

Computation Practical 9: Phylogenetics Visualisation

Exercise 1: Investigating an outbreak of CPE cases

Introduction

In this practical we will investigate the **CPE cases** we have collectively collected data for during the course. We will specifically focus on the ST78 *Klebsiella pneumoniae* isolates of the CPE cases. We will integrate the epidemiological and antibiotic susceptibility data collected using **EpiCollect** in previous practicals, along with the phylogenetic tree we generated. We will also contextualise our local hospital cases in relation to other isolates collected from other regions in the world.

We will make use of **Microreact** (<https://Microreact.org/>), a web application that provides an interactive visualization of datasets via phylogenetic trees, maps, timelines, and tables. But first we will download the epidemiological data collected using EpiCollect to draw conclusions on the origin and spread of the CPE outbreak cases.

Download collected data from EpiCollect

Visit the following URL to visualise the data collected by all participants:

<https://five.epicollect.net/project/thailand-cpe-outbreak>

Login using the provided credentials to access the project main page (Figure 1). Click on 'VIEW DATA' to access all entries in this project.

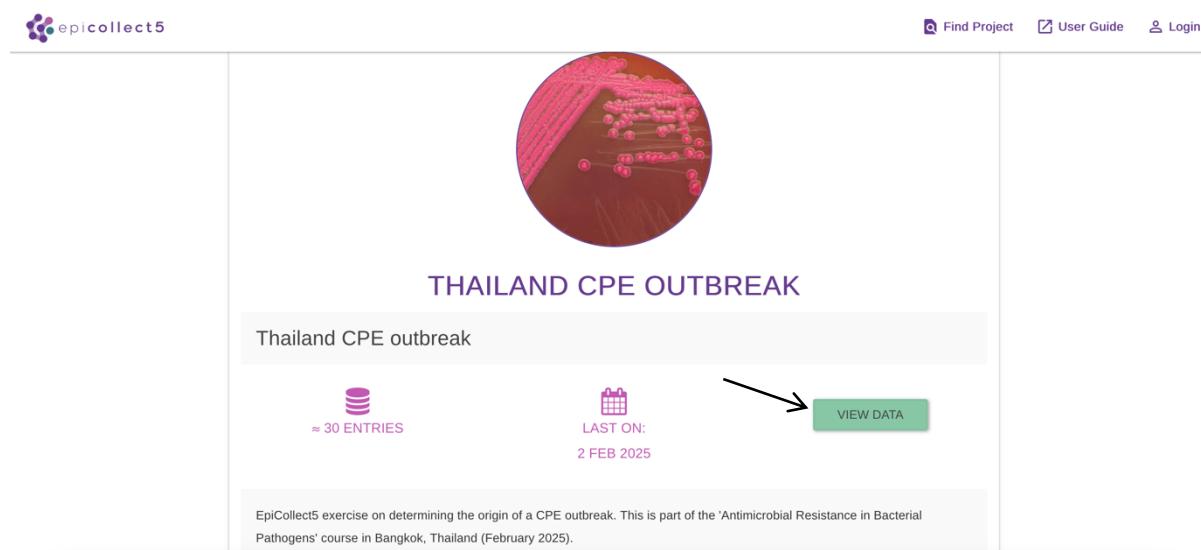


Figure 1 Thailand CPE outbreak EpiCollect project page

You will see then the data collected by all course attendees as a table (Figure 2).

View	Title	Created At	Barcode	Strain ID	Patient ID	Gender	Age	Longitude/Latitude
	cped010	2nd Feb, 2025	cpe010	cped010	P_022	Male	67	
	cpe093	2nd Feb, 2025	cpe093	cpe093	P_097	Male	77	13.758131, 100.536403
	cpe108	2nd Feb, 2025	cpe108	cpe108	P_108	Female	67	13.758093, 100.536418
	Cpe013	2nd Feb, 2025	cpe013	Cpe013	P_029	Female	70	13.757591, 100.536285
	cpe035	2nd Feb, 2025	cpe035	cpe035	P_008	Female	73	13.75792, 100.536427
	cpe004	2nd Feb, 2025	cpe004	cpe004	P_016	Female	59	13.757811, 100.536324
	cpe049	2nd Feb, 2025	cpe049	cpe049	P_059	Male	61	13.757539, 100.536095
	cpe025	2nd Feb, 2025	cpe025	cpe025	P_039	Male	63	13.758082, 100.536317
	cpe070	2nd Feb, 2025	cpe070	cpe070	P_084	Male	69	13.757582, 100.536306

Figure 2 Thailand CPE outbreak EpiCollect project data

To be able to download this data locally we need to Login first (arrow in Figure 2).

Click on ‘Download’ (Figure 3) to download all entries as a CSV file.

Created At	Barcode	Strain ID	Patient ID	Gender	Age
2nd Feb, 2025	cpe010	cped010	P_022	Male	67
2nd Feb, 2025	cpe093	cpe093	P_097	Male	77
2nd Feb, 2025	cpe108	cpe108	P_108	Female	67
2nd Feb, 2025	cpe013	Cpe013	P_029	Female	70
2nd Feb, 2025	cpe035	cpe035	P_008	Female	73
2nd Feb, 2025	cpe004	cpe004	P_016	Female	59
2nd Feb, 2025	cpe049	cpe049	P_059	Male	61

Figure 3 Download EpiCollect entries

Unzip the file ‘thailand-cpe-outbreak-csv.zip’ and open it with Excel. We will need to edit this file first to be able to read with MicroReact. Insert a new column on the left-hand side of the Excel spreadsheet. Right-click on the first column and select ‘Insert’ as show in Figure 4.

Figure 4 Editing EpiCollect entries using Excel

Cut the column ‘1_Barcod’ and paste it on the newly inserted column as shown in Figure 5.

Figure 5 Editing EpiCollect entries using Excel

Change the name of the first column to ‘Id’ and delete the empty column (Figure 6).

A	B	C	D	E	F	G	H	I	J
1	Id	ec52uid	created_at	uploaded_at	title	2_Strain_ID	3_Patient_ID	4_Gender	5_Age
2	cpe010	9063a68b-0911-4576-a05b-ab6e905edd5d	2025-02-02T11:11:49.000Z	2025-02-02T11:11:56.000Z	cped010	cped010	P_022	Male	67
3	cpe093	7256ada4-4743-4d0e-ba3-84fb29654b3f	2025-02-02T11:11:33.000Z	2025-02-02T11:11:43.000Z	cpe093	cpe093	P_097	Male	77
4	cpe10	5aaafc5b-a-cd64-4ad7-9fb8-02dc3c38771	2025-02-02T11:10:47.000Z	2025-02-02T11:11:25.000Z	cpe108	cpe108	P_108	Female	67
5	cpe013	251cd864-a3d-4b2d-a9b5-8ff8c465e1c3	2025-02-02T11:10:46.000Z	2025-02-02T11:11:19.000Z	Cpe013	Cpe013	P_029	Female	70
6	cpe035	7bc7d789-126d-47e-96b5-c27d3656b38	2025-02-02T11:10:43.000Z	2025-02-02T11:10:52.000Z	cpe035	cpe035	P_008	Female	73
7	cpe004	97fbba06-3b46-40f1-bfc3-b756e0856823	2025-02-02T11:10:15.000Z	2025-02-02T11:11:19.000Z	cpe004	cpe004	P_016	Female	59
8	cpe049	d9b3c408-905e-4b30-b251-e40a053ee5b	2025-02-02T11:10:06.000Z	2025-02-02T11:10:14.000Z	cpe049	cpe049	P_059	Male	61
9	cpe025	430955cf-4bb0-490d-8da2-8f20469e6112	2025-02-02T11:10:05.000Z	2025-02-02T11:10:26.000Z	cpe025	cpe025	P_039	Male	63
10	cpe070	40413bb8d-bbab-4a1d-9d43-c0969393152e	2025-02-02T11:10:03.000Z	2025-02-02T11:10:11.000Z	cpe070	cpe070	P_084	Male	69
11	cpe079	cfdf5247-c-e361-4cf-ba05-158c846fcfa15	2025-02-02T11:10:10.000Z	2025-02-02T11:13:42.000Z	cpe079	cpe079	P_071	Male	85
12	cpe061	c42d02a8-4b7c-487c-80b6-f0626647ccb5	2025-02-02T11:10:37.000Z	2025-02-02T11:10:22.000Z	cpe061	cpe061	P_018	Female	71
13	cpe058	c19961e9-1fa7-4d22-b3c4-didd9c3af98	2025-02-02T11:10:35.000Z	2025-02-02T11:10:42.000Z	cpe058	cpe058	P_019	Female	59
14	cpe022	a8609cd-1cf-e-440b-b106-c7cf3f6ae01e	2025-02-02T11:10:27.000Z	2025-02-02T11:09:36.000Z	cpe022	cpe022	P_036	Female	78
15	cpe110	8c7eb7e-76-c085-4e1c-8696-bbb3bca60884	2025-02-02T11:09:24.000Z	2025-02-02T11:09:44.000Z	cpe110	cpe110	P_110	Male	62
16	cpe076	921cccb3-26bc-4872-8bdc-14a47e9038	2025-02-02T11:09:20.000Z	2025-02-02T11:17:04.000Z	cpe076	cpe076	P_070	Male	39
17	cpe107	8e433bf-b266-4cd8c874-db37baad3c61	2025-02-02T11:09:12.000Z	2025-02-02T11:10:35.000Z	cpe107	cpe107	P_107	Male	61
18	cpe007	3e71896e-b18d-486b-adff-bc7e58bbd2b7	2025-02-02T11:09:11.000Z	2025-02-02T11:09:17.000Z	Cpe007	Cpe007	P_022	Male	67
19	cpe003	cc06c007-a8bf-4206-83bd-6e7ea592394b	2025-02-02T11:09:10.000Z	2025-02-02T11:09:17.000Z	cpe003	cpe003	P_015	Male	78
20	cpe073	82e0fe3f-222d-4d53-a2cc-984a6916cb6d	2025-02-02T11:09:00.000Z	2025-02-02T11:11:12.000Z	cpe073	cpe073	P_086	Female	59
21	cpe109	ffa8904b-5db1-4fd0-b392-fd6a8706fb5	2025-02-02T11:08:35.000Z	2025-02-02T11:08:44.000Z	cpe109	cpe109	P_109	Female	56
22	cpe059	7ad23766-263d-40a4-9f8c-4e2fa0b296f0	2025-02-02T11:08:24.000Z	2025-02-02T11:08:58.000Z	cpe059	cpe059	P_018	Female	71
23	cpe034	47da13bc-0b65-4f67-9801-e70b53b70989	2025-02-02T11:08:09.000Z	2025-02-04T07:26:26.000Z	cpe034	cpe034	P_008	Female	73
24	cpe090	4da1d00d-b8a0-4d83-a9e1-c3aa34fac9ea	2025-02-02T11:08:03.000Z	2025-02-02T11:11:18.000Z	cpe090	cpe090	P_091	Female	56
25	cpe001	63706bd-7dc3-489b-91ed-6eeff0684b93	2025-02-02T07:54.15.000Z	2025-02-02T07:54:23.000Z	cpe001	cpe001	P_010	Male	64
26									
27									
28									
29									
30									

Figure 6 Editing EpiCollect entries using Excel

Change the name of column ‘lat_6_LongitudeLatitude’ to ‘Latitude’ and ‘long_6_LongitudeLatitude’ to ‘Longitude’ (Figure 7). Save your changes and close.

