

Tracking transmission and evolution of pathogens

In this section, we will overlay associated data: metadata, location data, genomic predictions onto the phylogenetic tree to track transmission and evolution of pathogens. We will use the free to use web-tool Microreact developed by the Centre for Genomic Pathogen Surveillance (CGPS) team. In the microreact (<https://microreact.org/>) you can upload your phylogenetic tree and metadata in csv format. The metadata must contain an “ID” column which should contain the IDs matching those present in the phylogenetic tree. Here, we will use example data to explore functions of the microreact. Afterwards, we will carry out an exercise to track the transmission and spread of antimicrobial resistance.

Understanding the emergence, spread and transmission of pathogen with Microreact

Understanding the population structure of *Salmonella typhi* in Papua New Guinea (Dyson ZA et al, 2022). A total of 86 isolates collected over 30 years were whole genome sequenced. Of the 86 isolates 84 (98%) of them belonged to the lineage 2.1.7. Our aim here would be to study the diversity in this lineage. We will then analyze the isolates in context with isolates from other countries (n=49) to understand emergence and evolution of this lineage.

Given data:

We have already created a phylogenetic tree and have put metadata together with the geo-location of the individual isolates together in a microreact project. Place the link: https://microreact.org/project/icGPbF_O8, in your web-browser (preferably firefox web-browser) to enter the project as shown in figure 1. The tool microreact combines phylogenetic tree, metadata and the geolocation shown as three separate windows.

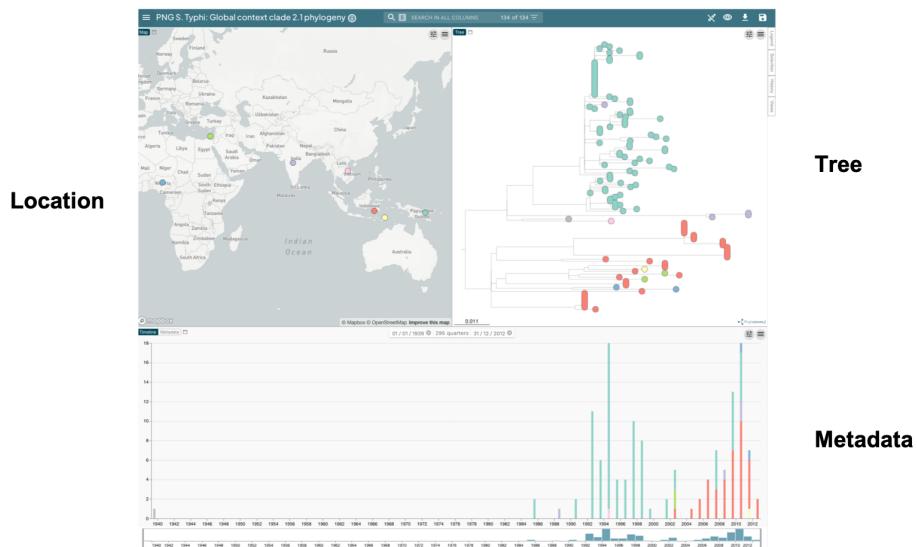


Figure 1

Now click on the tab “metadata” as shown by the circle1 in the figure 2 to see the associated metadata for the 133 isolates.

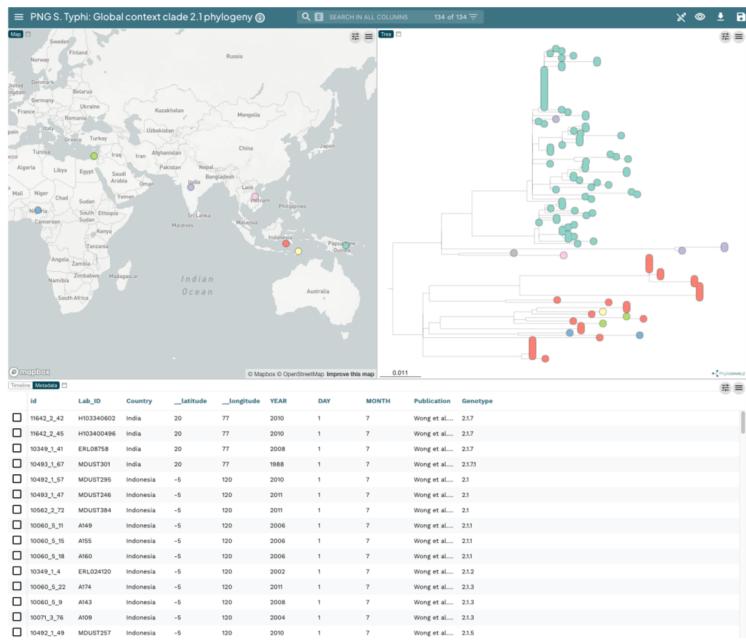


Figure 2

Microreact has a range of options for changing the layouts and overlaying metadata onto the phylogenetic trees to improve visualization. Click on the show controls menu (circle 1, figure 3) in the tree window, this will show various options to change the tree layout (circle 2) and display labels and metadata. For this practical purpose, keep the default rectangular layout of the tree. Click on the circle1 again to hide the options.

