



CURSO POSTGRADO: GENÓMICA FÚNGICA

Coordinador: Dra. Sofia Chulze (UNRC)

Profesor: Dr. Theo van der Lee (Wageningen UR, Países Bajos)

CV en <https://www.vcard.wur.nl/Views/Profile/View.aspx?id=1772&ln=eng>

Colaborador: Dra. Maria Veronica Fumero (UNRC)

El curso se dictará en inglés. Se trabajará en grupos pequeños o de manera individual con datos proporcionados por los docentes o presentes en bases de datos. Se dictarán clases teóricas y demostraciones, para luego enfocarse en la experiencia práctica. Posibilidad de BYOD.

- Objetivo general del curso:

Aprendizaje de herramientas bioinformáticas para el manejo de grandes volúmenes de datos en investigaciones biológicas basadas en secuenciación de nueva generación (next-generation sequencing, NGS). Se introducirán habilidades más comunes de computación como los comandos Unix y la programación en R, que luego se aplicarán a lo largo del curso. Se realizará el ensamblaje de genomas y análisis de su calidad, análisis de datos RNA-seq, obtención de variantes genómicas (SNPs, INDELS), genómica comparativa y evolución de genomas. La formulación de preguntas de investigación basadas en datos genómicos será otro de los objetivos del curso.

- El curso está principalmente orientado a micólogos, fitopatólogos, biólogos, taxónomos, bioinformáticos y afines.

CRONOGRAMA

- Fecha: 22 al 26 de Septiembre de 2017

- Lugar: Luján de Cuyo, Mendoza - Argentina (Laboratorio de Informática, Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo; Almt. Brown 500, Chacras de Coria)

- Clases teóricas por la mañana (9-12), clases prácticas por la tarde (13-17)

<http://amcspegazzini.weebly.com>

amcspegazzini@gmail.com

<https://www.facebook.com/amcspegazzini>

CONTENIDOS

Viernes 22/9 - Introducción a NGS y bioinformática. Ensamblaje de genomas

Teórico: Introducción a las tecnologías de NGS y herramientas bioinformáticas básicas para el análisis de datos genómicos.

Práctico 1: Análisis y procesamiento de secuencias de acuerdo a su calidad (*quality trimming*). Por qué realizar el *quality trimming*. Ensamblaje *de novo* y basado en un genoma de referencia. Diferencias entre cada estrategia y selección de la mejor de acuerdo a las características del tipo de datos. Validación del ensamblado. Control de calidad y optimización. Exportación y almacenamiento de datos. Tipos de archivos generados (*output files*).

Sábado 23/9 - Anotaciones genómicas y RNA-seq

Revisión del práctico 1.

Teórico: Anotaciones genómicas (*genome annotation*). Introducción a las herramientas de identificación y descripción de genes. RNA-seq y evaluación de datos de expresión génica a gran escala.

Práctico 2: RNA-seq y herramientas para análisis de datos de expresión génica a gran escala. Estrategias basadas en *BRAKER1 (BRAKER1 pipelines)*. Anotaciones estructurales (*evidence-based structural annotation*). Validación de la calidad mediante el score BUSCO. Anotaciones funcionales usando INTERPROSCAN. Análisis de expresión génica utilizando datos de RNA-seq. Edición del RNA.

Lunes 25/9 - Genómica comparativa. Señales genómicas de la evolución en hongos

Revisión del práctico 2.

Teórico: Genómica comparativa. Evolución de los genomas fúngicos. Secuencias repetidas y transposones en hongos. Efecto de posición. Diferencias entre secuencias genómicas y mitocondriales.

Práctico: Herramientas para el análisis comparativo de genomas fúngicos. Genotipado mediante SNPs, INDELS y variaciones estructurales. Búsqueda de similitudes y diferencias entre especies y cepas fúngicas. Análisis de regiones taxonómicamente relevantes (*barcoding regions*, proyecto BOLD).

<http://amcspegazzini.weebly.com>

amcspegazzini@gmail.com

<https://www.facebook.com/amcspegazzini>

Martes 26/9 - Metagenómica

Revisión del práctico 3.

Teórico: Estrategias para la identificación de nuevos patógenos en tejidos vegetales y matrices complejas utilizando datos de NGS. Desafíos de la Metagenómica. Estrategias para la generación de datos para la identificación (*host-reduction*).

Práctico: Implementación de estrategias y parámetros en CLC Genomics Workbench (QIAGEN). Optimización. Control de calidad. Exportación y almacenamiento de datos.

- Programas:

CLC-bio, Geneious, Galaxy, DNASTar, Homebrew.

PREPARACIÓN PARA EL CURSO

Antes del primer día del curso, los alumnos necesitarán poder trabajar en una terminal, en un entorno UNIX y editar archivos en el mismo entorno con un editor de texto.

Acá se demuestra cómo y qué deben hacer en varios sistemas operativos (OS, Windows, Linux).

a. Mac

Mac tiene UNIX incorporado, de manera que quien tenga Mac estará listo para el curso.

Mac tiene incorporado Python pero podrá variar la versión del mismo. El curso requiere Python 2.7 en adelante. Dependiendo de la versión de Python que posean, necesitarán instalarlo o actualizarlo. Se puede descargar desde python.org.

Hay varios editores de texto para Mac: nosotros sugerimos TextWrangler.

Los alumnos deberán conocer cómo conectarse a un servidor remoto con -ssh.

b. Windows

Windows es el sistema operativo más débil para interactuar con UNIX. No obstante, existen herramientas de libre descarga como PuTTY, MobaXterm, u otras, que se pueden descargar.

(Se necesitan el instalador y algunos complementos adicionales como NEdit, Python3, se les indicará como hacer esto).

Otra opción es utilizar Linux, mediante la instalación de Ubuntu (virtualbox).

<http://amcspegazzini.weebly.com>
amcspegazzini@gmail.com
<https://www.facebook.com/amcspegazzini>

Un buen editor de texto disponible para Windows es Notepad++.

c. Linux

Quienes estén familiarizados con Linux, están preparados y listos para el curso. Los editores de texto que pueden usar con Linux son: Gedit, jedit, nedit, emacs, vim, etc.

Se recomienda realizar la preparación con anterioridad al comienzo del curso. Para ello, los inscriptos deben dirigirse a mariaveronicafumero@gmail.com solicitando el instructivo en función del sistema operativo del que dispongan.

<http://amcspegazzini.weebly.com>
amcspegazzini@gmail.com
<https://www.facebook.com/amcspegazzini>