico.

CG7: Capacidad de estudiar y resolver problemas biológicos y biomédicos con el soporte de herramientas computacionales.

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

CB7: Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB8: Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB9: Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

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.

CT1: Capacidad para trabajar en equipo de forma colaborativa y con responsabilidad compartida en el diseño y comunicación de estrategias experimentales.

CT2: Capacidad de identificar fuentes de información científica solventes para fundamentar el estado de la cuestión de un problema bioinformático y poder abordar su resolución.

1.12. Contenidos del programa*

Programa

Bloque 1. Introducción a la genómica

Bloque 2. Métodos de Secuenciación

Bloque 3. Alineadores y Formatos para datos de NGS

Bloque 4. Introducción a las variantes genómicas.

Bloque 5. Mutaciones.

Bloque 6. SNPs.

Bloque 7. Epidemiología Molecular.

Bloque 8. Detección de variantes estructurales (CNVs) en poblaciones humanas, bases de la variabilidad.

Bloque 9. Principales fuentes de datos y herramientas para el análisis e interpretación de resultados.

Programa detallado

1. **Bloque 1. Introducción a la genómica.** Introducción general al genómica. Estado del arte. Proyectos genómicos internacionales. Genómica y sociedad.
2. **Bloque 2. Métodos de Secuenciación.** Historia de la secuenciación y métodos clásicos. Métodos de secuenciación masiva, tecnologías actuales (454, SOLiD, Illumina, Ion, Nanopore, etc)
3. **Bloque 3. Alineadores y Formatos para datos de NGS.** Alineadores y tipos de algoritmos de alineamiento (hashing, alineadores Burrows-Wheeler). Archivos estándar de secuenciación masiva: FASTQ files. Archivos estándar de alineamientos: SAM/BAM files.
4. **Bloque 4. Introducción a las variantes genómicas.** Relevancia e impacto. Interpretación clínica. Asociación de variantes con respuesta a fármacos. Fuentes de interacción con fármacos. Aplicaciones traslacionales.



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

5. **Bloque 5. Mutaciones.** NGS y detección. Anotación. Interpretación y consecuencias clínicas. Consecuencias transcripcionales. Predicción de impacto y evidencias clínicas. Mutaciones en cáncer. Detección de artefactos. Herramientas de visualización y anotación de variantes. Priorización y selección.
6. **Bloque 6. SNPs.** Genes y fenotipo. SNPs y enfermedad. Métodos estadísticos. Haplotipo y desequilibrio de ligamiento. Tag-SNPs. Proyecto HapMap. Análisis de polimorfismos en poblaciones mediante (GWAS). Diseños de GWAS y control de calidad. Bases de datos y meta-análisis.
7. **Bloque 7. Epidemiología Molecular.** Epidemiología clásica. Epidemiología del cáncer. Modelos epidemiológicos. Datos ómicos y big data en epidemiología. Métodos multimarcador. Integromica. Anotación génica post-GWAS.
8. **Bloque 8. Detección de variantes estructurales (CNVs) en poblaciones humanas, bases de la variabilidad.** Introducción a las CNVs. CNVs y expresión génica. Métodos de detección de CNVs con aCGH. SNP y CNVs. Detección de CNVs con NGS. Limitaciones técnicas.
9. **Bloque 9. Principales fuentes de datos y herramientas para el análisis e interpretación de resultados.** ICGC portal. TCGA. cBioportal. LINCS, PanDrugs etc.

*La asignatura incluye además seminarios *ad hoc* impartidos por expertos en el área y que se impartirán dentro de la asignatura “Seminarios”.

1.13. Referencias de consulta

La mayor parte de la bibliografía serán artículos de la literatura científica actual que serán distribuidos o indicados a los alumnos antes o durante las clases. La mayoría de los temas expuestos se pueden estudiar en:

Libros

- Exploring Personal Genomics. Joel T. Dudley, Konrad J. Karczewski. Oxford. 2013.
- Genomic and Precision Medicine: Foundations, Translation, and Implementation. Geoffrey S. Ginsburg, Huntington F Willard. 2016

Artículos

- Hyman, D. M., Taylor, B. S. & Baselga, J. Implementing Genome-Driven Oncology. Cell 168, 584-599 (2017).
- Biankin, A. V. The road to precision oncology. Nat Genet 49, 320-321 (2017).
- Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17, 333-351 (2016).