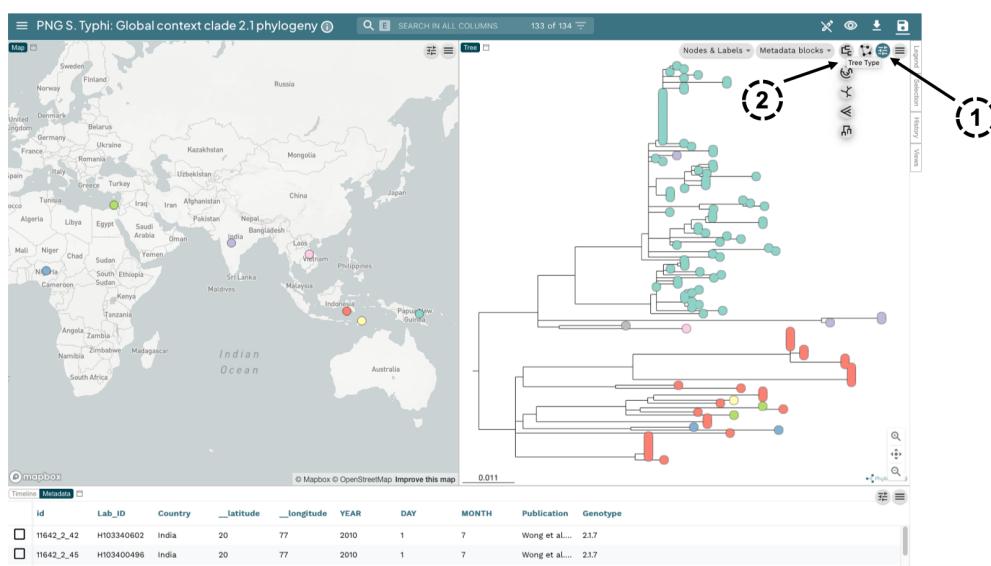


Figure 3

Diversity of *Salmonella* typhi from Papua New Guinea

First we will colour the tips of the tree to reflect the countries they are isolated from. For this, click on the eye icon (circle 1, figure 4) and select “country” for the “colour column” option.

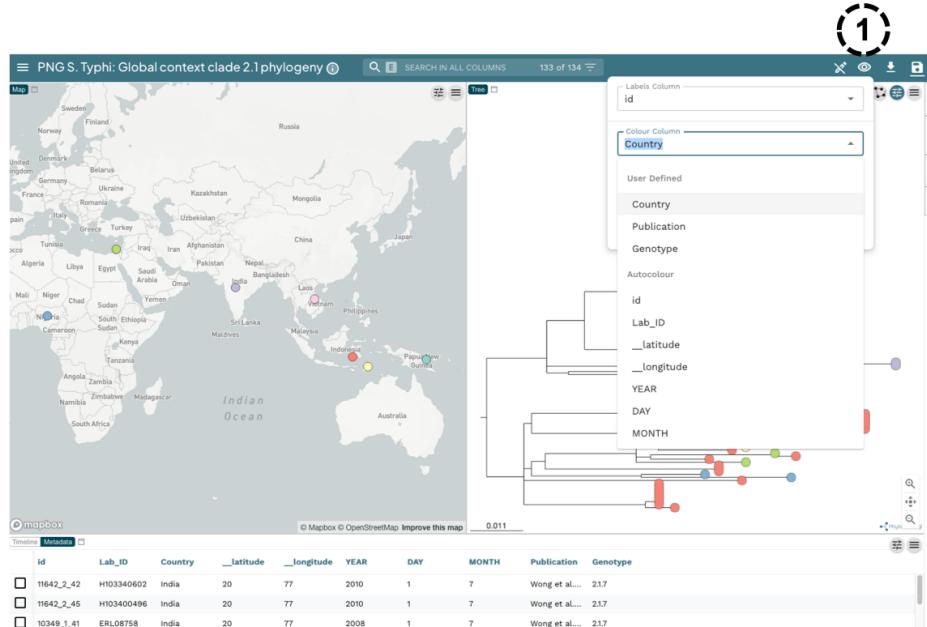


Figure 4

Now click on the legend option on the right side of the browser (circle 1, figure 5). This will show the country the samples belong to.



Figure 5

Now we select the internal node when all the isolates from Papua New Guinea (circle 1, figure 6). We can see that all 84 isolates form a single monophyletic cluster with just 1 isolate from India (figure 6). Further, this cluster can be divided into two subclusters. Therefore, it can be inferred that the isolates from Papua New Guinea have a unique genetic structure and have minor evidence of inter-country transmission, shown by lack of isolates from other countries within this cluster. Single isolate from India was a case of traveler associated infection. The authors compared genetic differences to show that these isolates were less diverse than the isolates from neighboring countries. Also, they found that, unlike isolates from other countries most of the isolates from Papua New Guinea were susceptible to most of the drugs used to treat *Salmonella typhi* infections.