D	E	F	G	H	I	L	M
1	at	title	2_Strain_ID	3_Patient_ID	4_Gender	5_Age	
2	02T11:11:56.000Z	cped010	cped010	P_022	Male	67	
3	02T11:11:43.000Z	cpe093	cpe093	P_097	Male	77	13.758131
4	02T11:11:25.000Z	cpe108	cpe108	P_108	Female	67	100.536403
5	02T11:11:19.000Z	Cpe013	Cpe013	P_029	Female	70	13.757591
6	02T11:10:25.000Z	cpe035	cpe035	P_008	Female	73	13.75792
7	02T11:11:19.000Z	cpe004	cpe004	P_016	Female	59	100.536324
8	02T11:10:14.000Z	cpe049	cpe049	P_059	Male	61	13.757539
9	02T11:10:26.000Z	cpe025	cpe025	P_039	Male	63	100.536317
10	02T11:10:11.000Z	cpe070	cpe070	P_084	Male	69	13.757582
11	02T11:13:42.000Z	cpe079	cpe079	P_071	Male	85	100.536378
12	02T11:10:22.000Z	cpe061	cpe061	P_018	Female	71	100.536284
13	02T11:09:42.000Z	cpe058	cpe058	P_019	Female	59	
14	02T11:09:36.000Z	cpe022	cpe022	P_036	Female	78	13.757919
15	02T11:09:44.000Z	cpe110	cpe110	P_110	Male	62	100.536222
16	02T11:10:04.000Z	cpe076	cpe076	P_070	Male	39	
17	02T11:10:35.000Z	cpe107	cpe107	P_107	Male	61	13.757841
18	02T11:09:17.000Z	Cpe007	Cpe007	P_022	Male	67	100.53638
19	02T11:09:17.000Z	cpe003	cpe003	P_015	Male	78	
20	02T11:11:12.000Z	cpe073	cpe073	P_086	Female	59	
21	02T11:08:44.000Z	cpe109	cpe109	P_109	Female	56	13.758121
22	02T11:08:58.000Z	cpe059	cpe059	P_018	Female	71	100.536387
23	04T07:26:26.000Z	cpe034	cpe034	P_008	Female	73	13.757827
24	02T11:11:18.000Z	cpe090	cpe090	P_091	Female	56	100.536176
25	02T07:54.23.000Z	cpe001	cpe001	P_010	Male	64	100.536487
26							
27							
28							
29							
30							

Figure 7 Editing EpiCollect entries using Excel

Data visualization with Microreact

Start by opening a new window in Firefox and typing <https://microreact.org/> in the address bar. Click on “Upload” and browse for the phylogenetic tree (Kpn_ST78.cpe058.rmRCB_iqtree.contree.tree) and metadata files (cpe_cases.epicollect_data.csv) to create a new MicroReact project (Figure 8 and 9).

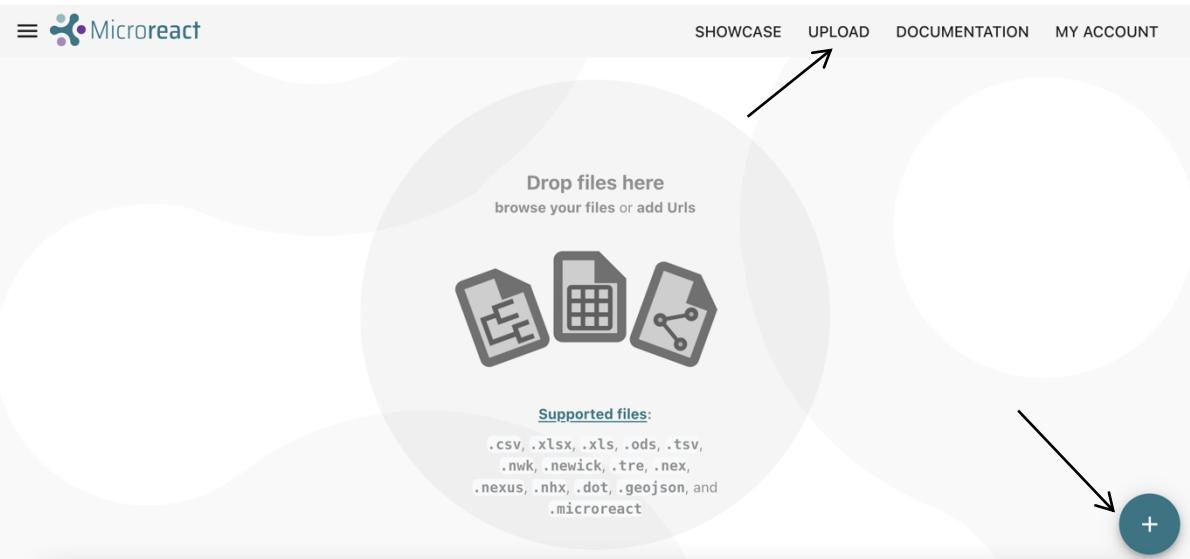


Figure 8 Uploading files on Microreact

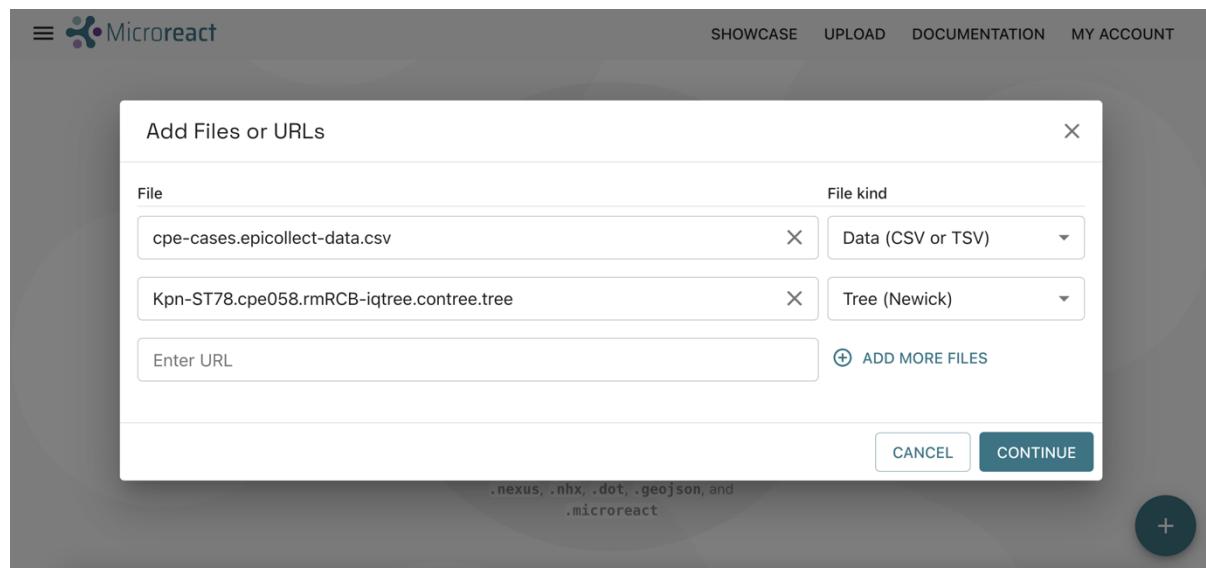


Figure 9 Uploading files on Microreact

Once the tree and metadata files are loaded you will be directed to a new window where files will be automatically detected as Data (CSV or TSV) file (cpe_cases.epicollect_data.csv) and Tree (Newick) file (Kpn_ST78.cpe058.rmRCB_iqtree.contree.tree). In this new window click on ‘Continue’. In the next window, make sure the column ‘barcode’ is selected as the ‘ID Column’ (Figure 10) and then click on ‘Continue’. The ‘ID column’ is the column in the metadata file that must match the strain labels in the phylogenetic tree (i.e. tip or leave labels).

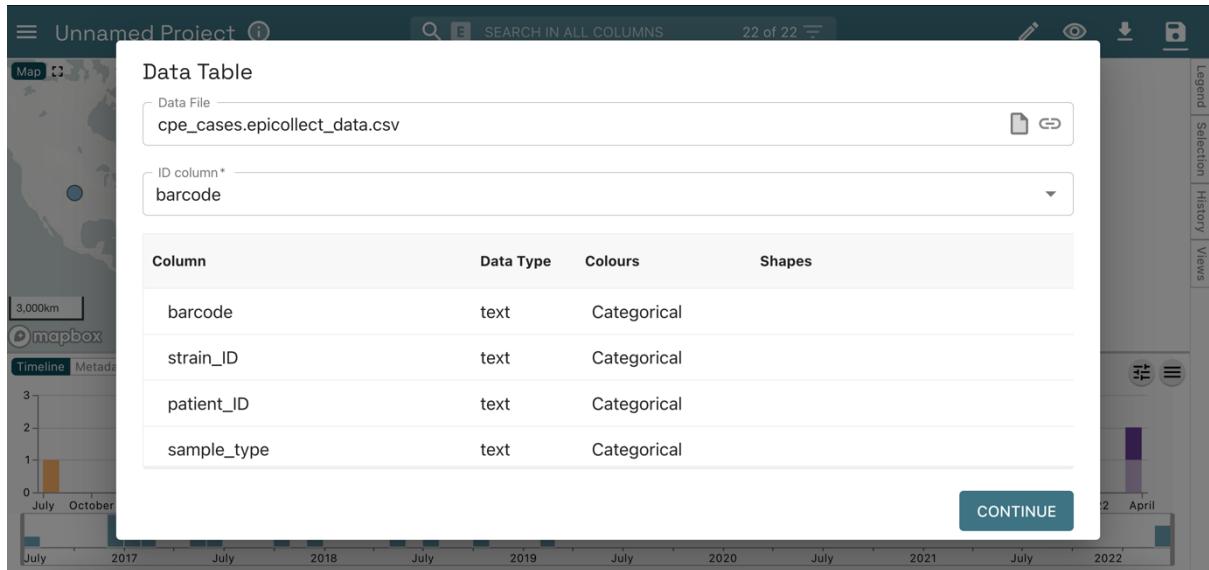


Figure 10 Selecting ID column in Microreact

Once the form is completed your data will be utilized to create a MicroReact project. You should now have a view like the one shown in Figure 11.

