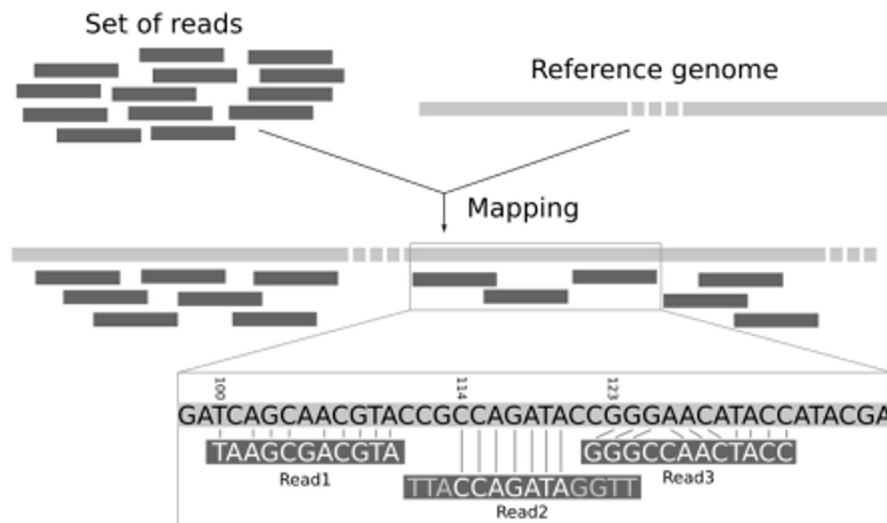# Constructing a Phylogeny

# Alignement

The first step to construct a phylogeny is to build an *alignment*

*This will align the reads to a specific location in the genome.*
*The reads can be aligned to a reference genome*



*Reference genome*
- Representative of the species or a lineage/clone
- ATCC700669 genomes (NCBI accession number FM211187)
- Should be of high quality

# SNP alignment

*Why do we build a snp alignment?*

Identification of variable sites between the reference genome and our data.

# Phylogeny

*Phylogeny* studies relationships between species, individuals or genes.
It helps to understand how genes, genomes, species evolve.

*Phylogenetic trees* based on whole genome data tell us about the relationships of bacterial isolates to each other.
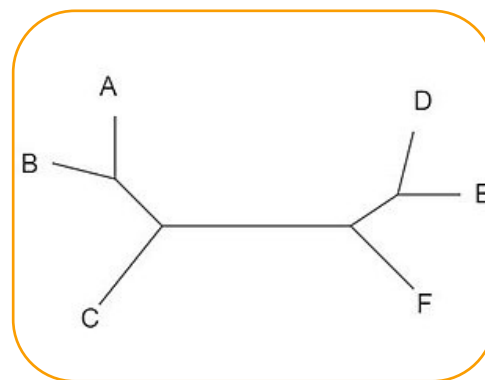
# Root the phylogeny

*Root position* affects interpretation because it tells us the direction of evolution. Thus, affecting statements about relatedness between samples

It allows us to say that "*A is more closely related to B than it is to C*"
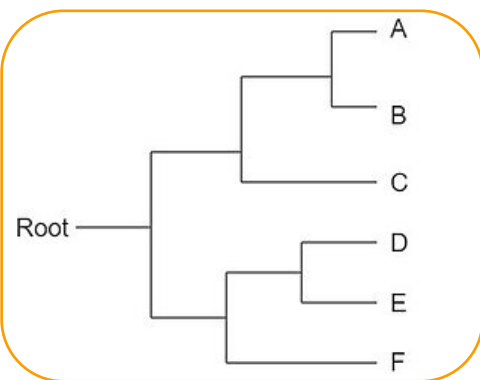


*Rooted phylogenetic* tree shows ancestry relationship



*Unrooted phylogenetic only shows relatedness of organisms*

# Phylogeny

- Species-wide tree

- Reads map to a general reference

- Lineage-specific reads will be mapped onto the reference

- As compared with lineage-specific phylogeny, lower resolution can be viewed

- GPSC and CC correlate well with the phylogeny
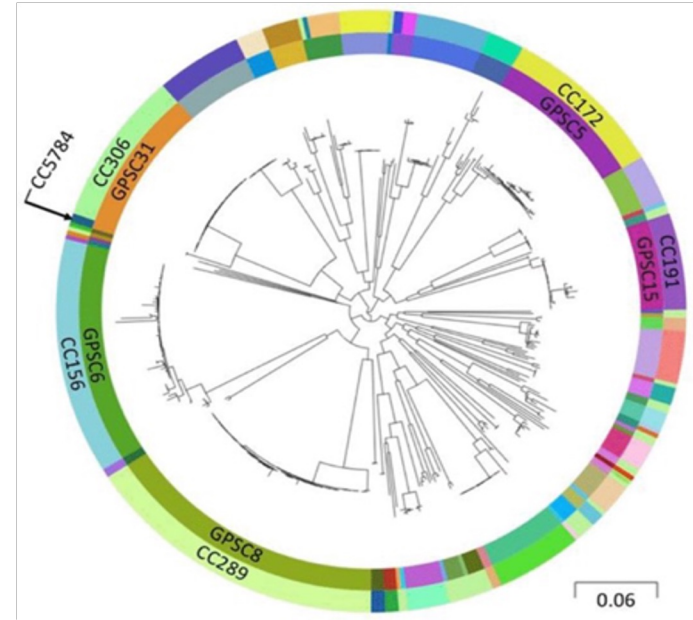
# Microreact

Open data visualization of genomic surveillance data.



merged_data.csv

argentina_snp_tree.nhx