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

- Alioto, T. S. et al. A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing. *Nat Commun* 6, 1-13 (2015).
- Mardis, E. R. DNA sequencing technologies: 2006-2016. *Nature Protocols* 12, 213-218 (2017).
- Baselga, J. Bringing precision medicine to the clinic: from genomic profiling to the power of clinical observation. *Annals of Oncology* 24, 1956-1957 (2013).
- Danecek, P. et al. The Variant Call Format and VCFtools. *Bioinformatics* (2011). doi:10.1093/bioinformatics/btr330
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760 (2009).

Material electrónico de trabajo: los documentos electrónicos de trabajo (material adicional, ejercicios del curso, etc.) se facilitarán a los alumnos durante las clases.

2. Métodos docentes

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

*Clases de teoría:

Actividad del profesor

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

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.

*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:



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

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. Supervisar el trabajo de los grupos de trabajo en el laboratorio. Suministrar el guión de prácticas a completar en el laboratorio.

Se utilizan el método expositivo tanto en tutorías como en el laboratorio con cada grupo de trabajo. Los medios utilizados son los entornos de programación, editores, compiladores y los ordenadores del propio laboratorio para la ejecución y análisis de los programas realizados.

Actividad del estudiante:

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.

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.

3. Tiempo de trabajo del estudiante

		Nº de horas	Porcentaje
Presencial	Clases teóricas	30 h (20 %)	54 h (36%)
	Clases prácticas	20 h (13%)	
	Tutorías, seminarios y realización de pruebas escritas	4 h (3%)	
No presencial	Estudio semanal	44 h (46%)	96 h (64%)
	Realización de actividades prácticas	14 h (15%)	
	Preparación del examen (convocatoria ordinaria)	16 h (16%)	
	Preparación del examen (convocatoria extraordinaria)	22 h (23%)	
Carga total de horas de trabajo: 25 horas x 6 ECTS		150 h	

4. Métodos de evaluación y porcentaje en la calificación final

- Todas las partes de la asignatura, teoría, prácticas puntúan sobre 10.
- La calificación final de la asignatura en la convocatoria ordinaria se obtiene a partir de las notas de teoría y participación en el aula mediante la ecuación:

$$\text{Calificación: } 0.9 \cdot \text{Teoría} + 0.1 \cdot \text{Participación en el aula}$$



Asignatura: Fundamentos de Secuenciación de Última Generación y Genómica Traslacional
Código: 33088
Centro: Escuela Politécnica Superior
Titulación: Máster de Bioinformática y Biología Computacional
Nivel: Máster
Tipo: Obligatoria

5. Cronograma*

Semana	Contenido	Horas presenciales	Horas no presenciales
1	Bloque 1. Introducción a la genómica	3	5
1	Bloque 2. Métodos de Secuenciación	3	5
1-2	Bloque 3. Alineadores y Formatos para datos de NGS	3	5
2	Bloque 4. Introducción a las variantes genómicas y sus aplicaciones traslacionales.	3	5
2-3	Bloque 6. Análisis de polimorfismos en poblaciones mediante SNPs (GWAS).	9	20
4	Bloque 7. Epidemiología Molecular.	6	10
4-5	Bloque 5. Mutaciones: NGS y detección. Anotación. Interpretación y consecuencias clínicas.	17	35
6	Bloque 8. Detección de variantes estructurales (CNVs) en poblaciones humanas, bases de la variabilidad.	3	5
6	Bloque 9. Principales fuentes de datos y herramientas para el análisis e interpretación de resultados.	2	5
6	Examen en convocatoria ordinaria	1	

*Este cronograma tiene carácter orientativo. Los seminarios específicos de esta asignatura no están incluidos en el cronograma ya que se imparten dentro del marco de la asignatura "Seminarios".