

Visualización Científica

Proyecto nº2 de 2 (20% de la nota a presentación al examen)

Fecha límite para enviar el informe del proyecto: 15/06/18 (Viernes)

Para este proyecto, necesitara preparar un informe, respondiendo a las siguientes preguntas **poniendo los gráficos correspondientes adentro del informe**. Debe acompañar el envío del informe **con todo el código que ha escrito para obtener los gráficos**. El puntaje máximo que se puede obtener en la sesión practica es de 21 puntos, el que será convertido a nota 7. Con 1-3 días de atraso, la nota será limitada al 4. Con más atraso, la nota será 1.

Control de Calidad [7 puntos]

1. Explica en unas tres líneas o más cual es la función de las siguientes herramientas [1 punto]:
 - NGS QC toolkit (<http://www.nipgr.res.in/ngsqctoolkit.html>)
 - FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>)
2. Elige una métrica que NGS QC toolkit y FastQC tienen en común y como varia la representación de esas métricas entre ellos. ¿Qué representación considera la mejor y por qué? [2 puntos]
3. Con Python/R (o ambos), crea su propia representación de unas de las métricas disponibles en NGS QC toolkit y/o FastQC y descríbela (por favor enviar todo su código con el informe y poner el grafico resultante en el cuerpo del informe con descripción). [4 puntos]

Para el grafico, usa unos de los FASTQ aquí:

<ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/>

UCSC Genome Browser [8 puntos]

Escenario: Tres pacientes relacionados tienen una comunicación interauricular (ing. atrial septal defect) en el corazón. Al secuenciar sus genomas, los investigadores han encontrado una delección en el cromosoma 8 ocupando las siguientes coordenadas 11,750,209-11,750,217 (genoma de referencia: hg38).

1. Usando el 'Custom Track' del UCSC Genome Browser, representa esta delección. [3 puntos]

Para responder a las siguientes preguntas, por favor solo usar el Genome Browser y los recursos vinculados:

2. Identifica el gen y el exón que se ve afectado por la delección. [1 punto]
3. Verifique lo que se sabe de algunas alelas en esta región. [1 punto]

4. ¿En qué órganos se expresan el gen afectado por la delección y como eso se relaciona a la enfermedad encontrada en los pacientes? [1 punto]
5. ¿Podríamos usar un modelo mamífero para ver si desarrollo también una enfermedad cardiaca? ¿Como lo ha determinado? [1 punto]
6. Elimina las pistas (ing. tracks) que no sirven para responder a las previas preguntas y capturar la imagen resultante. [1 punto]

Cobertura de Secuencias Alineadas [6 puntos]

Selecciona dos archivos BAM (y BAI) que contienen exomas humanos:

<ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/>

(bajo “exome_alignment”, los archivos que dicen “mapped”)

Usando BEDtools y R (o Python):

1. Representa en dos gráficos (uno para cada exoma) la cobertura en la región donde se ubica la delección que describimos en la sección anterior. [2 puntos]
2. Representa en dos gráficos (uno para cada exoma) la cobertura para todo el gen afectado. [2 puntos]
3. ¿Cuáles son las diferencias y similitudes en cobertura entre los dos exomas? [2 puntos]