

In silico serotyping

Install SeroBA (Epping et al 2018) as per instructions at <https://github.com/sanger-pathogens/seroba#installation> and git clone the database from the following link <https://github.com/sanger-pathogens/seroba.git>

Files required to run serotyping using SeroBA:

1. paired-end fastq files
2. Database
3. sample list (only for running on multiple samples)

Run in silico serotyping on a single sample :

`serotype runSerotyping <full path to the database> <read 1> <read 2> <output folder prefix>`

Run in silico serotyping on multiple samples:

1. create a list of sample names and save it as `samplelist` (e.g. the sample name for 24371_8#283_1.fastq.gz is 24371_8#283)
2. `for f in $(cat samplelist); do seroba runSerotyping <path to the database> ${f}_1.fastq.gz ${f}_2.fastq.gz ${f}; done`
3. `seroba summary ./`

Output:

`summary.tsv`