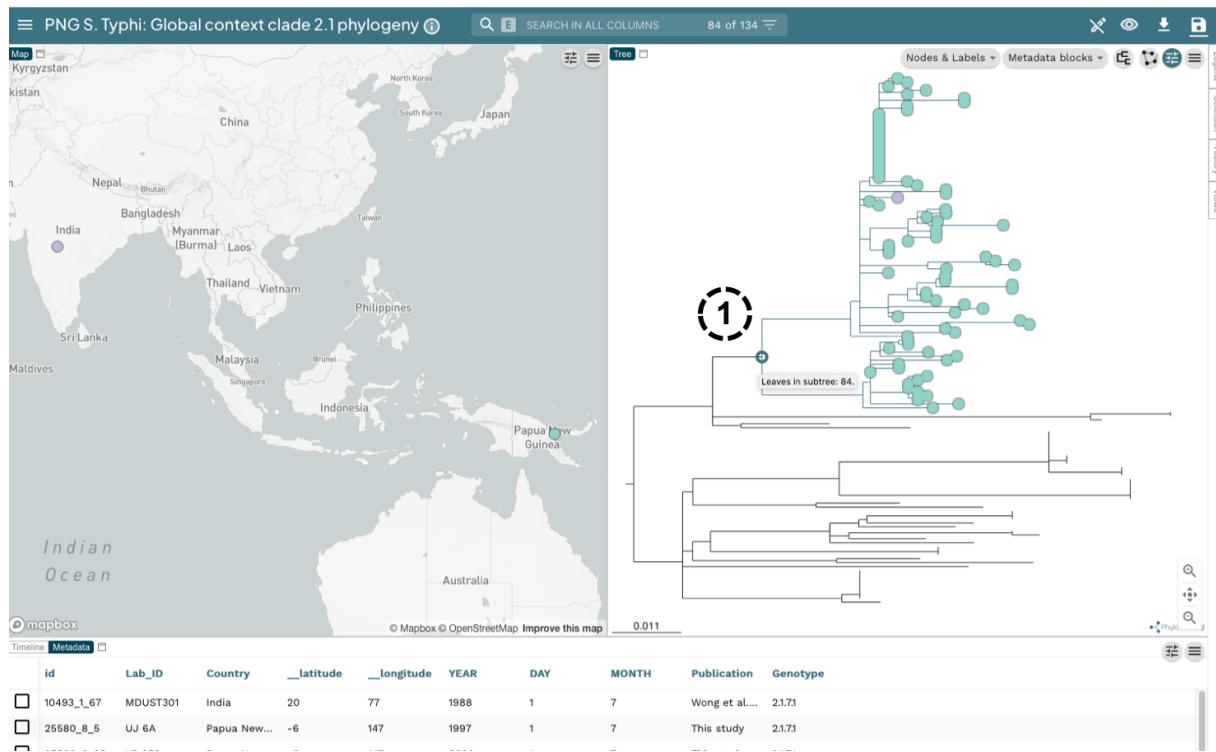


Figure 6

Origins of Papua New Guinea lineage 2.1.7

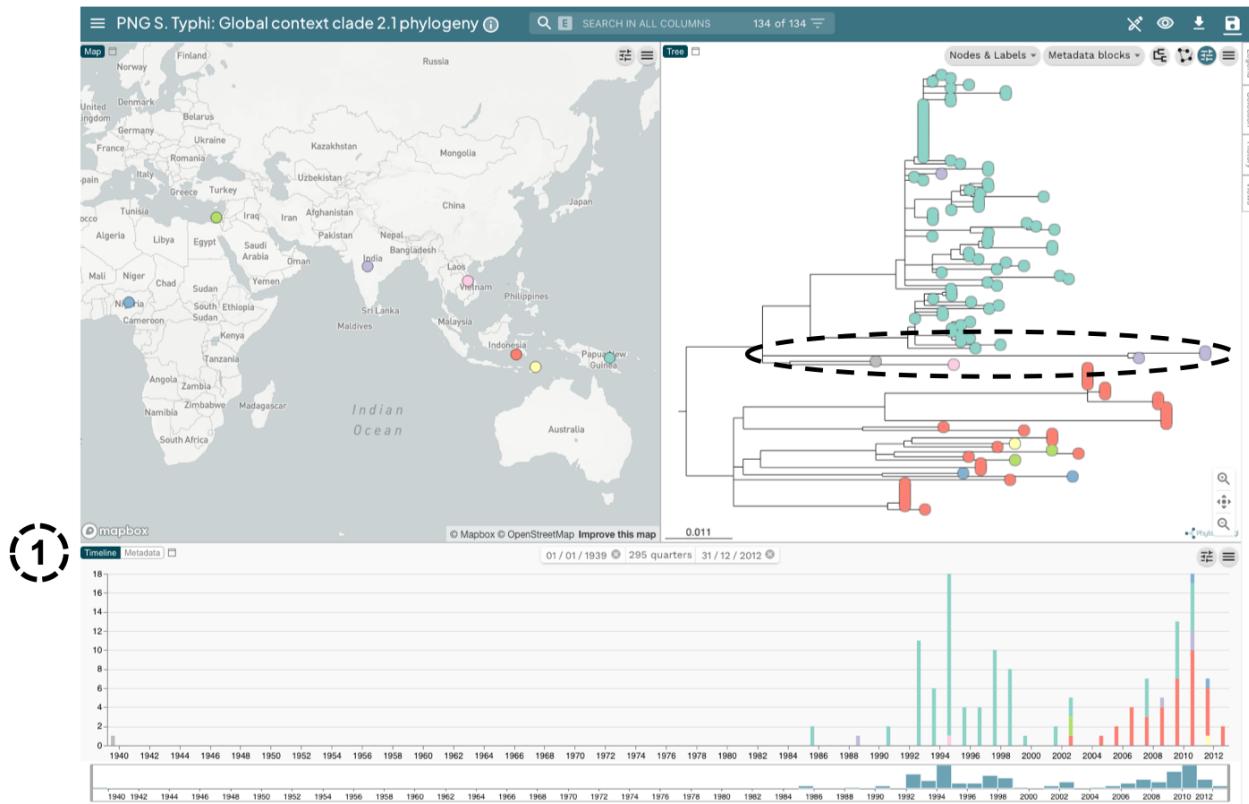


Figure 7

Now, click on white space outside the tree to select all the isolates. On the metadata window select the timeline tab (circle 1, figure 7) to see the time when samples were collected. From the phylogenetic tree and the metadata, we can see that most of the line 2.1.7 isolates from Papua New Guinea were collected prior to the isolates from other countries. The isolates closest to the node separating Papua New Guinea isolates are from India and Vietnam (encircled by the dashed lines, figure 7). This shows that these isolates shared a common ancestry with Papua New Guinea isolates which might have originated in southeast asia.