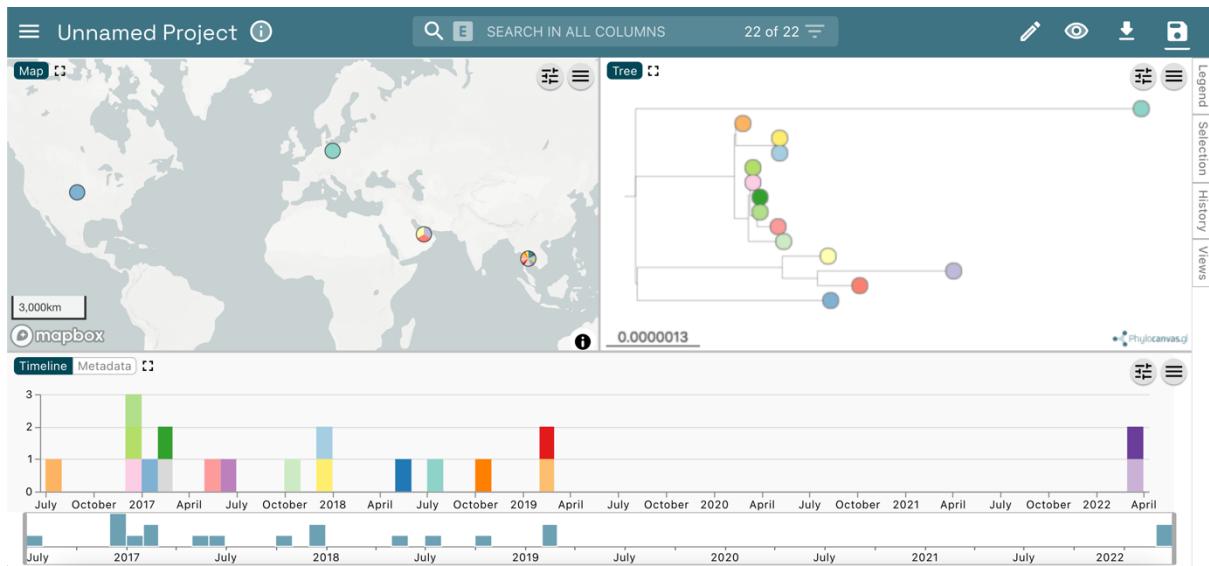


Figure 11 Thailand CPE cases on Microreact

You should see a Map, Tree, and Timeline panels. You can use click-drag-zoom to navigate both the tree and the map.

You can reorient the tree in different layouts to make it easier to view (Figure 12). Click the control panel symbol  then click the tree-view control button  then select the ‘Hierarchical tree’ button (the last option at the bottom).

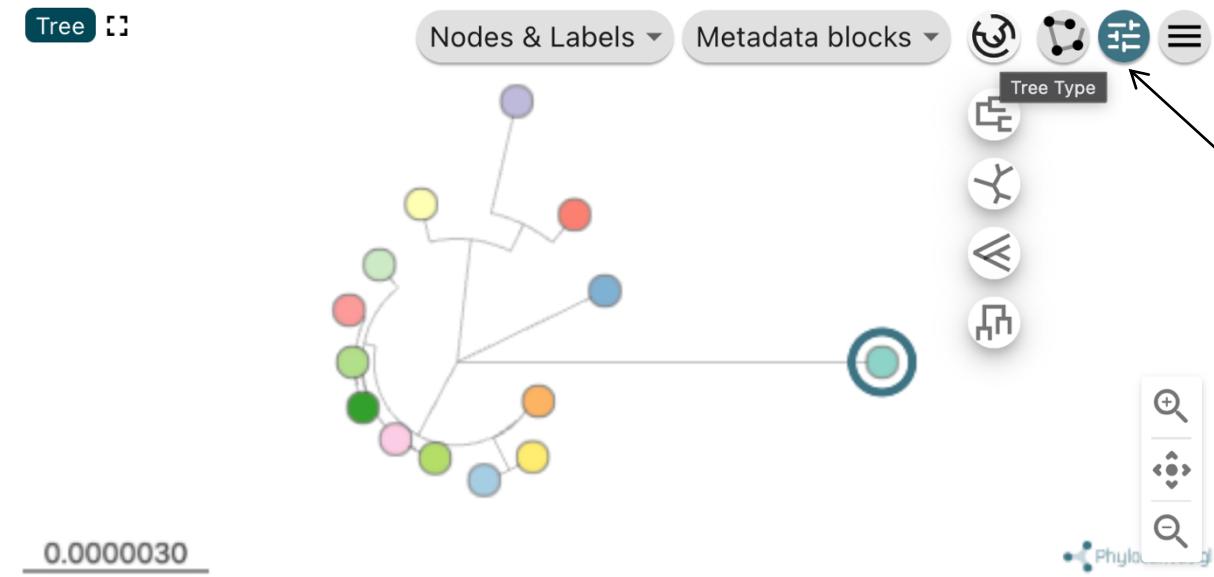


Figure 12 Changing the phylogenetic tree layout on Microreact.

By default, the tree tips should be coloured by strain_ID. You can click on the ‘Legend’ button to view the colour legend.

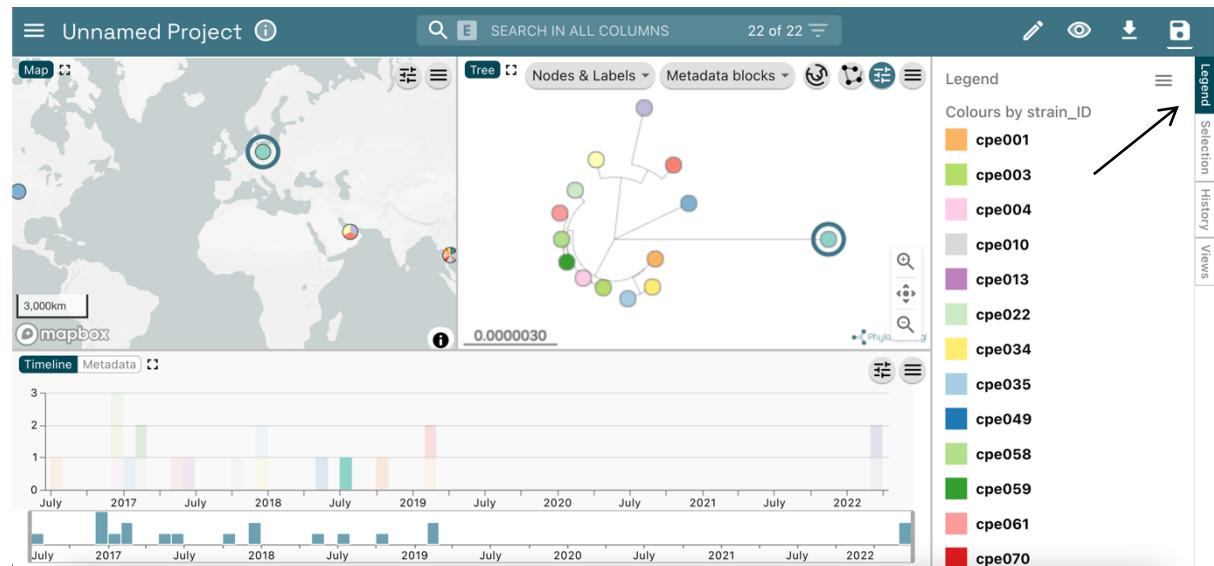


Figure 13 Changing the phylogenetic tree layout on Microreact.

Next, we will also display the strain IDs on the phylogenetic tree. Click on the ‘Labels, Colours, and Shapes’  icon and select ‘strain_ID’ under the lists ‘Labels Column’ and ‘Colour Column’ (arrows in Figure 14). You can also click the ‘Legend’ tab (top right) in the Microreact window to see a legend.

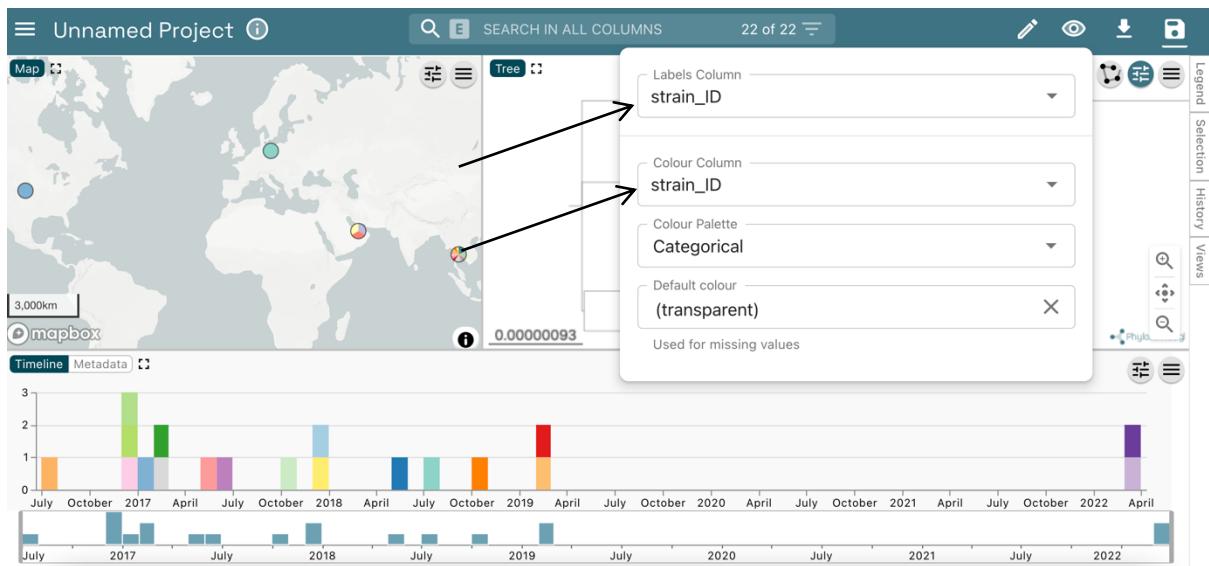


Figure 14 Changing the tip labels and colours.

