

Práctica:

Anotación KEGG Orthology y Reconstrucción de Vías con KEGG Mapper

Sara Cuadros Orellana

KOALA (KEGG Orthology And Links Annotation) es la herramienta de anotación interna de KEGG.

BlastKOALA y GhostKOALA son herramientas de KOALA que asignan números K a los datos de secuencia del usuario mediante búsquedas BLAST y GHOSTX, respectivamente, contra un conjunto no redundante de GENES KEGG.

Los archivos de la base de datos para GhostKOALA son algo diferentes de los de BlastKOALA. Para cada grupo de organismos KEGG a nivel de género o familia, se genera un conjunto de datos no redundante al tomar todos los genes codificadores de proteínas del genoma representativo y genes adicionales de los otros genomas con dos criterios. Uno es el mismo que en BlastKOALA, diferentes números K, y el otro es **exclusivo de GhostKOALA**, diferentes grupos de CD-HIT, que se calculan con un límite de identidad del 50%. Además, el archivo de base de datos para virus es creado por CD-HIT con un límite de identidad del 90% de la categoría de virus de KEGG GENES. **Estas adiciones están destinadas a analizar composiciones taxonómicas de metagenomas.**

Ingrese a la página de GhostKOALA: <https://www.kegg.jp/ghostkoala/>

Trabajaremos con datos del mismo proyecto que analizamos en el Tutorial 1, sin embargo, necesitaremos un archivo multi-fasta de aminoácidos.

Sample Hypersaline water prokaryotic community from SS33 pond from Santa Pola Saltern, Spain: <https://www.ebi.ac.uk/metagenomics/samples/SRS478661>

```
> Analysed associated runs
  > Analyses
    > Download
      > Predicted CDS (.faa)
        SRR979792_FASTQ_CDS_unannotated.faa
```

O directamente a través de uno de estos enlaces:

https://www.ebi.ac.uk/metagenomics/api/v1/analyses/MGYA00002655/file/SRR979792_FASTQ_CD_S.faa.gz

<https://drive.google.com/file/d/144n-cogHflfxFPuk4UyLRx7o3pzXpE2S/view?usp=sharing>

Subir el archivo, informar su dirección de correo y observar la llegada del mensaje. Hay que seguir un **enlace** para que el trabajo sea iniciado. El término del análisis también se comunicará a su dirección e-mail.

Explorar los resultados y Reconstruir Vías usando **KEGG Mapper** ([enlace en resultados](#)).