Exercise: Tracking transmission and spread of antibiotic resistance

In this exercise we will aim to investigate the origins of an outbreak of extensively drug-resistant *Salmonella typhi*. We have already provided the phylogenetic tree file (Klemm2018_tree.nwk) and metadata file with geo-locations added (Klemm2018_metadata.csv). We will upload these two files in microreact to create a project.

Step1: Creating a microreact project

Copy the microreact link: <https://microreact.org/> into the firefox web-browser. You will see a window as shown in figure 8. Now, click on the “upload” option to the top of the window as marked by the circle 1 in the figure 8.

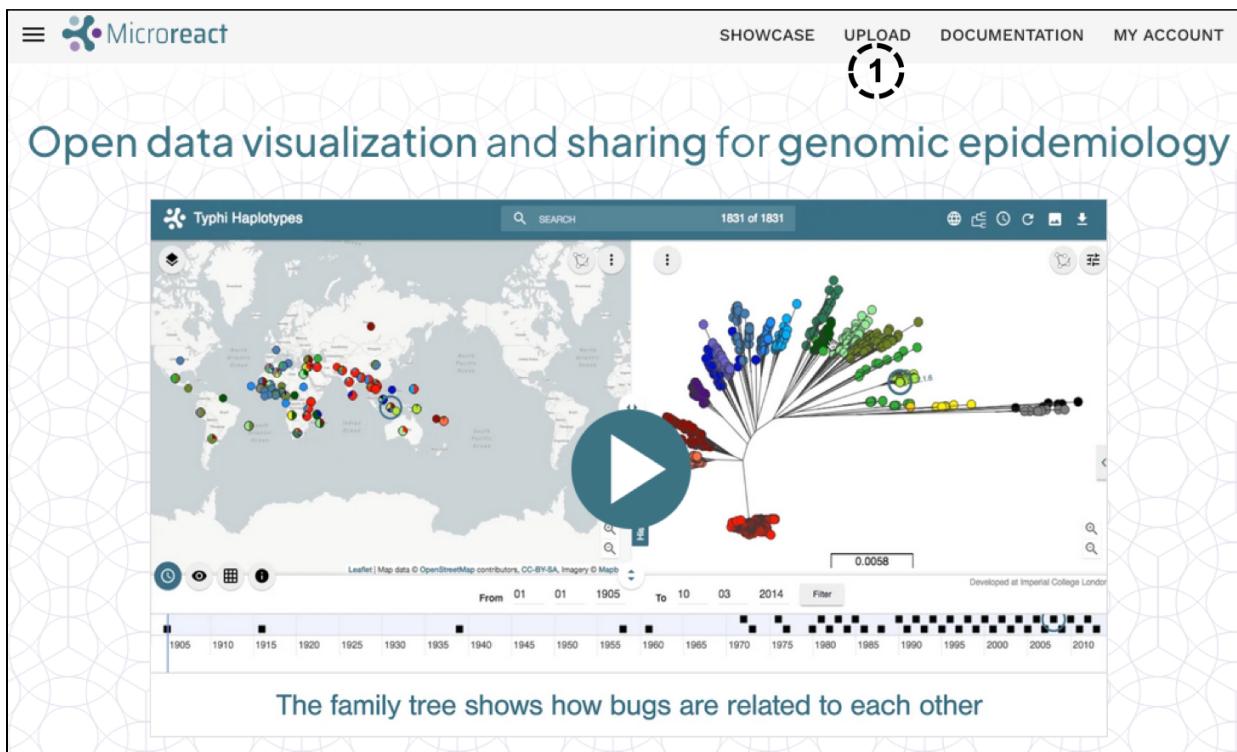


Figure 8

Step2 : Upload the files provided to you by dragging them into the window or you can click on the “browse your files” option as shown by circle 1 figure 9. When prompted click continue, which will create a window with the three panels as shown in the figure 10.