You should be able to see the clustering of *K. pneumoniae* ST78 isolates with respect to the contextual ST78 isolates (Figure 15).

Question: is there any phylogenetic evidence of a **single-source outbreak** among the *K. pneumoniae* ST78 isolates we are investigating? Or multiple circulating clones? Consider too the SNP distances calculated by *pairsnp* or *MEGA*. Is there any evidence of phylogenetic clustering based on geography?

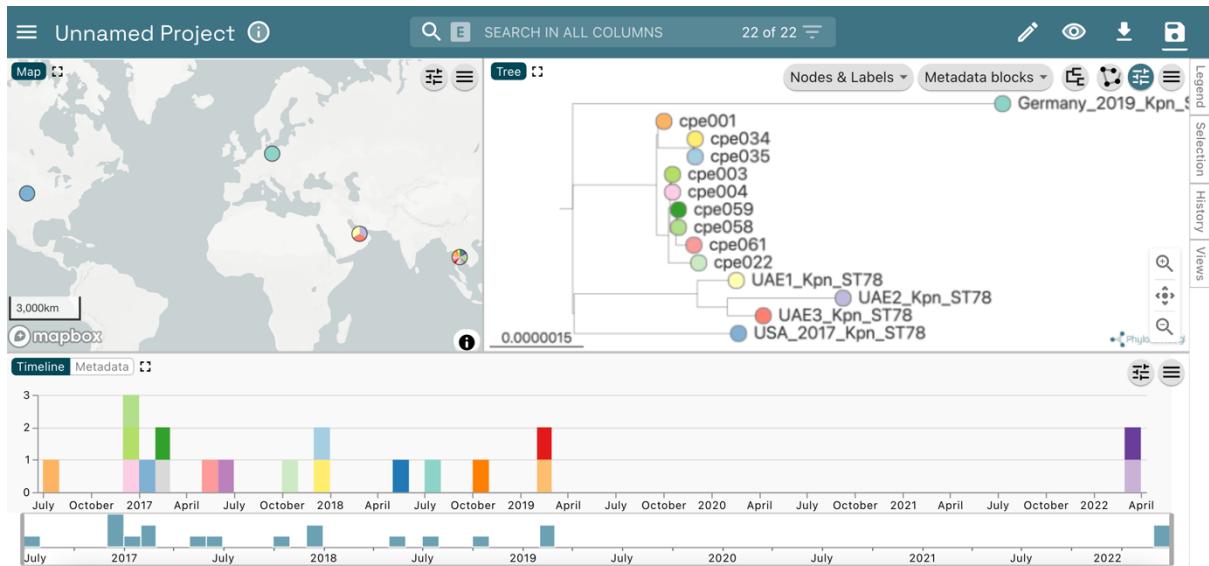


Figure 15 Phylogenetic clustering.

Next, click on the ‘Labels, Colours, and Shapes’ icon and select ‘patient_ID’ under the lists ‘Labels Column’ and ‘Colour Column’ (arrows 1 and 2 in Figure 16). You can also click the ‘Legend’ tab (top right) in the Microreact window to see a legend.

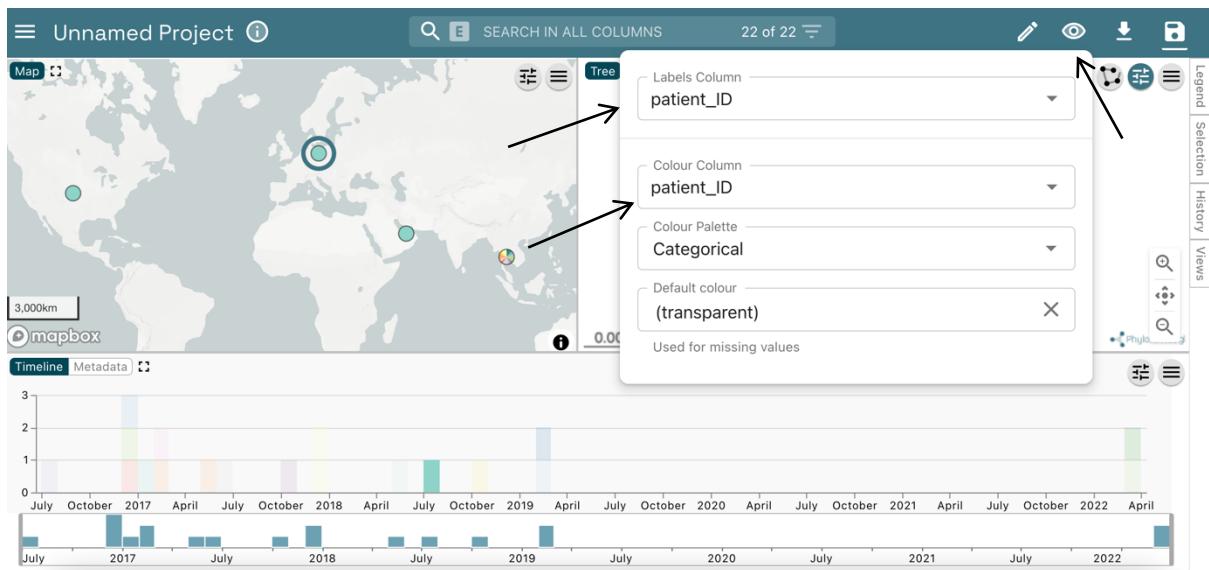


Figure 16 Changing the tip labels and colours.

Click on the ‘Show controls’ button  (Figure 17) to then click the ‘Nodes & Labels’ menu (above the tree) and toggle the ‘Leaf Labels’ button  to switch on labelling (arrow 2). Unselect ‘Align Leaf Labels’ (Figure 17).

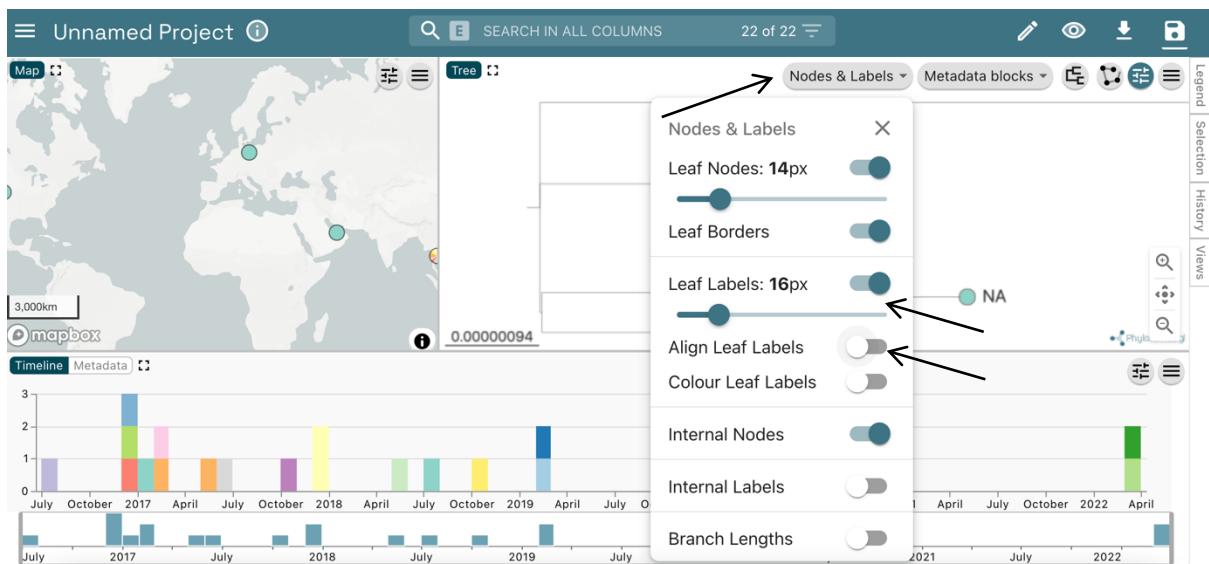


Figure 17 Showing tip labels on the tree.

Question: can you identify any patient with more than one isolate on the phylogenetic tree? (note: ignore the contextual isolates with ‘NA’ patient_ID, i.e., those with not available patient_ID, which happen to have the same colour)

Next, zoom into the clade of CPE outbreak cases and identify the clustering (sub-clades) of patients. Identify which patients have the most ancestral isolates in each of these sub-clades and consider the sampling time of CPE cases (visible in the Timeline panel, Figure 18).



Figure 18 Phylogenetic and temporal clustering of CPE cases.

Question: based on the temporal distribution and phylogenetic clustering of CPE cases, which patients might be among the **early sources of the outbreak?** Which patients could be ruled out as sources of the outbreak? Important note: we cannot infer the origin of an outbreak based solely on phylogenetic and temporal data, we will need epidemiological data too (i.e., hospital contacts, hospital room stays, etc.).

Finally, we will explore the distribution of antibiograms and AMR genetic determinants.

It is possible to view multiple variables against the tree simultaneously. First, click on the ‘Show controls’ button and under ‘Nodes & Labels’ menu to toggle both the ‘Leaf Labels’ and ‘Align Lead Labels’ buttons to switch on and make sure leaf labels are visible and aligned. Next, click the ‘Metadata blocks’ menu button and check all the boxes that relate to phenotypic antibiotic resistance (that is, from amikacin to piperacillin_tazobactam). See Figure 19.

