Para realizar el análisis de secuencias de Mycobacterium Tuberculosis. vamos a realizar los siguientes pasos

Seguiremos este diagrama de flujo.

**1.- Docker**

Primero verificamos la instalación correcta de Docker en nuestro equipo.

*docker run --name conda-ubuntu\_WGS -d -v TU\_CARPETA\_EN\_WINDOWS:/home/WGS\_Delfin/ continuumio/miniconda3 tail -f /dev/null*

*docker exec -it conda-ubuntu\_WGS /bin/bash*

Procedemos a instalar y actualizar Anaconda y bioconda.

*conda config --add channels defaults*

*conda config --add channels bioconda*

*conda config --add channels conda-forge*

**2.- Instalar MTBseq**

*conda install -c bioconda mtbseq*

Procedemo s configurar GATK

Nos desplazamos a la carpeta de MTBseq

cd /opt/conda/share/mtbseq-1.0.4-2/

dentro de esta carpeta descargamos GATk con el siguiente link

*wget https://storage.googleapis.com/gatk-software/package-archive/gatk/GenomeAnalysisTK-3.8-0-ge9d806836.tar.bz2*

Descomprimimos el archivo

*tar -xvf GenomeAnalysisTK-3.8-0-ge9d806836.tar.bz2*

Verificamos la instalación

Check the version:

*MTBseq --version*

Check the dependencies:

*MTBseq --check*

3.- Correr MTBseq

3.1-ENSAMBLE

*MTBseq –step TBbwa –threads 4*

3.2-CALIBRACIÓN DEL ENSAMBLE

*MTBseq –step TBrefine –threads 4*

3.3-FORMATO DE ACUMULACIÓN

*MTBseq –step TBpile –threads 4*

3.4-TRANSFORMACIÓN DE ACUMULACIÓN A LISTA

*MTBseq –step TBlist –threads 4*

3.5-LLAMADA DE VARIANTES

*MTBseq –step TBvariants –threads 4*

3.6-VALORACIÓN DEL ENSAMBLE

*MTBseq –step TBstats –threads 4*

3.7-CLASIFICACIÓN DEL LINAGE

*MTBseq –step TBstrains –threads 4*

3.8-TABLA COMPARATIVxxA DE SNPs

*MTBseq –step TBjoin --snp\_vars –samples samples.txt --project WGS\_delfin –threads 4*

3.9-POSPROCESAMIENTO DE TABLA DE SNPS EXPORTAR A OTROS FORMATOS

*MTBseq –step TBamend --snp\_vars –samples samples.txt --project WGS\_delfin –threads 4*