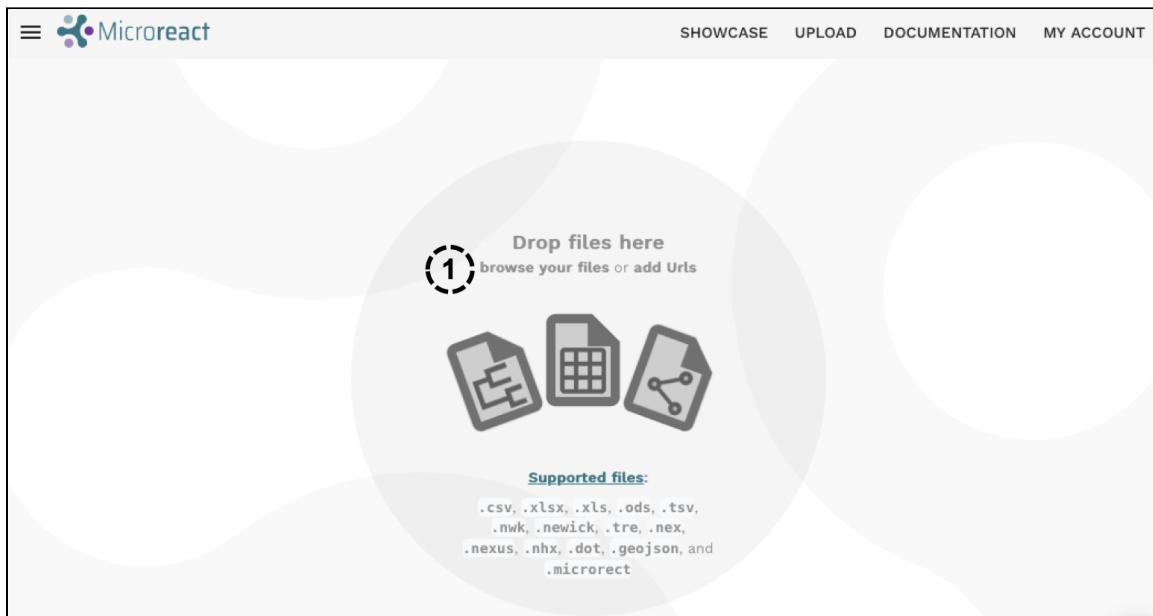


Figure 9

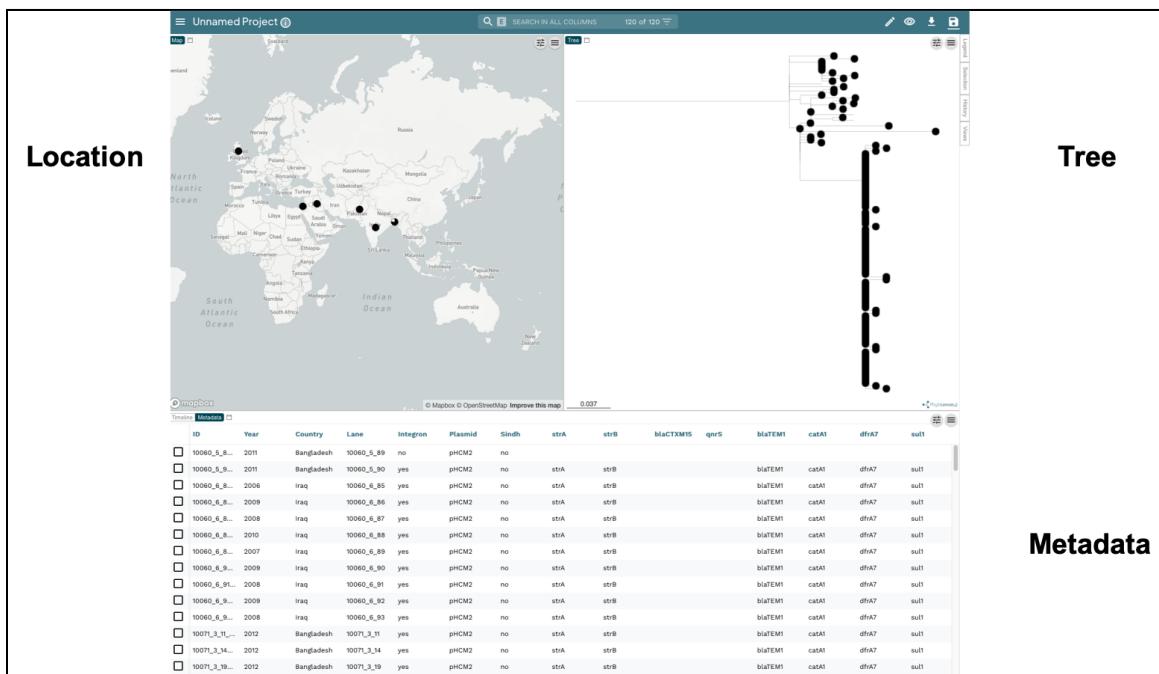


Figure 10

Identifying the origin

Now click on the show option icon (circle 1) and from the tree layout options (circle 2) select the vertical layout option (circle 3) as shown in figure 11.

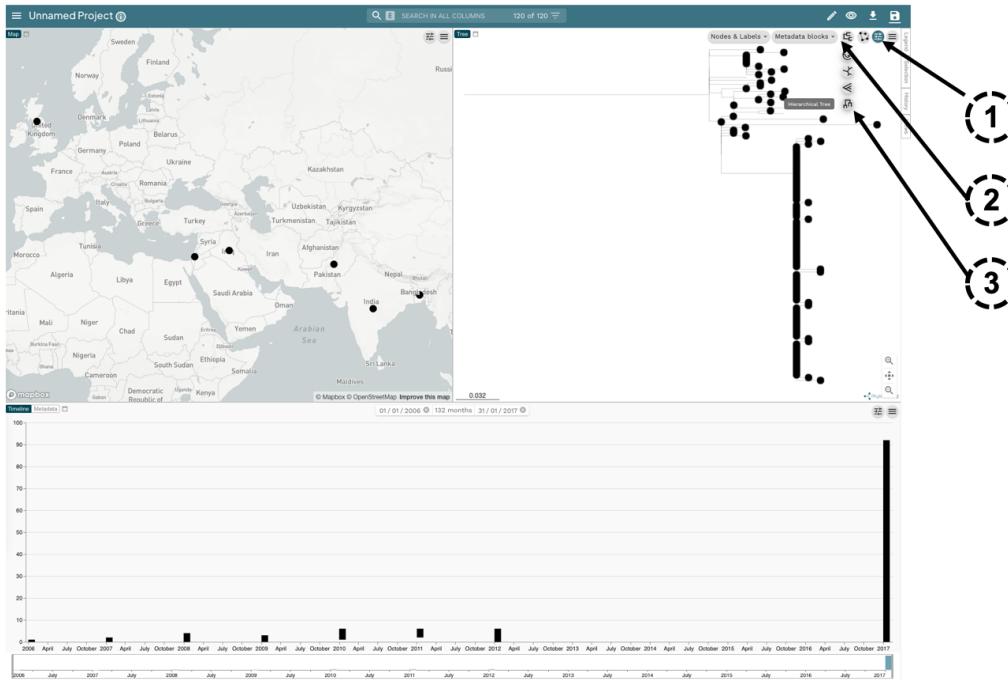


Figure 11

Now we will identify the isolates collected during the outbreak in Sindh province, click on the eye option (circle 1) and from the options in the “labels column” tab (circle 2) select “sindh” (circle 3) as shown in the figure 12.