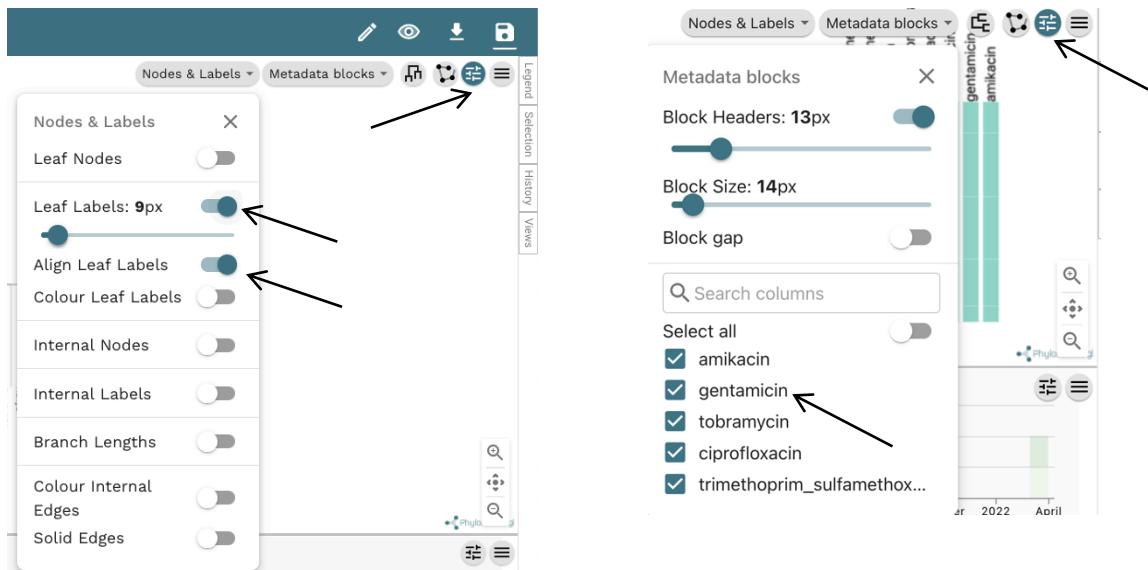


Figure 19 Options to display metadata block.

This will allow us to visualise the entire antibiogram of our local CPE cases and that of contextual strains (Figure 20).

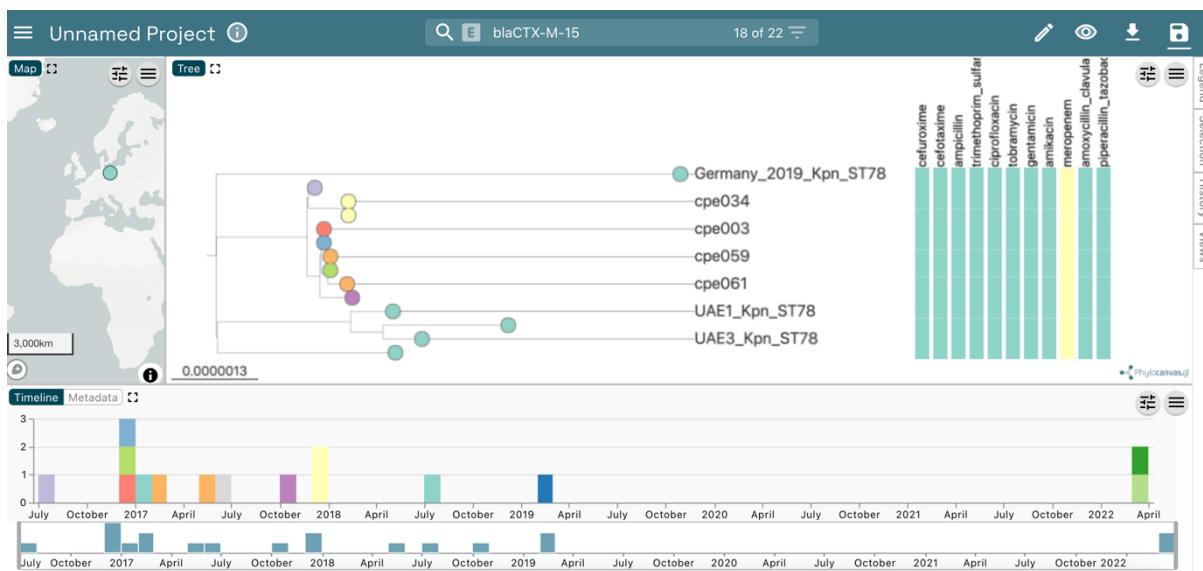


Figure 20 Displaying the antibiogram of CPE cases.

Label the tips of samples with AMR genetic determinants while displaying phenotypic AMR in the metadata blocks. See Figures 21 and 22 as examples.



Figure 21 Displaying aminoglycoside resistance genes and resistance phenotypes.

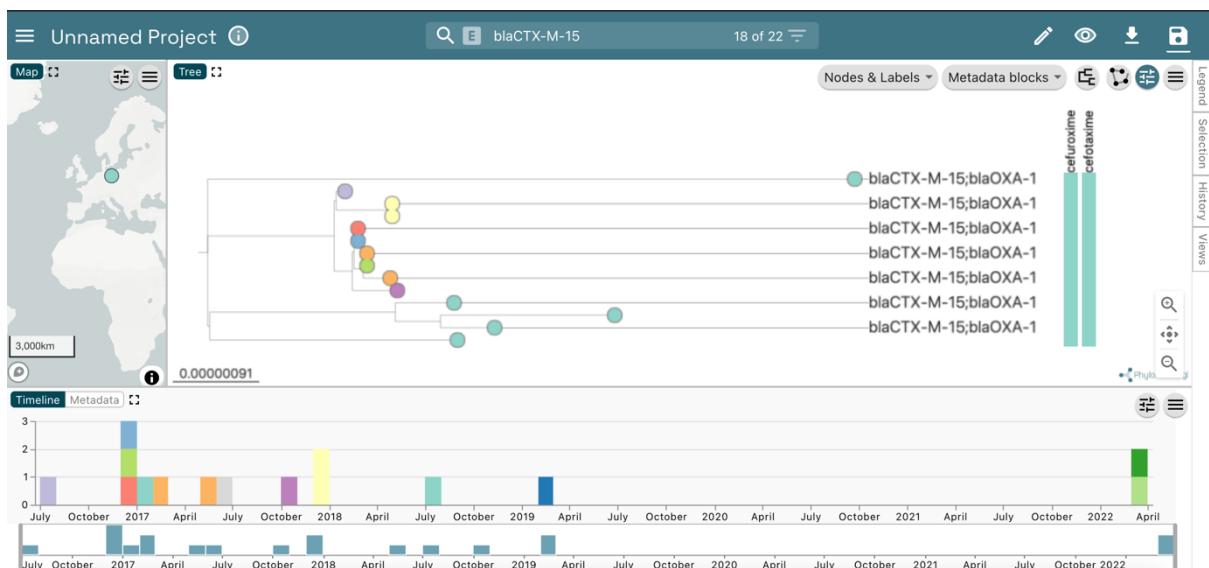


Figure 22 Displaying ESBL resistance genes and resistance phenotypes.

Spend some time exploring the distribution of AMR genes and their relationship with phenotypic resistance.

Question: can you identify any AMR genetic determinants specific to the local CPE cases, and which are shared with the contextual strains?

Exercise 2: Extensively Drug-Resistant *Salmonella enterica* Serovar Typhi in Pakistan

Introduction

Typhoid fever is a bacterial infection caused by the bacterium *Salmonella enterica* serovar Typhi. The bacteria spreads via contaminated food and water, particularly in communities with poor sanitation, and is highly contagious. Symptoms include fever, stomach pain, headache and constipation or diarrhoea. If left untreated, the infection can be fatal. Vaccination, access to clean water, and improved sanitation are effective means to prevent typhoid.

Typhoid can be treated with antibiotics, although antibiotic-resistant *Salmonella* Typhi (*S. Typhi*) strains have become increasingly prevalent. Historically, typhoid has been treated with ampicillin, trimethoprim-sulfamethoxazole, and chloramphenicol.¹⁷ *S. Typhi* strains that are resistant to all these three first-line antibiotics are labelled as multidrug resistant (MDR). Resistance to fluoroquinolones has also been reported in regions where MDR infections are treated with these second-line antibiotics. Ceftriaxone, a third generation cephalosporin, and azithromycin, a macrolide, are reserved to treat multidrug-resistant infections.

In November 2016, an unusually high number of ceftriaxone-resistant cases were identified in the province of Sindh, Pakistan, primarily from the cities of Hyderabad and Karachi. The provincial public health authorities in Pakistan launched an investigation into possible sources and control measures including an emergency vaccination campaign were put in place. Around the same time, scientists at Public Health England identified a strain of typhoid with the same high levels of resistance from an individual in the United Kingdom returning from Pakistan.

The cases were resistant to five antibiotics in total: chloramphenicol, ampicillin, trimethoprim-sulfamethoxazole, fluoroquinolones, and third-generation cephalosporins. Although MDR, quinolone-resistant and sporadic ceftriaxone resistance cases had been reported before in Pakistan, this was the first time such high level of drug resistance was seen in Typhoid cases. The investigators labelled these cases as extensively drug-resistant (XDR). The doctors were left with few antibiotics available to effectively treat the infection. Scientists at the Wellcome Sanger Institute in Cambridge were approached by scientists at Aga Khan University with a request for genetic analysis of samples.¹⁸

Available data

Whole-genome sequencing was carried out on 87 of the XDR *S. Typhi* strains isolated in Sindh, Pakistan, over a 6-month period between November 2016 and March 2017.

Twelve contemporaneous ceftriaxone-susceptible isolates collected from the same region were also sequenced for context. The authors found that all XDR isolates and 11 out of 12 of the contextual (ceftriaxone-sensitive) isolates belonged to the same clade (4.3.1/H58 clade). They constructed a maximum-likelihood phylogenetic tree with the XDR and contextual strains from Pakistan and previously published *S. Typhi* genomes from the same clade (4.3.1/H58 clade). We will make use of this phylogenetic tree along with metadata on the country of origin, year of isolation and antibiotic resistance genetic determinants to determine the origin of these XDR *S. Typhi* cases.

Creating a Microreact project

The phylogeny and metadata files were used to create a MicroReact project (<https://microreact.org/project/xdrtypyphi>). You can make use of this URL to navigate to the project page or follow the instructions below (Figures 1 to 3) to create a project page from scratch.

