



Julio - Agosto 2023
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. Guillermo Lamolle & Dr. Héctor Musto.

Colaboradores: Lic. Hernán Juan, Mag. Javier Calvelo.

Contenido:

Como resultado de los avances en la tecnología de secuenciación se ha generado una revolución en diversas áreas de la biología. Estas 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. (Musto & Iriarte – Lunes 24/07)
2. Comandos básicos en la terminal. Concepto de pipeline y *path*. (Iriarte – Martes 25/07)
3. Comandos para el manejo de textos I. (Iriarte – Miércoles 26/07)
4. Comandos para el manejo de textos II. (Iriarte – Jueves 27/07)
5. Herramientas del paquete *Emboss*: Introducción al *emboss* y manejo básico de secuencias. (Iriarte - Viernes 28/07)
6. Introducción a la programación. (Iriarte - Lunes 31/07)
7. Programación en el *Shell I* (“*Bash scripting*”): Manejo de variables y listas, salidas y entradas. Integración con otros lenguajes de programación. (Iriarte - Martes 01/08)
8. Programación en el *Shell II* (“*Bash scripting*”): *Loops* (for, while, until) y Condicional (if). (Iriarte – Miércoles 02/08)
9. Programación en el *Shell III* (“*Awk y Sed*”). (Alvarez-Valín – Jueves 03/08)
10. Introducción a R: Manejo de datos, entradas y salidas. Tipos de objetos. (Lamolle - Viernes 04/08)
11. Estadística básica y gráficos simples en R. (Lamolle - Lunes 07/08)





12. Manejo básico de datos en R. (Iriarte - Martes 08/08)
13. Análisis de datos de secuenciación masiva en R. (Jara - Miércoles 09/08)
14. Análisis de enriquecimiento funcional en R. (Iriarte - Jueves 10/08)

Práctico (3 hr.):

Docentes: Dr. A. Iriarte, Dr. F. Alvarez-Valin, Mag. J. Calvelo, Dr. G. Lamolle, 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 24/07)
2. Instalar y correr programas. Ejemplos: *Muscle*, *FastTree*, *Bwe*, *Spades* y otros. (Martes 25/07)
3. Manejo de textos: *sed*, *uniq*, *grep*, *sort* y otros. (Miércoles 26/07)
4. Manejo de textos: *awk*. (Jueves 27/07)
5. Comandos del paquete *Emboss*, ejemplos más utilizados. (Viernes 28/07)
6. Comparación de scripts escritos en distintos lenguajes y ejemplos. (Lunes 31/07)
7. *Scripting Shell* I. Introducción al scripting y uso de variables. (Martes 01/08)
8. *Scripting Shell* II. Recorrer listas, ejecutar acciones repetidas. Uso de *loops*. (Miércoles 02/08)
9. *Scripting Awk & Sed*. Introducción a la programación con *awk*. (Jueves 03/08)
10. *RStudio*. Manejo de datos, filas y matrices en R. (Viernes 04/08)
11. Análisis estadísticos en R. (Lunes 07/08)
12. Manejo básico de datos con *Tidverse*. (Martes 08/08)
13. Análisis de *RNAseq* en R. (Miércoles 09/08)
14. Análisis genómicos funcionales. (Jueves 10/08)

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 Golemund y Hadley Wickham, 2017.

Fecha: del 24 de julio al 10 de agosto.

Presencial: Salón Rojo, Instituto de Higiene.

Horario: 9:00 – 13:30 hs.

Carga Horaria: Teóricos 21 hs. // Prácticos 42 hs. // Evaluación 3 hs.

Créditos: 8 (PEDECIBA).

Cupos: 20.

Escribir a: airiarteo@gmail.com

Web: www.higiene.edu.uy/ddbp/lbc/