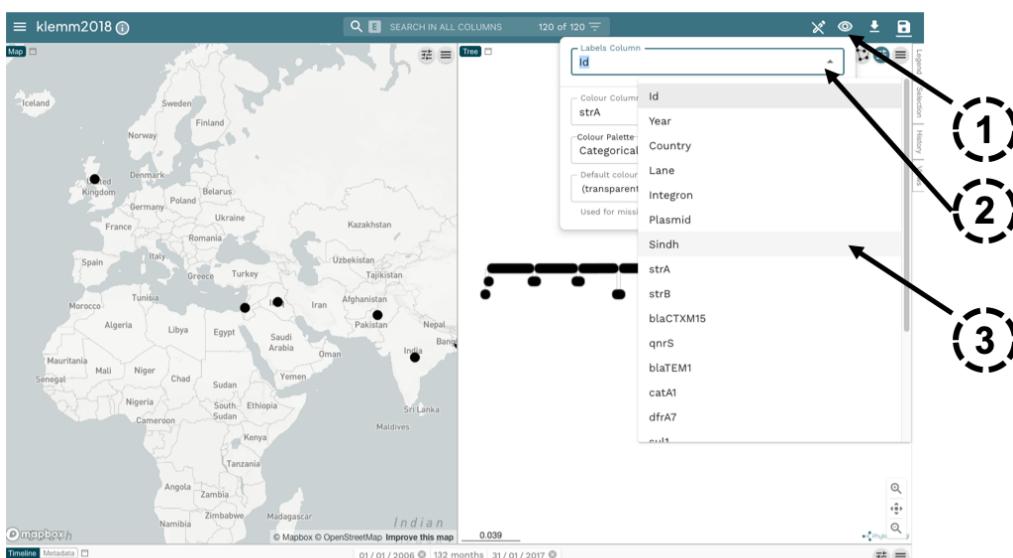


Figure 12

In order to see the labels for each isolate, click on the show options (circle 1), and under the Nodes and Labels tab (circle 2) select “leaf labels” (circle 3) and adjust the size of the labels by dragging the gliding window (circle 4). For better visibility turn off the “align leaf labels” option (circle 5) as shown in the figure 13. Under the metadata window select the timeline option and drag the bar to see isolates collected after 2016 as the outbreak occurred in 2017 (circle 6).

Now you will see a window as shown in figure 14. Here we can clearly see the outbreak clone formed a monophyletic clade characteristic of the outbreak. This lack of diversity and clonal genetic makeup of the isolates suggest a single source of the outbreak. If we look over in the geolocations window, we can see all these isolates were collected from Pakistan except for one but this was the case of a travel associated infection from Pakistan.