Start by opening up a new window in Firefox and typing <https://Microreact.org/> in the address bar. Click on “Upload” as show in Figure 1 to create a new project.

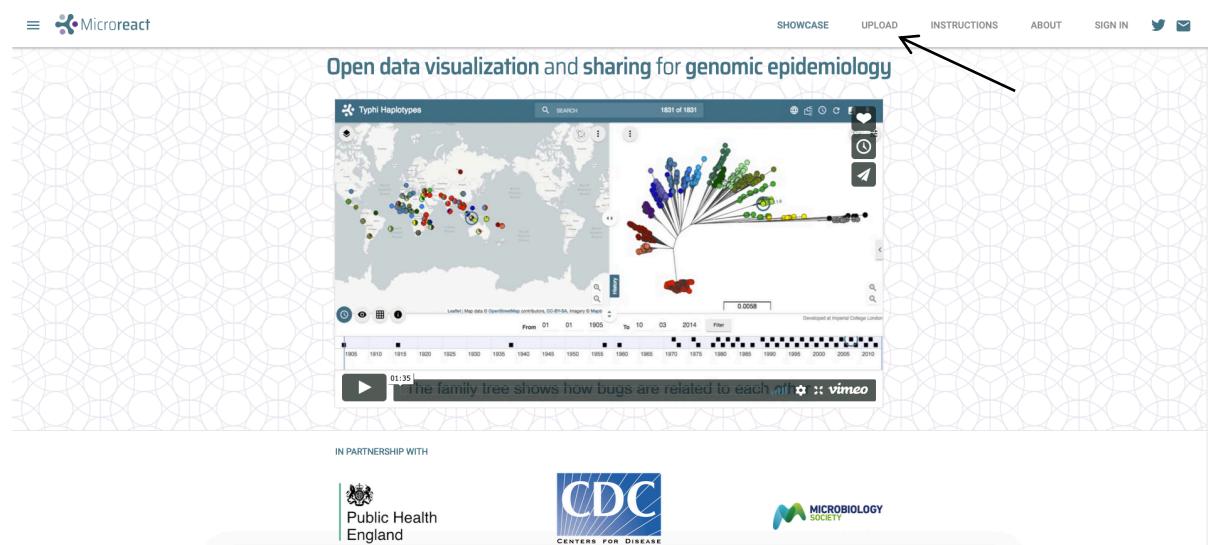


Figure 1 Microreact home page

Drag and drop the files Klemm2018_metadata.csv and Klemm2018_tree.nwk into the browser window.

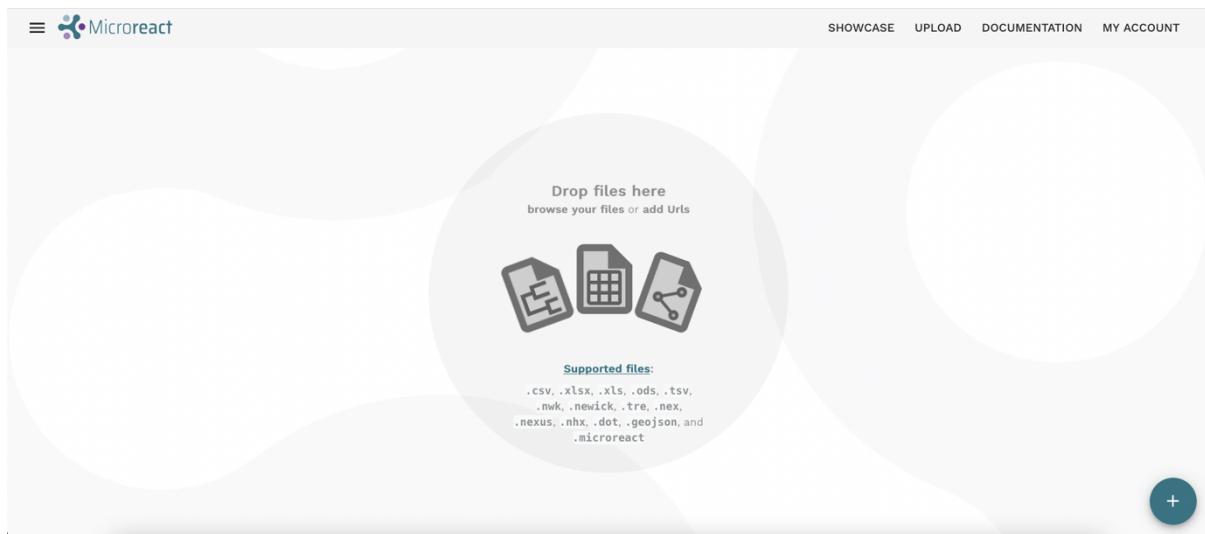


Figure 2 Microreact upload page

The files will start uploading automatically, once done you should see a dialogue box like the one in **Figure 3**, indicating that the files have been correctly identified as a data file (CSV or TSV format) and tree file (newick format). Click Continue.

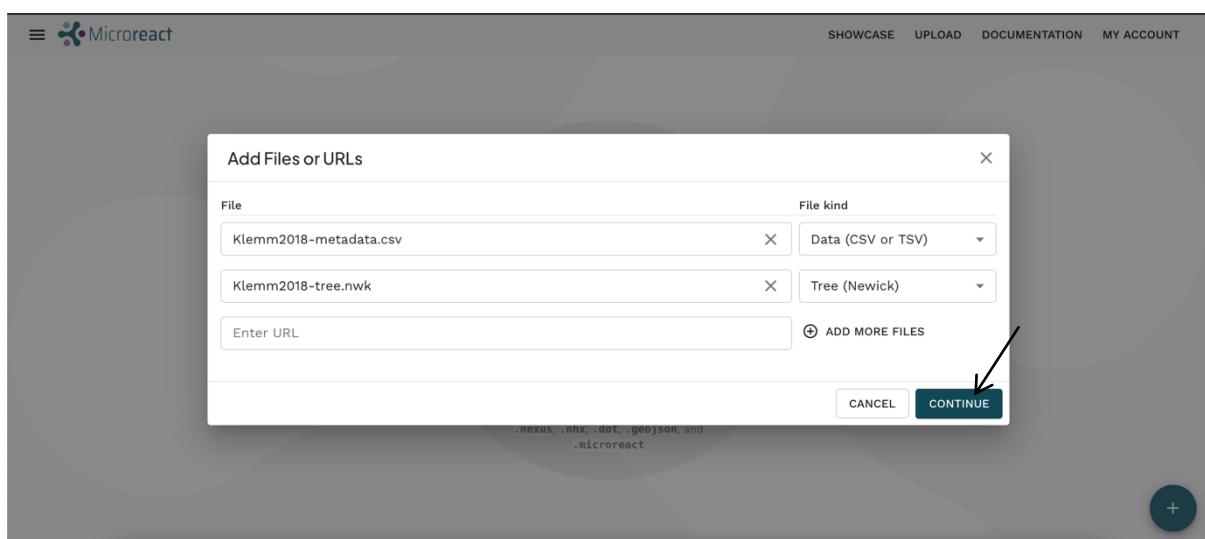


Figure 3 Microreact uploaded files

Once the files are loaded you should see a Microreact setup window (**Figure 4**) asking to indicate which column in the data file contains the 'labels' column – this is to ensure the data can be matched correctly to the tree file. Select the 'Id' column (arrow 1) and click Continue (arrow 2).

Tree

Tree File
Klemm2018_tree.nwk

Labels Column* **1**
Id

Choose the metadata column which contains tree leaf labels

Id	Lane	22420_1_10_P...	22420_1_10	2017	Pakistan	yes	IncY	outbreak	blaCTXN
22420_1_11_P...	22420_1_11	2017	Pakistan	yes	IncY	outbreak	blaCTXN		
22420_1_12_P...	22420_1_12	2017	Pakistan	yes	IncY	outbreak	blaCTXN		
22420_1_13_P...	22420_1_13	2017	Pakistan	yes	IncY	outbreak	blaCTXN		
22420_1_14_P...	22420_1_14	2017	Pakistan	yes	IncY	outbreak	blaCTXN		

CONTINUE

Figure 4 Microreact setup

You should now have a view like the one shown in Figure 5. You can use the save button (top right) to name and save the project for future use if you like.

