

Curso -Taller: Genética de poblaciones aplicada a medicina

Perfil: El curso/taller esta dirigido a investigadores y estudiantes de pregrado y posgrado interesados en conocer (y practicar) las estrategias y métodos de análisis de datos de variantes genéticas de ADN (tipo SNP) útiles en estudios biomédicos, medicina forense y de antropología molecular, usando la genómica computacional como herramienta para la ejecución de análisis.

Los requisitos deseados de los interesados son: conocimientos básicos de línea de comandos Unix/Linux, Leguaje de programación Shell scripting, lenguaje R así como conocimientos básicos de biología molecular, medicina y tecnologías de nueva generación que producen datos de ADN (genotipos).

Profesor titular: Juan Carlos Fernández López

E-mail: jfernandez@inmegen.gob.mx

Duración: 11 Sesiones (33 horas)

Fecha: Del 25 de agosto al 1 de octubre de 2020

Horario: martes y jueves de 18:00 a 21:00 horas

OBJETIVO GENERAL DEL CURSO

Conocer y aplicar las estrategias de análisis de datos básicas de variantes genéticas (tipo SNP) en Genética de Poblaciones aplicada a estudios biomédicos, de genética de poblaciones en general.

OBJETIVOS ESPECÍFICOS

- Discutir la aplicación y el impacto que ha tenido la Genética de Poblaciones en la Medicina Genómica moderna en los últimos 15 años.
- Conocer y practicar los comandos básicos en UNIX/Linux más comúnmente usados en la manipulación de archivos de genotipos (ADN) de gran tamaño. (50% práctico).
- Conocer los formatos de archivos de variantes genéticas (ADN) más usados en el análisis de datos en Genética de Poblaciones Médica: Linkage Format y VCF.
- Conocer las pruebas de estadística básicas en el análisis de variantes aplicadas a medicina usando el lenguaje de programación R. (50% práctico)

- Aprender a usar la herramienta de análisis de asociación de genoma completo PLINK (v1.07 y v1.9). (50% práctico)
- Conocer, discutir y practicar los controles de calidad de los datos necesarios en los estudios genéticos y genómicos en medicina genómica. (50% práctico)
- Discutir la estratificación poblacional en los estudios biomédicos y la utilidad del Análisis de Componentes Principales (PCA). (50% práctico)
- Comprender y discutir la “ancestría global”, su utilidad en estudios de genómica humana y su interpretación. (50% práctico)
- Comprender y discutir la “ancestría local” y su utilidad en el estudio de enfermedades. (50% práctico)
- El alumno será capaz de desarrollar un proyecto en medicina genómica usando los métodos y estrategias aprendidas en el curso.

TEMARIO

Sesión 1.

Martes 25 de agosto

- **Presentación del Curso**

- **Introducción: Análisis de Datos en Genética de Poblaciones (en Medicina Genómica)**

Sesión 2.

Jueves 27 de agosto

Unix: Comandos usados en la manipulación de datos de alta densidad

Sesión 3.

Martes 1 de septiembre

Formatos de archivos de genotipos y herramientas básicas de análisis

Sesión 4.

Jueves 3 de septiembre

Introducción al manejo estadístico de datos en el lenguaje de programación R

Profesora invitada: Alicia Huerta (INCMNSZ)

Sesión 5.

Martes 8 de septiembre

Aprendiendo la herramienta de análisis de asociación de genoma completo PLINK (v1.07-v1.9)

Sesión 6.

Jueves 10 de septiembre

Control de Calidad de los Datos (QC) y su relevancia

Sesión 7.

Jueves 17 de septiembre

El efecto de la estratificación poblacional en los estudios de asociación y su corrección usando el Análisis de Componentes Principales (PCA)

Sesión 8.

Martes 22 de septiembre de 10:00 a 13:00 hrs.

Inferencia del componente ancestral (global), su utilidad y su interpretación en estudios de biomedicina y antropología molecular.

Profesora invitada: Sandra Romero (Inmegen)

Sesión 9.

Jueves 24 de septiembre de 10:00 a 13:00 hrs.

Ancestría local y sus aplicación en el estudio de enfermedades.

Profesora invitada: Sandra Romero (Inmegen)

Sesión 10.

Martes 29 de septiembre

Nuevas tecnologías genómicas de ADN.

Profesor invitado: Alfredo Hidalgo (Inmegen)

Sesión 11.

Jueves 1 de octubre

- La importancia de la paleogenómica en la Medicina Genómica

Profesor invitado: Federico Sánchez (Inmegen)

- Conclusiones del taller

BIBLIOGRAFÍA

Initial sequencing and analysis of the human genome
ES Lander, LM Linton, B Birren, C Nusbaum, MC Zody... - 2001

Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. I Silva-Zolezzi, A Hidalgo-Miranda, J Estrada-Gil... - Proceedings of the National Academy of Sciences, 2009

A map of human genome variation from population-scale sequencing
1000 Genomes Project Consortium - Nature, 2010

The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. A Moreno-Estrada, CR Gignoux, JC Fernández-López... - Science, 2014

UNIX/LINUX

<https://es.wikipedia.org/wiki/GNU/Linux>

Package: PLINK (v1.07, v1.9)

Author: Shaun Purcell

URL: <http://pngu.mgh.harvard.edu/purcell/plink/>

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007)

PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics, 81.

The R Project for Statistical Computing

<https://www.r-project.org/>

Reich D, Price AL, Patterson N (2008) Principal component analysis of genetic data. Nature Genetics 40, 491-2

Development of a panel of genome-wide ancestry informative markers to study admixture throughout the Americas. JM Galanter, JC Fernandez-Lopez, CR Gignoux... - PLoS genetics, 2012

Structure Software

Pritchard Lab, Stanford University

<https://web.stanford.edu/group/pritchardlab/structure.html>

D.H. Alexander, J. Novembre, and K. Lange. Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19:1655–1664, 2009.

Wang, H., Sofer, T., Zhang, X. et al. Local Ancestry Inference in Large Pedigrees. Sci Rep 10, 189 (2020). <https://doi.org/10.1038/s41598-019-57039-w>

Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico
SIGMA Type 2 Diabetes Consortium - Nature, 2014