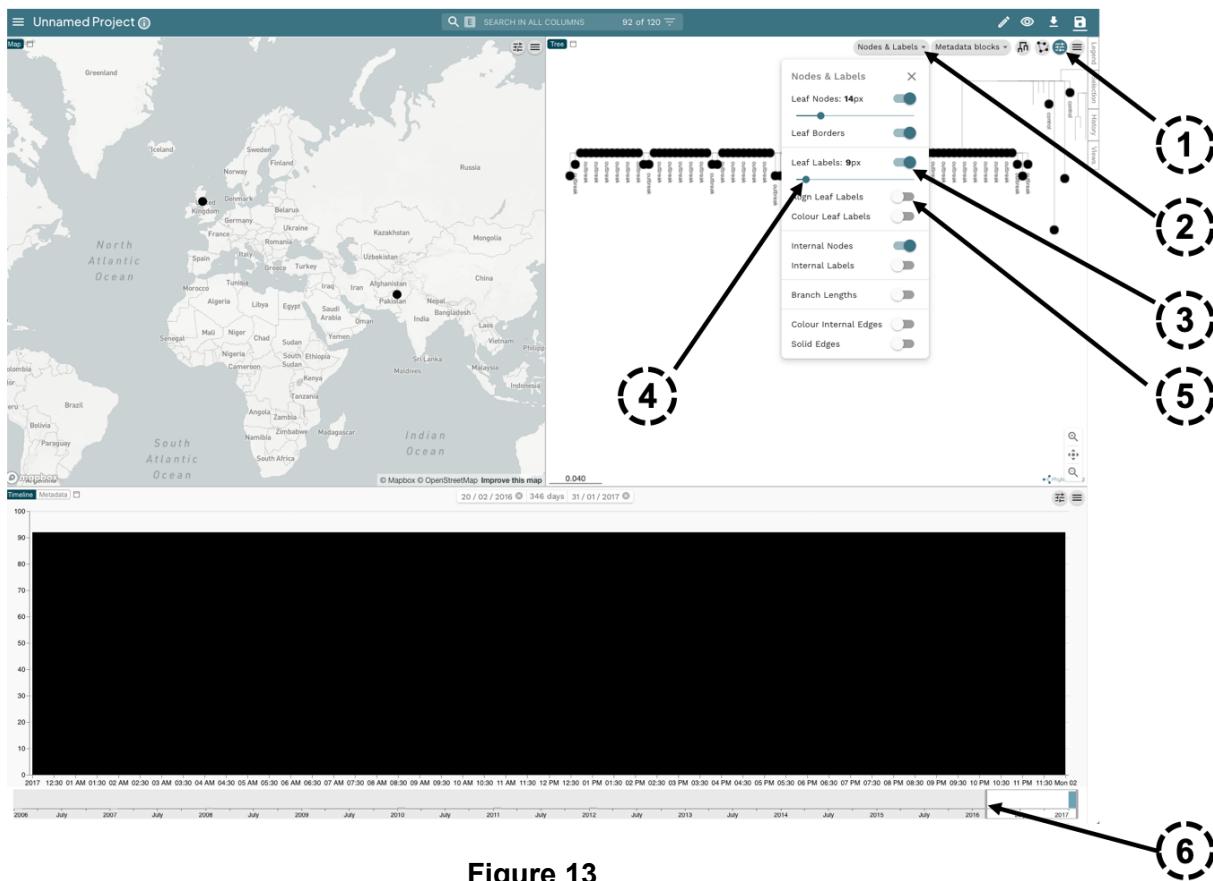


Figure 13

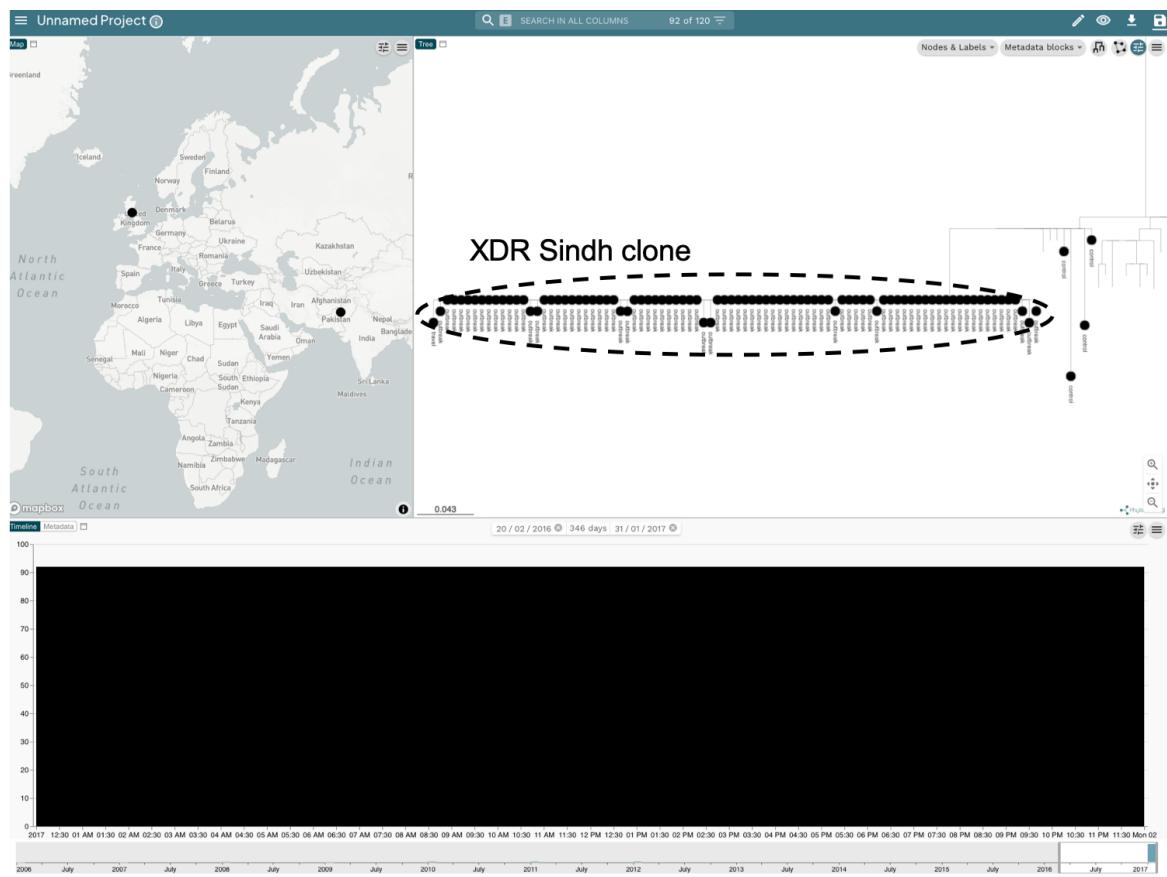


Figure 14

Similarly, if we select all the isolates by moving the slide in the timeline window (shown above) back to 2006 and also change the labels to countries (as shown in the figure 12). We will see a window as shown in figure 15.

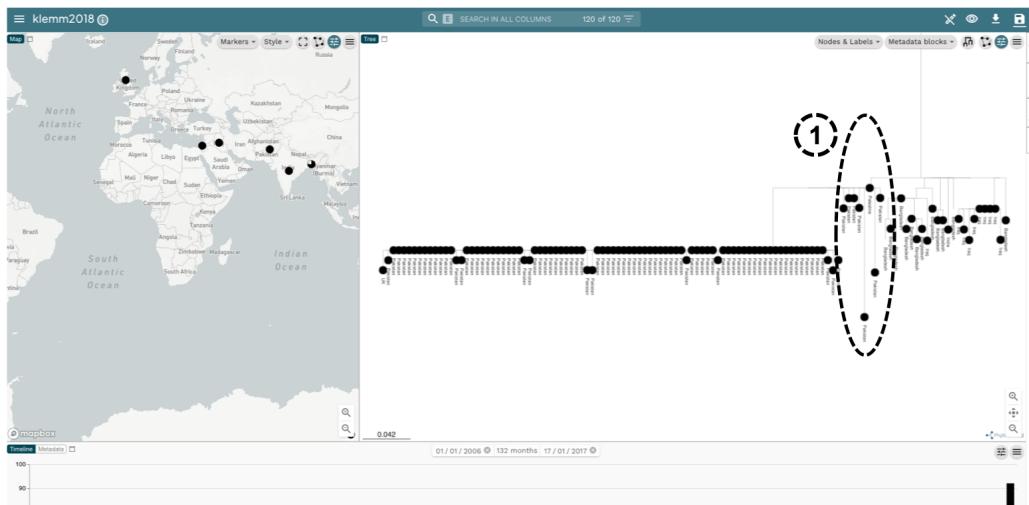


Figure 15

In figure 15, if we zoom in on the penultimate node of this clone we can see all the isolates were collected from Pakistan and these isolates share an ancestor with the XDR clone. So it can be inferred that this clone emerged and caused an outbreak in Pakistan.