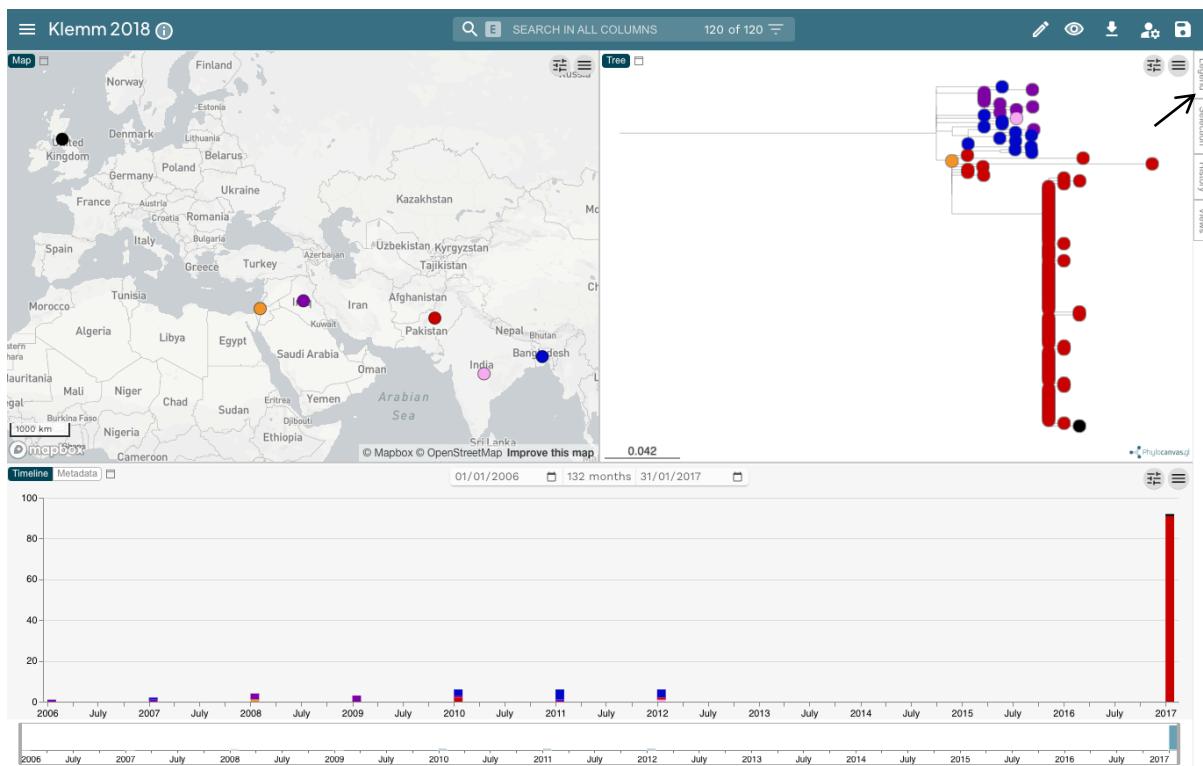


Figure 5 Klemm2018 Microreact project

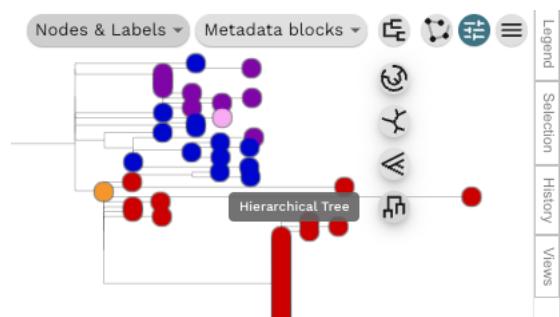
Identifying the origin and spread of extensively drug-resistant *S. Typhi* in Pakistan

We will use the phylogeny and the map to identify the origin of the XDR *S. Typhi* cases in Pakistan.

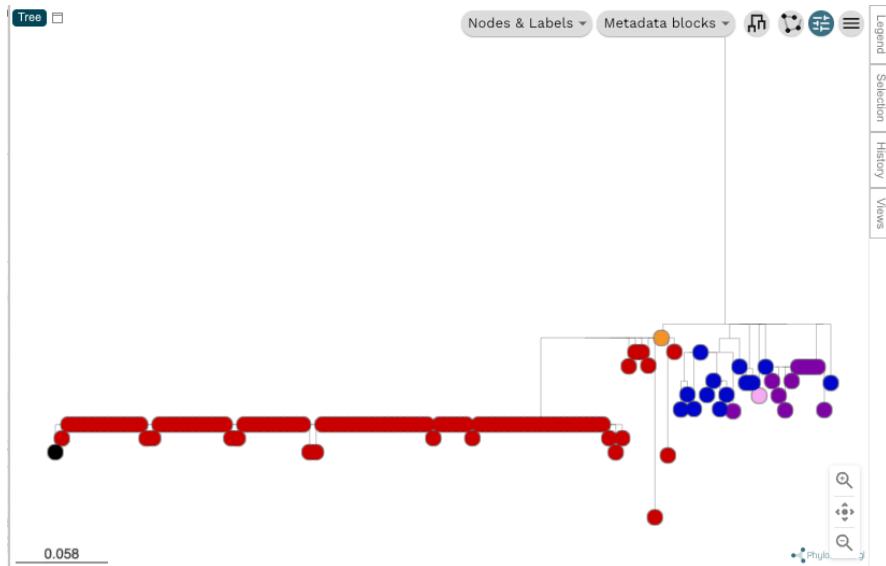
By default, the tree tips should be coloured by Country (as per Figure 5). The map view shows the colours assigned to each country, you can also click the 'Legend' tab (top right) in the Microreact window (arrow in Figure 5) to see a legend. It should be easy to see that most of the isolates from Pakistan (red) cluster together, separately from strains from other countries.

We can reorient the tree horizontally to make it easier to view.

Click the control panel symbol  then click the tree-view control button  then select the 'Hierarchical tree' button (the last option at the bottom)



The tree should now be oriented like this:

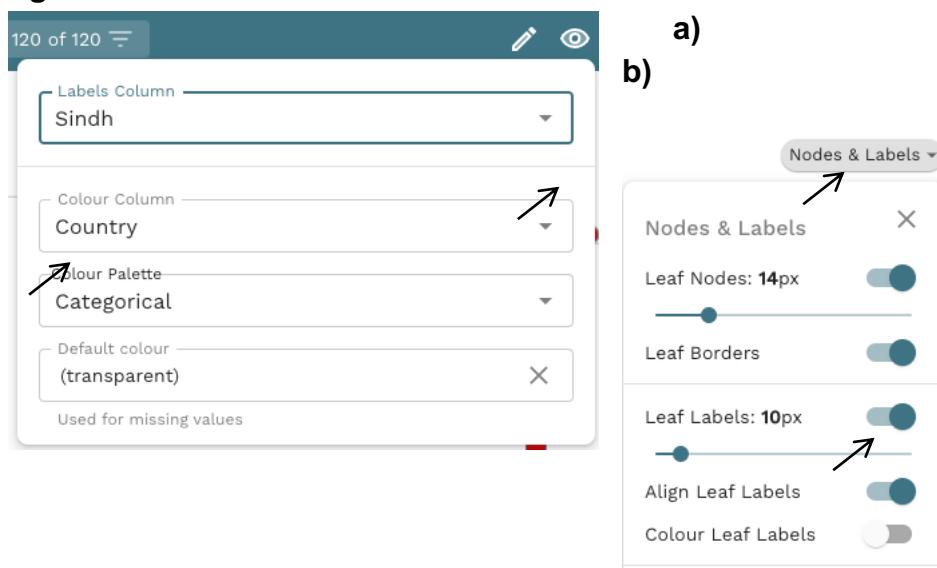


Now let's look at the dates each strain was isolated, and which ones were epidemiologically associated with the suspected outbreak.

First, to label the outbreak strains, click the eye symbol at the top right, and then set the 'Labels Column' to 'Sindh' (Figure 6a).

Then click the 'Nodes & Labels' menu (above the tree) and toggle the 'Leaf Labels' button to switch on labelling (Figure 6b). Change the size to 10px (using the sliding scale) to make it easier to see all the labels in one view.

Figure 6



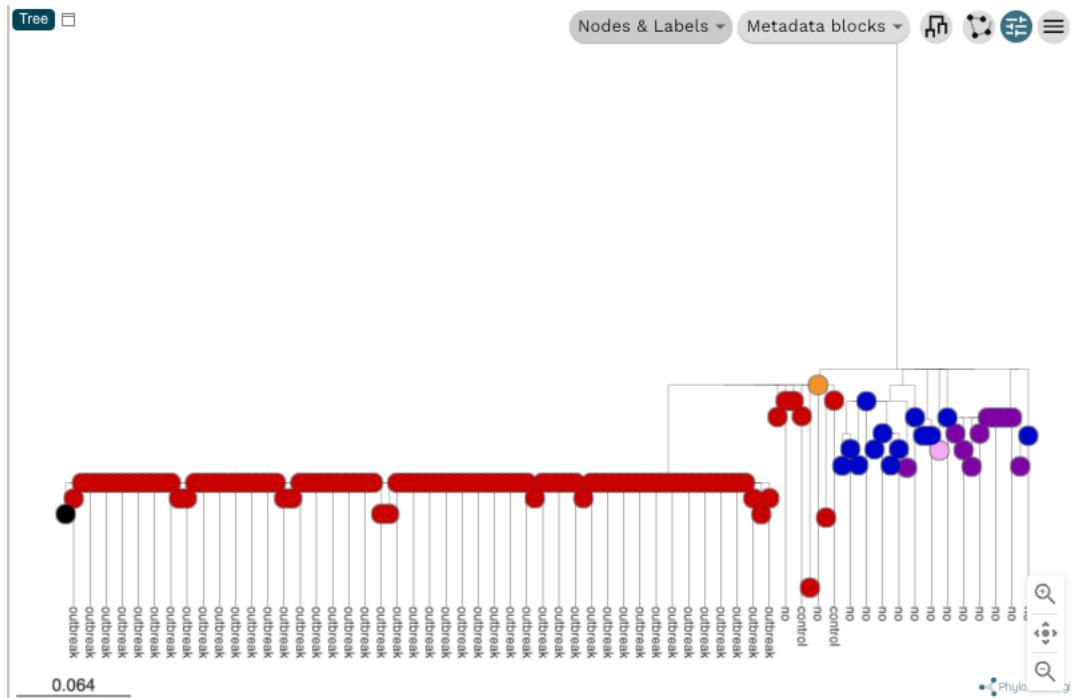


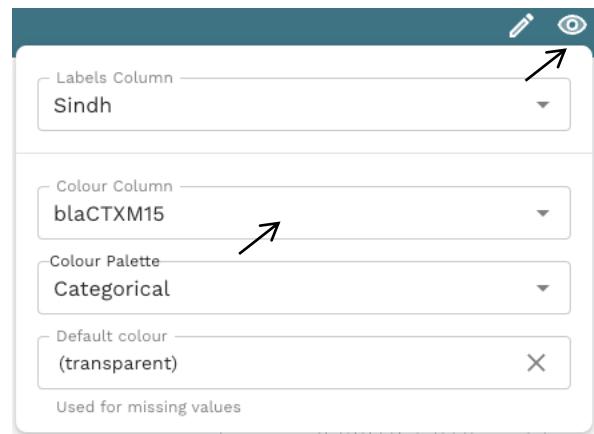
Figure 7 Displaying the tree with hierarchical layout, country of isolation, and outbreak status.

You should now see a view like the one in Figure 7, where it becomes clear that the isolates sequenced from the suspected outbreak in Sindh all form a tight cluster, distinct from the non-outbreak strains from Pakistan and other strains from other countries.

The authors used whole-genome sequencing to build a phylogeny of all the *S. Typhi* strains but also to investigate the genetic determinants of antibiotic resistance in this outbreak clone and in the rest of samples. You can use Microreact to view the distribution of antibiotic resistance genes against the phylogeny, to explore the genetic properties of the outbreak strain in the context of related strains, in two ways.

One option is to colour the tips of the tree by a specific AMR gene of interest. Click the eye symbol top right and change ‘Colour Column’ to a resistance gene – for example, the blaCTXM15 gene, which is responsible for the phenotype of resistance to third-generation cephalosporins such as ceftriaxone.

You should now have a view that looks like Figure 8.



