



AGOSTO 2021

Introducción a la línea de comandos y a la programación para análisis bioinformáticos.

Coordinador: Dr. Andrés Iriarte

Docentes participantes: Dr. Guillermo Lamolle, Dr. Fernando Alvarez-Valin, Dr. Héctor Romero & Dr. Héctor Musto

Colaboradores: Mag. Eugenio Jara, Mag. Javier Calvelo, Lic. Mauricio Langleib, Lic. Virginia Cantera.

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, Programación.

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 26/07)
2. Comandos básicos en la terminal. Concepto de pipeline y *path*. (Iriarte – Martes 27/07)
3. Comandos para el manejo de textos I. (Iriarte – Miércoles 28/07)
4. Comandos para el manejo de textos II. (Iriarte – Jueves 29/07)
5. Herramientas del paquete *Emboss*: Introducción al *emboss* y manejo básico de secuencias. (Iriarte - Viernes 30/07)
6. Introducción a la programación. (Iriarte - Lunes 02/08)
7. Programación en el *Shell I* (“*Bash scripting*”): Manejo de variables y listas, salidas y entradas. Integración con otros lenguajes de programación. (Iriarte - Martes 03/08)
8. Programación en el *Shell II* (“*Bash scripting*”): *Loops* (for, while, until) y Condicional (if). (Iriarte – Miércoles 04/08)
9. Programación en el *Shell III* (“*Awk y Sed*”). (Alvarez-Valín – Jueves 05/08)
10. Introducción a R: Manejo de datos, entradas y salidas. Tipos de objetos. (Lamolle – Viernes 06/08)





11. Estadística básicas y gráficos simples en R. (Lamolle - Lunes 09/08)
12. Manejo de datos en R I. (Romero & Langleib - Martes 10/08)
13. Manejo de datos en R II. (Romero & Langleib - Miércoles 11/08)
14. Análisis de datos de secuenciación masiva en R. (Jara - Jueves 12/08)
15. Análisis de enriquecimiento funcional en R. (Iriarte - Viernes 13/08)

Práctico (2 hr. 45 min.):

Docentes: Dr. A. Iriarte, Mag. E. Jara, Dr. F. Alvarez-Valin, Dr. G. Lamolle, Dr. H. Romero, Mag. J. Calvelo, Lic. M. Langleib & Lic. V. Cantera.

1. Conexión remota vía *ssh*. Desplazamiento a través de la terminal. Uso de comandos básicos en la terminal. (Lunes 26/07)
2. Instalar y correr programas. Ejemplos: *Muscle*, *FastTree*, *Bwe*, *Spades* y otros. (Martes 27/07)
3. Manejo de textos: *sed*, *uniq*, *grep*, *sort* y otros. (Miércoles 28/07)
4. Manejo de textos: *awk*. (Jueves 29/07)
5. Comandos del paquete *Emboss*, ejemplos más utilizados. (Viernes 30/07)
6. Comparación de scripts escritos en distintos lenguajes y ejemplos. (Lunes 02/08)
7. *Scripting Shell I*. Introducción al scripting y uso de variables. (Martes 03/08)
8. *Scripting Shell II*. Recorrer listas, ejecutar acciones repetidas. Uso de *loops*. (Miércoles 04/08)
9. *Scripting Awk & Sed*. Introducción a la programación con *awk*. (Jueves 05/08)
10. *RStudio*. Manejo de datos, filas y matrices en R. Tipos de objetos. (Viernes 06/08)
11. Análisis estadísticos en R. (Lunes 09/08)
12. Ejemplos simples de *loops* y funciones en R. (Martes 10/08)
13. Filtrado, edición y visualización de datos biológicos. (Miércoles 11/08)
14. Análisis de *RNAseq* en R. (Jueves 12/08)
15. Análisis genómicos funcionales. (Viernes 13/08)

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 26 de Julio al 13 de Agosto.

Virtual: Sala Zoom.

Horario: 9:00 – 13:30 hs.

Carga Horaria: Teóricos 23 hs. // Prácticos 41 hs. // Evaluación 3 hs.

Créditos: 9 (PEDECIBA).

Cupos: 15.

Escribir a: airiarteo@gmail.com

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

