# *Mycobacterium tuberculosis* NGS made easy: data analysis step-by-step

We hope that you are enjoying the training, and that all the things that you are learning will be useful in your current and future research!

Please, read the questions, and provide the answers in the table below for **at least one** of the points in 1) 2) and 3)

1. **Mapping and Variant Calling**
    a. After SNP calling, we observe a genomic region that has almost double the sequencing depth than the rest of the genome. What is a possible explanation for this phenomenon?

2. **Clustering and DR prediction**
    a. We find two patients with rifampicin-resistant tuberculosis to be linked by recent transmission (in cluster). Based on our genomic analysis, we also know that rifampicin resistance evolved *de novo* in both patients, as opposed to one patient transmitting an already resistant strain to the other. What specific information did we need to look at to be able to know this?
    b. Would you be able to say whether there are imported cases of tuberculosis? Why?

3. Phylogenetics
    a. What is the difference between a rooted and an unrooted phylogenetic tree? Why is the former more informative?

| Your Name | Question number | Answer |
|---|---|---|
| | | |

| | | |
|---|---|---|
| Olga Shavuka | 3 | A rooted tree shows the evolutionary direction and common ancestor, while an unrooted tree only shows relationships between strains. A rooted tree is more informative because it helps understand the order of evolution and transmission over time. |

|  |  |  |
|---|---|---|
| Michelle Amma Yeboah Amoako-Nimako | 2 | To determine that rifampicin resistance evolved independently in both patients, we need to compare the specific resistance-conferring mutations in the *Mycobacterium tuberculosis* genomes. If we observe different resistance mutations despite clustering, then this suggests that there is de novo evolution rather than direct transmission of a resistant strain.<br><br>Yes, imported TB cases may be identified if isolates cluster more closely with strains from another geographic region than with local strains. Phylogenetic comparison with global reference datasets can therefore suggest imported cases of TB. |

| | | |
|---|---|---|
| Noutin Michodigni | 1. **Mapping and Variant Calling** | The observed increase in sequencing depth in that genomic region is likely due to the presence of **repetitive sequences**, particularly from the **PE/PPE gene family** in the *Mycobacterium tuberculosis* genome. These regions contain highly similar or duplicated sequences, causing multiple reads to map to the same locus during alignment. As a result, the coverage appears artificially elevated compared to the rest of the genome |

| | | |
|---|---|---|
| Saidou Mahmoudou (Gabon | 3 | A rooted phylogenetic tree explicitly indicates the direction of time and the most recent common ancestor of the group of species under study, whereas an unrooted tree shows only the relationships and evolutionary distances between organisms, without specifying the chronological order of divergence<br>the rooted tree is more informative because it provides layers of crucial biological information that an unrooted tree does not |

|  |  |  |
|---|---|---|
|  |  |  |

|  |  |  |
|---|---|---|
|  |  |  |

|  |  |  |
|---|---|---|
|  |  |  |

|  |  |  |
|---|---|---|
|  | 1. |  |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

|  |  |  |
|---|---|---|
|  |  |  |

|  |  |  |
|---|---|---|
|  |  |  |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

|  |  |  |
|---|---|---|
|  |  |  |

|  |  |  |
|---|---|---|
|  |  |  |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

|  |  |  |
|---|---|---|
|  |  |  |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

| | | |
|---|---|---|
| | | |

|  |  |  |
|---|---|---|
|  |  |  |