



LA ACADEMIA DE BIOLOGÍA EXPERIMENTAL Y APLICADA, EL POSGRADO EN BIOCENCIAS, EL CUERPO ACADÉMICO DE BIOLOGÍA Y ECOLOGÍA DE LA CONSERVACIÓN Y EL DICTUS, INVITAN AL CURSO:

INTRODUCCIÓN A LA BIOINFORMÁTICA PARA EL ANÁLISIS DE DATOS DE SECUENCIACIÓN GENÓMICA

INSTRUCTORES: DR. ALEXSANDRE GUTIÉRREZ BARRAGÁN Y DR. FERNANDO G. RAZO MENDÍVIL

CONTACTOS: alexandre.gutierrez@unison.mx ; fernando.razo@unison.mx

FECHA: 02 DE DICIEMBRE AL 13 DE DICIEMBRE DE 2024, DE 10:00 A 12:00 HRS

DURACIÓN: 20 HORAS

LUGAR: SALA DE CÓMPUTO, EDIFICIO 7H, UNIVERSIDAD DE SONORA

CUPO MÁXIMO: 20 ESTUDIANTES

OBJETIVO DEL CURSO

Capacitar a investigadores y estudiantes en el análisis de datos obtenidos mediante secuenciación genómica en Illumina para la obtención de sitios variantes (SNPs por sus siglas en inglés)

OBJETIVOS ESPECÍFICOS

1. Comprender el proceso de secuenciación de lecturas cortas en plataformas de secuenciación de segunda generación (Next Generation Sequencing), como Illumina
2. Manipular directorios y archivos y extraer información mediante comandos básicos en Linux
3. Conocer herramientas computacionales empleadas comúnmente para el procesamiento de datos de secuenciación genómica y la obtención de sitios variantes, como polimorfismos de un solo nucleótido (SNPs)

METODOLOGÍA Y TEMÁTICA

Los profesores expondrán una introducción general sobre cada tema. La presentación deberá fomentar la discusión mediante el intercambio de ideas y preguntas, para asimilar y razonar el fundamento del análisis a realizar. Posteriormente, los instructores mostrarán las herramientas que permiten el procesamiento de los datos en servidores de la Universidad de Sonora.

El curso comenzará con un repaso histórico de la secuenciación hasta llegar a la secuenciación de segunda generación (Next Generation Sequencing), para que los asistentes comprendan el proceso de secuenciación. Posteriormente, se dará una introducción a los comandos básicos en Linux para proporcionar herramientas de manipulación de archivos y extracción de información. Empleando este conocimiento básico, se procederá con la teoría y práctica para el análisis bioinformático de datos de secuenciación, iniciando por análisis y control de calidad, ensamble de genomas, alineamiento de lecturas y control de calidad de mapeo. Finalmente, se mostrarán algunas opciones para la obtención de variantes (SNPs), tanto en técnicas de representación reducida, como en datos obtenidos por secuenciación de genomas completos a baja profundidad.



PRE-REQUISITOS

Conocimiento básico teórico y práctico del análisis de datos genéticos.

OPCIONAL: Computadora personal con al menos 8GB de memoria RAM.

TEMARIO

1. Secuenciación genómica
 - a. Historia de la secuenciación
 - b. Secuenciación NGS/Illumina
2. Introducción a la bioinformática
 - a. Lenguaje Linux
 - b. Comandos básicos
3. Calidad de la secuenciación
 - a. Análisis de calidad con fastqc
 - b. Control de calidad con trimmomatic y/o fastp
4. Ensamble y anotación de genomas
5. Alineamiento/mapeo de lecturas con bwa
6. Control de calidad de alineamiento
 - a. Filtrado de lecturas mapeadas con samtools
 - b. Marcado y eliminación de duplicados con MarkDuplicates
 - c. Eliminación de regiones superpuestas (overlap)
 - d. Realineamiento alrededor de InDels con GATK
 - e. Visualización de alineamiento en IGV
7. Obtención de variantes
 - a. Llamado de SNPs para datos de GBS y/o RADseq
 - b. Inferencia de genotipos en ANGSD para datos de lcWGS
 - i. Parámetros de filtrado
 - ii. Estimación de la verosimilitud de los genotipos
 - iii. Llamado de genotipos

CRONOGRAMA DE ACTIVIDADES

DÍA	FECHA	TEMAS	INSTRUCTOR	HORAS
1	02/12/2024	Secuenciación genómica	ALEXSANDRE GUTIERREZ	2
2	03/12/2024	Introducción a la bioinformática	FERNANDO RAZO	2
3	04/12/2024	Calidad de la secuenciación	ALEXSANDRE GUTIERREZ	2
4	05/12/2024	Ensamble y anotación de genomas	FERNANDO RAZO	2
5	06/12/2024	Ensamble y anotación de genomas	FERNANDO RAZO	2
6	09/12/2024	Alineamiento de secuencias	ALEXSANDRE GUTIERREZ	2
7	10/12/2024	Control de calidad del alineamiento	ALEXSANDRE GUTIERREZ	2
8	11/12/2024	Obtención de variantes (GBS)	FERNANDO RAZO	2
9	12/12/2024	Obtención de variantes (lcWGS)	ALEXSANDRE GUTIERREZ	2
10	13/12/2024	Obtención de variantes (lcWGS)	ALEXSANDRE GUTIERREZ	2
TOTAL				20 hrs