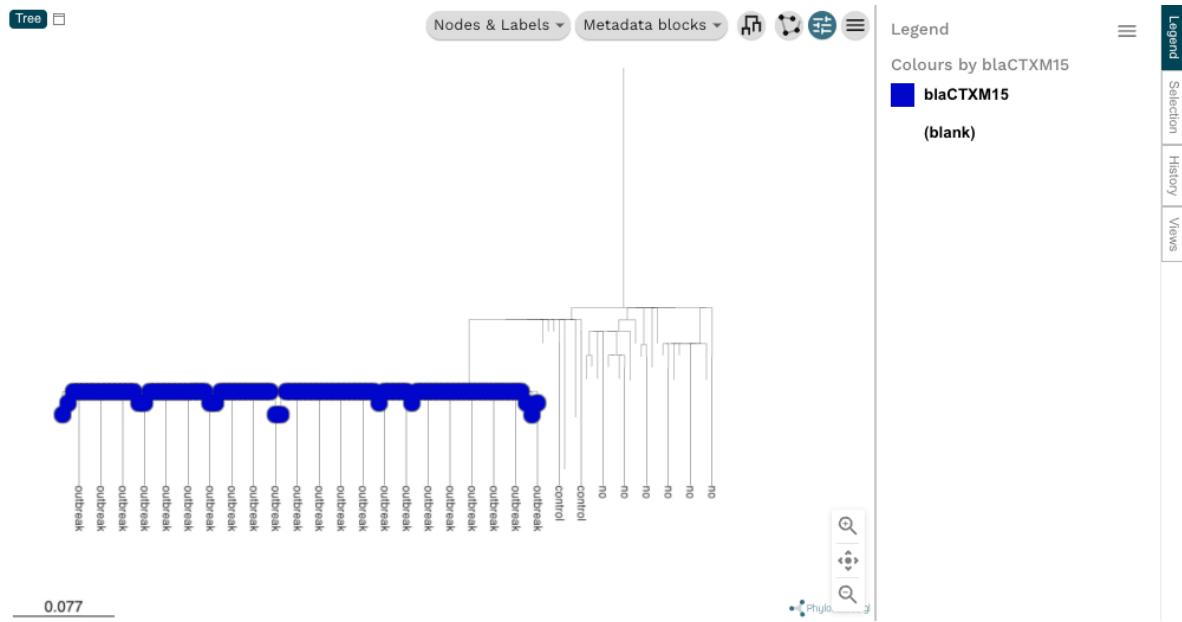


Figure 8 Displaying the tree with a resistance determinant overlaid

You can explore the distribution of other AMR determinants, such as the *qnrS* gene (plasmid-borne fluoroquinolone resistance), *catA1* gene (which confers resistance to chloramphenicol), *blaTEM-1* (ampicillin), *dfrA7* (trimethoprim), *sul1* and *sul2* (sulfamethoxazole), and *strA* and *strB* (streptomycin). Which ones are specific to the outbreak and which are common in other strains?

It is also possible to view multiple variables against the tree simultaneously. Click the ‘Metadata blocks’ menu button and check all the boxes that relate to resistance determinants (from blaCTXM15 down to parE_L416F).

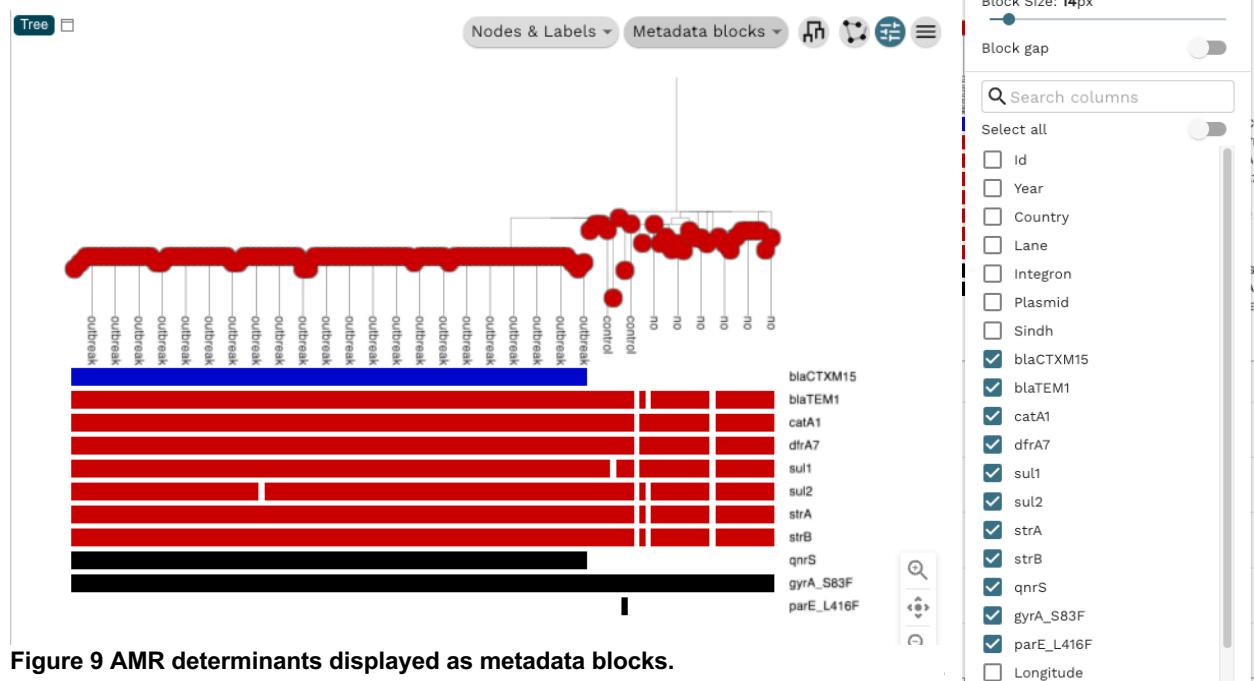


Figure 9 AMR determinants displayed as metadata blocks.

You will notice that almost all samples in the tree (including Sindh outbreak strains) harbour *catA1* gene (chloramphenicol resistance), *blaTEM-1* (ampicillin), *dfrA7* (trimethoprim), *sul1* and *sul2* (sulfamethoxazole), and *strA* and *strB* (streptomycin). All these genes are present in the same composite AMR transposon that is integrated into the chromosome.

The label ‘Integron’ under ‘Label by’ indicates whether samples have this AMR transposon integrated or not. The stability of this AMR cassette (i.e. it is present in almost all samples from the same phylogenetic lineage and rarely lost) is characteristic of chromosomally-integrated transposons which, once integrated into the chromosome, are inherited vertically. All samples also have the S83F mutation in *gyrA* which increases the minimum inhibitory concentration of fluoroquinolones.

The Sindh outbreak clone carries two additional AMR genes: *blaCTX-M-15* which confers resistance to ceftriaxone (a third-generation cephalosporin) and *qnrS* which mediates resistance to ciprofloxacin, both carried on the same IncY plasmid. You can see the presence of the plasmid either by adding ‘Plasmid’ to the list of variables shown in the ‘Metadata blocks’, or by setting the ‘Colour Column’ to ‘Plasmid’ (click on the eye symbol to get this menu).

Now let’s consider what is the most plausible origin of the Sindh XDR outbreak clone. First, set the ‘Colour Column’ back to ‘Country’ (click on the eye symbol to get this menu), to show the country of isolation of all samples on the tree. Second, identify the genetically closest strains to the outbreak clone (circled in Figure 10). You will notice that, except for one sample isolated from Palestine in 2008, these strains were sampled in Pakistan between 2010 and 2017. The fact that the most genetically related strains to the Sindh XDR outbreak clone were also isolated in the same country earlier in time, suggest the outbreak clone derived from an endemic Pakistan clone as opposed to being imported from a different country.

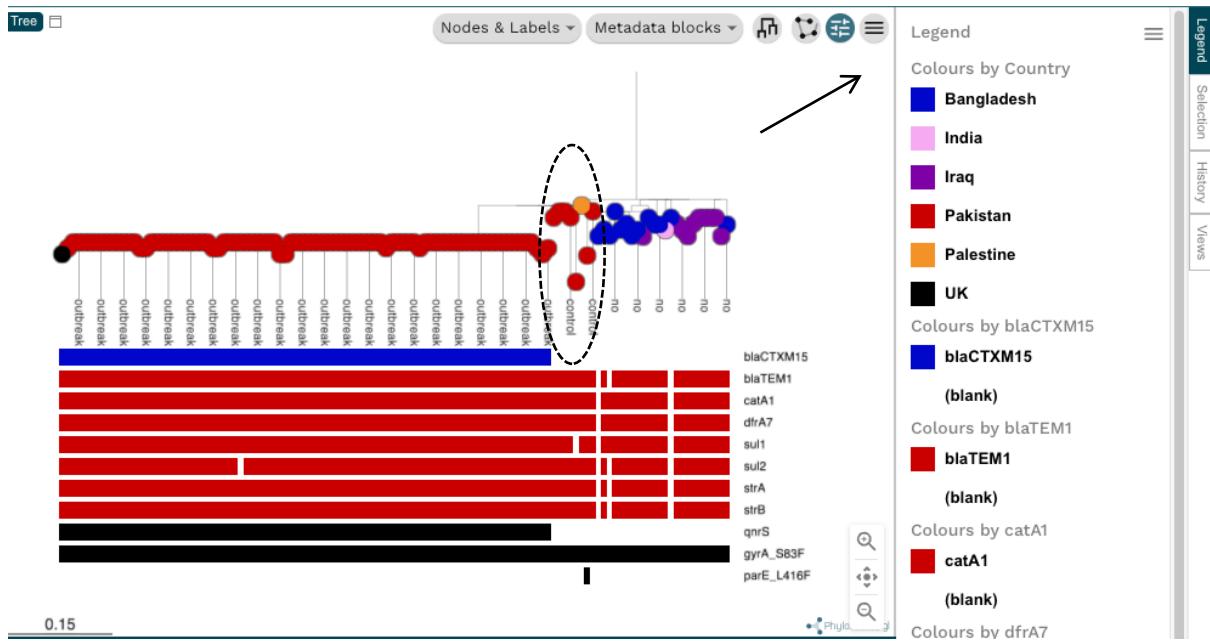


Figure 10 Identifying the origin of the Sindh XDR outbreak clone