Tracking spread of antimicrobial resistance

Now we will look deeper into the genetic resistance harbored by the outbreak isolates. We know from the study that the outbreak isolates were extensively drug resistant and were difficult to treat with usual drugs. The study identified these isolates as XDR meaning very few options were left to treat such infection. Using the metadata and the phylogenetic tree we will now investigate the mechanisms of resistance for these outbreak isolates and what led to the emergence of this XDR clone. Using microreact it is easier to compare all the data related to an isolate with others without any command-line use.

We can directly search the presence of various resistance determinants by typing the gene name into the search box on the top of the microreact project window as shown (circle 1) in the figure 16.

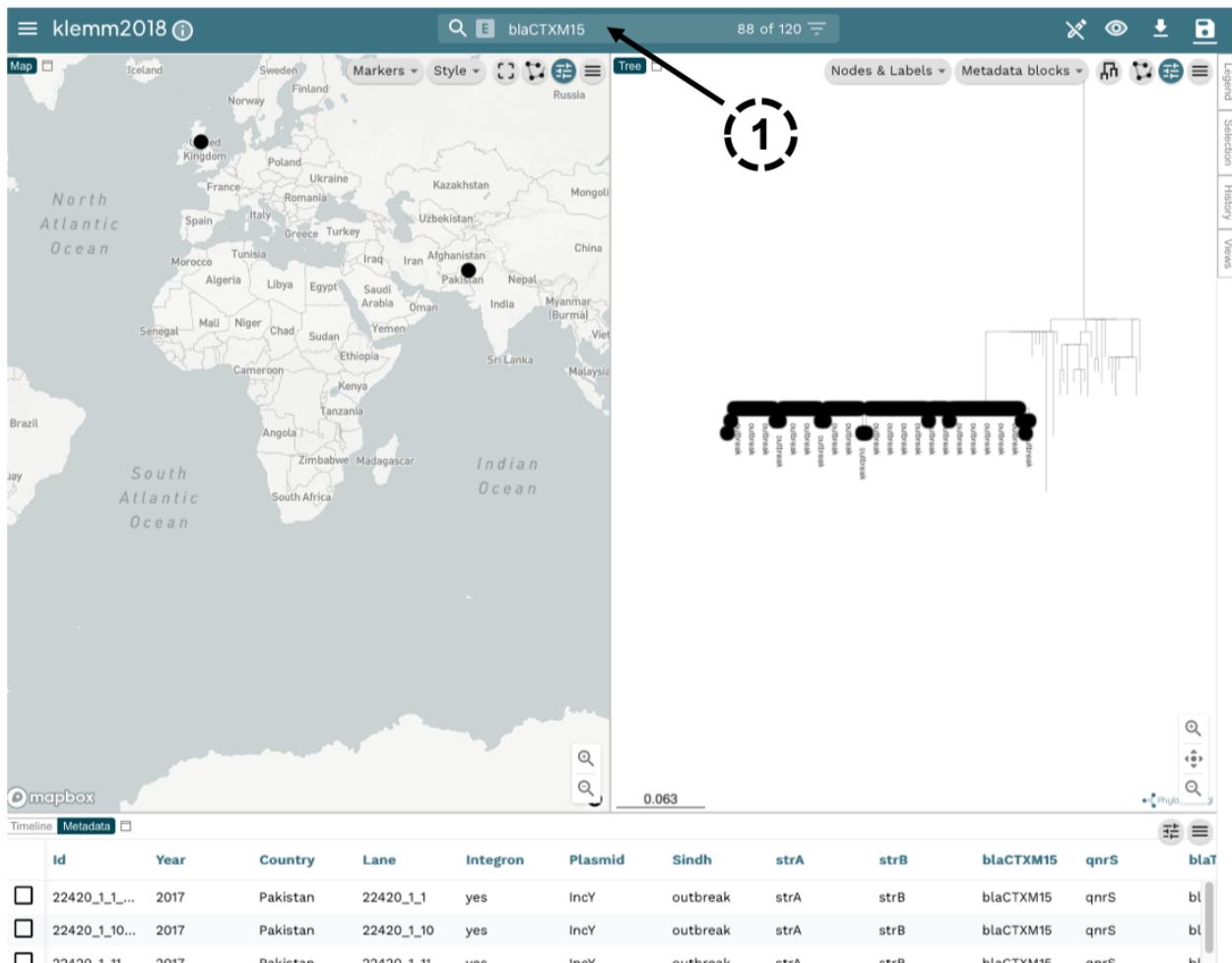


Figure 16

You are welcome to identify the isolates with resistance conferring genes such as *catA1* (confers resistance to chloramphenicol), *dfrA7* (trimethoprim), *sul1* and *sul1* (sulfamethoxazole), *strA* and *strB* (streptomycin). The outbreak isolates all contained *blaCTX-M-15* (confers resistance to third-generation cephalosporins) and *qnrS* (confers resistance to fluoroquinolones). In the study the authors identified that these genes specific to the outbreak clone were carried on a plasmid *incY*. This shows that XDR clone emerged in Pakistan and acquired the resistance via the acquisition of the *incY* plasmid.

Note: This practical is adapted from course material originally developed by Francesc Coll for the Wellcome Connecting Science course “Bacterial Genomes: Antimicrobial Resistance in Bacterial Pathogens”