PEDECIBA

INSTITUTO DE HIGIENE
Prof. Arnoldo Berta
FACULTAD DE MEDICINA

Febrero - Marzo 2024
Introducción a la línea de comandos y a la programación para análisis bioinformáticos.

Coordinador: Dr. Andrés Iriarte

Docentes participantes: Dr. Andrés Iriarte, Dr. Eugenio Jara, Dr. Fernando Álvarez-Valín, Dr. Germán Traglia, Dr. Héctor Musto & Dra. Paula Vico.
Colaboradores: Lic. Hernán Juan.

Contenido:

Como resultado de los avances en la tecnología de secuenciación se ha generado una revolución en diversas áreas de la biología. Estás metodologías generan una enorme cantidad de información, genomas completos, genes e incluso permite estimar con precisión sus niveles de expresión. Por sus características estos datos sólo pueden ser analizados mediante herramientas bioinformáticas. Muchas de estas herramientas se desarrollan sin una interfaz gráfica, y las que la tienen suelen desarrollarse más lentamente o en versiones no actualizadas. Adquirir manejo en su entorno es fundamental para lograr un uso eficiente de las mismas. Este curso plantea introducir a los estudiantes en la línea de comando, elemento básico para el análisis de secuencias, genomas y datos secuenciación, y en a la programación. El curso está orientado a estudiantes avanzados de grado y estudiantes de posgrado de áreas biológicas sin formación en programación o bioinformática.

Palabras claves: Linux, Biología computacional, Lenguaje de programación R.

Programa:

Teórico (1 hr. 30 min.):

1. Introducción al Curso. Historia de la Bioinformática. Introducción a sistema operativo Linux. Formatos, scripts y lenguajes de programación. (Iriarte/Musto – Lunes 17/02)
2. Comandos básicos en la terminal. Concepto de pipeline y *path*. (Iriarte – Martes 18/02)
3. Comandos para el manejo de textos I. (Iriarte – Miércoles 19/02)
4. Comandos para el manejo de textos II. (Iriarte – Jueves 20/02)
5. Herramientas del paquete *Emboss*: Introducción al *emboss* y manejo básico de secuencias. (Iriarte - Viernes 21/02)
6. Introducción a la programación. (Iriarte - Lunes 24/02)
7. Programación en el *Shell* I ("*Bash scripting*"): Manejo de variables y listas, salidas y entradas. Integración con otros lenguajes de programación. (Iriarte - Martes 25/02)
8. Programación en el *Shell* II ("*Bash scripting*"): Bucles y Condicional. (Iriarte – Miércoles 26/02)
9. Programación en el *Shell* III ("*Awk y Sed*"). (Alvarez-Valín – Jueves 27/02)
10. Introducción a CONDA, sistema de gestión de paquetes y ambientes. (Iriarte - Viernes 28/02)
11. Introducción a R. (Traglia - Lunes 10/03)
12. Manejo de datos en R: entradas y salidas. Tipos de objetos. (Traglia - Martes 11/03)
13. Análisis de datos de secuenciación masiva en R: Análisis metagenómico (Jara - Miércoles 12/03)

14. Análisis de datos de secuenciación masiva en R: RNA-seq (Jara - Jueves 13/03)
15. Análisis de enriquecimiento funcional en R. (Iriarte - Viernes 14/03)

***Práctico (3 hr.)*:**
Docentes: Dr. A. Iriarte, Dr. F. Alvarez-Valin, Dr. G. Traglia, Dr. P. Vico, Lic. H. Juan, Dr. Eugenio Jara.

1. Conexión remota vía *ssh*. Desplazamiento a través de la terminal. Uso de comandos básicos en la terminal. (Lunes 17/02)
2. Instalar y correr programas. Ejemplos: *Muscle*, *FastTree*, *Bwe*, *Spades* y otros. (Martes 18/02)
3. Manejo de textos: sed, uniq, grep, sort y otros. (Miércoles 19/02)
4. Manejo de textos: *awk*. (Jueves 20/02)
5. Comandos del paquete *Emboss*, ejemplos más utilizados. (Viernes 21/02)
6. Comparación de scripts escritos en distintos lenguajes y ejemplos. (Lunes 24/02)
7. *Scripting Shell* I. Introducción al *scripting* y uso de variables. (Martes 25/02)
8. *Scripting Shell* II. Recorrer listas, ejecutar acciones repetidas. Uso de *loops*. (Miércoles 26/02)
9. *Scripting Awk & Sed*. Introducción a la programación con *awk*. (Jueves 27/02)
10. Manejo de CONDA, ejemplos. (Viernes 28/02)
11. *RStudio*. Manejo de datos: filas y matrices. (Lunes 10/03)
12. Estadística básica y gráficos simples. (Martes 11/03)
13. Análisis metagenómico. (Miércoles 12/03)
14. Análisis de *RNA-seq*. (Jueves 13/03)
15. Análisis genómicos funcionales. (Viernes 14/03)

Evaluación: Examen individual, con preguntas abiertas y ejercicios, 3 horas.

Texto recomendado:

- *Practical Computing for Biologists*, de Haddock and Dunn, Ed. Sinauer, 2010.
- *A Beginner's Guide to R*, de Zuur, Ieno and Meesters, Ed. Springer, 2009.
- *R para Ciencia de Datos, de Garrett Grolemund y Hadley Wickham,* 2017.

Fecha: del 17 de febrero al 14 de marzo.
Presencial: Salón Rojo, Instituto de Higiene.
Horario: 12:30 – 17:00 hs.
Carga Horaria: Teóricos 22 hs. 30 min. // Prácticos 45 hs. // Evaluación 3 hs.
Créditos: 9 (PEDECIBA).
Cupos: 25.

Escribir a: airiarteo@gmail.com
Web: www.higiene.edu.uy/ddbp/lbc/