

**UNIVERSIDAD INTERAMERICANA DE PUERTO RICO  
RECINTO METROPOLITANO  
DEPARTAMENTO DE CIENCIAS NATURALES  
PROGRAMA DE MAESTRÍA EN CIENCIAS EN MICROBIOLOGÍA MOLECULAR**

**PRONTUARIO**

**I. INFORMACIÓN GENERAL**

Título del Curso	: <b>PROGRAMACIÓN Y MINERÍA DE DATOS</b>
Código y Número	: <b>MOMI 5320</b>
Créditos	: 3
Término Académico	:
Profesor	:
Horas de Oficina	:
Teléfono de la Oficina	:
Correo Electrónico	:

**II. DESCRIPCIÓN**

Aplicación de plataformas (*pipelines*) de código abierto para el análisis de las secuencias biológicas. Requiere 45 horas de conferencia-laboratorio.

**III. OBJETIVOS**

Se espera que al finalizar el curso, el estudiante pueda:

1. Aplicar lenguajes de programación y plataformas computacionales de procesamiento paralelo de alta performance para análisis de datos biológicos.
2. Formular un enfoque ético al manejo y transformación de los datos científicos y de la propiedad intelectual.

**Competencias del perfil del egresado que se atienden en este curso:**

1. Evaluar las plataformas más útiles para analizar datos de microbiología molecular utilizando herramientas emergentes.
2. Valorar la importancia de los estándares éticos relacionados con la conducta científica en la investigación, el respeto por la confidencialidad y la defensa de la propiedad intelectual.

#### IV. CONTENIDO TEMÁTICO

- A. Introducción a la programación: el ambiente UNIX
- B. Tutorial de Unix
  - 1. Archivos tabulares
  - 2. Combinación de comandos
  - 3. Modificar ficheros de texto en el terminal (Interfaces de línea de comandos)
  - 4. Crear conexión remota
  - 5. Copiar ficheros entre maquinas
  - 6. Cambiar permiso de ficheros
  - 7. Editor de texto *Emacs*
- C. Introducción a *Python* y *IPython* (funciones interactivas)
- D. Estadística inferencial y figuras
- E. Diseño de *Primers* en *Primer Prospector* y usando herramientas del web
- F. Aplicaciones de Perl en la bioinformática
  - 1. Manejo de datos de micro arreglos
  - 2. Detección de genes
- G. Creación de pipelines de trabajo
- H. Redistribución (*parsing*) de secuencias
- I. Iteraciones en *Python*
- J. Ensamblaje de secuencias de DNA
- K. *Clustering* jerárquico y algoritmos de agrupamiento
- L. Herramientas en *Galaxy* para flujos de trabajo con secuencias de DNA y proteínas

#### V. ACTIVIDADES

- 1. Conferencias ilustradas en formato de *power point*
- 2. Conferenciantes invitados
- 3. Laboratorio de computadoras
- 4. Ejercicios y simulaciones en la computadora

#### VI. EVALUACIÓN

La evaluación del curso estará basada en la ejecución de 1 proyecto, 2 exámenes parciales y un examen final. La nota final se calculará a base de 100% de la siguiente forma:

Un proyecto y 2 exámenes parciales.....	75 %
(25% cada uno)	
Examen final.....	25 %
	-----
	100%

### **Asistencia a Clases y Exámenes**

La asistencia a clases es obligatoria. El estudiante que necesite ausentarse a una clase deberá contactar al profesor antes de la clase por teléfono o por correo electrónico. No habrá exámenes de reposición, excepto por motivos de enfermedad. En dicho caso, las reposiciones se ofrecerán con una debida excusa médica durante el período de exámenes finales en las horas de oficina del profesor.

## **VII. NOTAS ESPECIALES**

### **A. Servicios auxiliares o necesidades especiales**

Todo estudiante que requiera servicios auxiliares o asistencia especial deberá solicitar los mismos al inicio del curso o tan pronto como adquiera conocimiento de que los necesita, a través del registro correspondiente, en el Programa de Orientación con el Sr. José Rodríguez.

### **B. Honradez, fraude y plagio**

La falta de honradez, el fraude, el plagio y cualquier otro comportamiento inadecuado con relación a la labor académica constituyen infracciones mayores sancionadas por el Reglamento General de Estudiantes. Las infracciones mayores, según dispone el Reglamento General de Estudiantes, pueden tener como consecuencia la suspensión de la Universidad por un tiempo definido mayor de un año o la expulsión permanente de la Universidad, entre otras sanciones.

### **C. Uso de dispositivos electrónicos**

Se desactivarán los teléfonos celulares y cualquier otro dispositivo electrónico que pudiese interrumpir los procesos de enseñanza y aprendizaje o alterar el ambiente conducente a la excelencia académica. Las situaciones apremiantes serán atendidas, según corresponda. Se prohíbe el manejo de dispositivos electrónicos que permitan acceder, almacenar o enviar datos durante evaluaciones o exámenes.

### **D. Cumplimiento con las disposiciones del Título IX**

La Ley de Educación Superior Federal, según enmendada, prohíbe el discrimen por razón de sexo en cualquier actividad académica, educativa, extracurricular, atlética o en cualquier otro programa o empleo, auspiciado o controlado por una institución de educación superior independientemente de que esta se realice dentro o fuera de los predios de la institución, si la institución recibe fondos federales.

Conforme dispone la reglamentación federal vigente, en nuestra unidad académica se ha designado un(a) Coordinador(a) Auxiliar de Título IX que brindará asistencia y orientación con relación a cualquier alegado incidente constitutivo de discrimen por sexo o género, acoso sexual o agresión sexual. Se puede comunicar con el Sr. George Rivera, Coordinador Auxiliar al teléfono 787-250-1912 extensión 2262 o 2147 o al correo electrónico [griverar@metro.inter.edu](mailto:griverar@metro.inter.edu).

El Documento Normativo titulado Normas y Procedimientos para Atender Alegadas Violaciones a las Disposiciones del Título IX es el documento que contiene las reglas institucionales para canalizar cualquier querrela que se presente basada en este tipo de

alegación. Este documento está disponible en el portal de la Universidad Interamericana de Puerto Rico ([www.inter.edu](http://www.inter.edu)).

## VIII. RECURSOS EDUCATIVOS

### Lecturas Suplementarias

Adler P, Kolde R, Kull M, Tkachenko A, Peterson H, Reimand J, Vilo J. Mining for coexpression across hundreds of datasets using novel rank aggregation and visualization methods. *Genome Biology* 2009;

Barret T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muerter RN, Edgar R. NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Research* 2009; 37:D885-D890.

## IX. BIBLIOGRAFÍA

### Libros

Brian Kernighan, et al (1999), *The Practice of Programming*, W. Addison-Wesley, ISBN-13: 978-0201615869

Mark Lutz (2006) *Programming python 3<sup>rd</sup> edition*, O'Reilly Media, Inc. ISBN13: **978-0596009250**

**Revisado**  
**FGV Noviembre 2013**