



User Manual for the MTB NGS Analysis Platform
for Drug Resistance Interpretation and Lineage Classification
of *Mycobacterium tuberculosis*

1. When accessing the website <https://mtb.nbt.or.th/>, you will see the main page of the National Biobank of Thailand: TB Web Service, as shown in the figure below.

MTB Analysis Platform

TB Landscape User Manual Contact us Login

วัณโรค TUBERCULOSIS

รู้เร็ว รักษาหาย ไม่แพร่กระจาย

Mycobacterium tuberculosis (MTB) causes a major public health problem. With the high prevalence of MTB in Thailand, the World Health Organization (WHO) has historically identified Thailand as one of the **high tuberculosis (TB) burden countries**, a group characterized by a significant number of cases across TB infection, co-infection with HIV (TB/HIV), and multidrug-resistant TB (MDR-TB). The increasing incidence of drug-resistant TB significantly hampers the success of TB control programs due to higher treatment failure rates. Often, newly diagnosed TB patients are prescribed multiple drugs empirically because standard TB drug sensitivity testing can take up to two months. Such practices can inadvertently promote a higher incidence of TB drug resistance. Consequently, the WHO recommends the use of **TB whole-genome sequencing (WGS)** within standard TB control programs, encompassing diagnosis, prediction of TB drug resistance, and TB spread management. WGS offers advantages in terms of speed and cost-effectiveness, making it increasingly a standard practice in public health. However, TB WGS generates large sequencing data files, requiring complex bioinformatic operations that can be challenging for TB interpretation personnel. Addressing this challenge, we are developing a computational platform to assist in accurate TB drug resistance prediction using **Next-Generation Sequencing (NGS)** analysis. The resulting computational prediction of drug-resistant TB aims to promote the construction of an up-to-date national TB genomic portal, offering a prototype for TB drug programs in Thailand.

MTB NGS Analysis Platform

Full scale NGS analysis platform for research support. Including lineage and drug resistant prediction, genome browser to explore your data, phylogenetic tree, outbreak analysis, structural variant analysis etc.

GET START

Logos of various Thai institutions and the BIOTEC NSTDA logo.

2. To access the TB NGS Analysis Platform, scroll down to the **TB NGS Analysis Platform** section and click the **GET START** button.



TB NGS Analysis Platform

Full scale NGS analysis platform for research support. Including lineage and drug resistant prediction, genome browser to explore your data, phylogenetic tree, outbreak analysis, structural variant analysis etc.

GET START

3. After clicking the **GET START** button, you need to register in order to log in and use the TB NGS Analysis Platform. Fill in the required information as displayed and wait for a confirmation email regarding account approval.

Sign Up

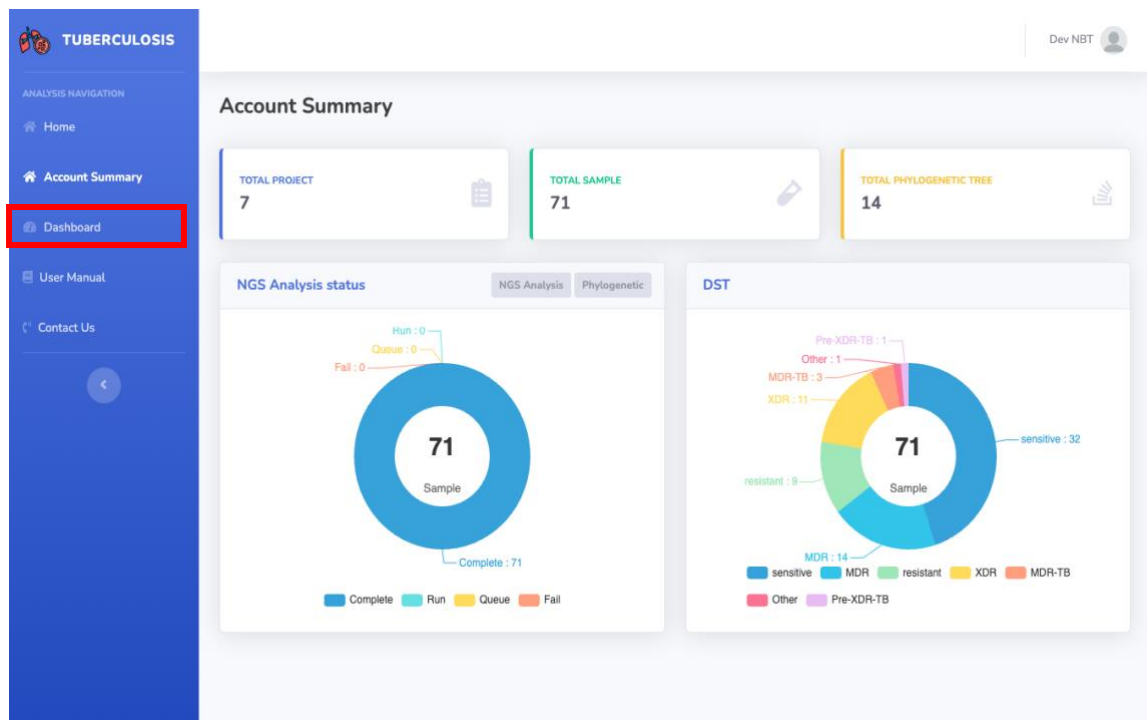
Sign up with your regular account

Sign Up

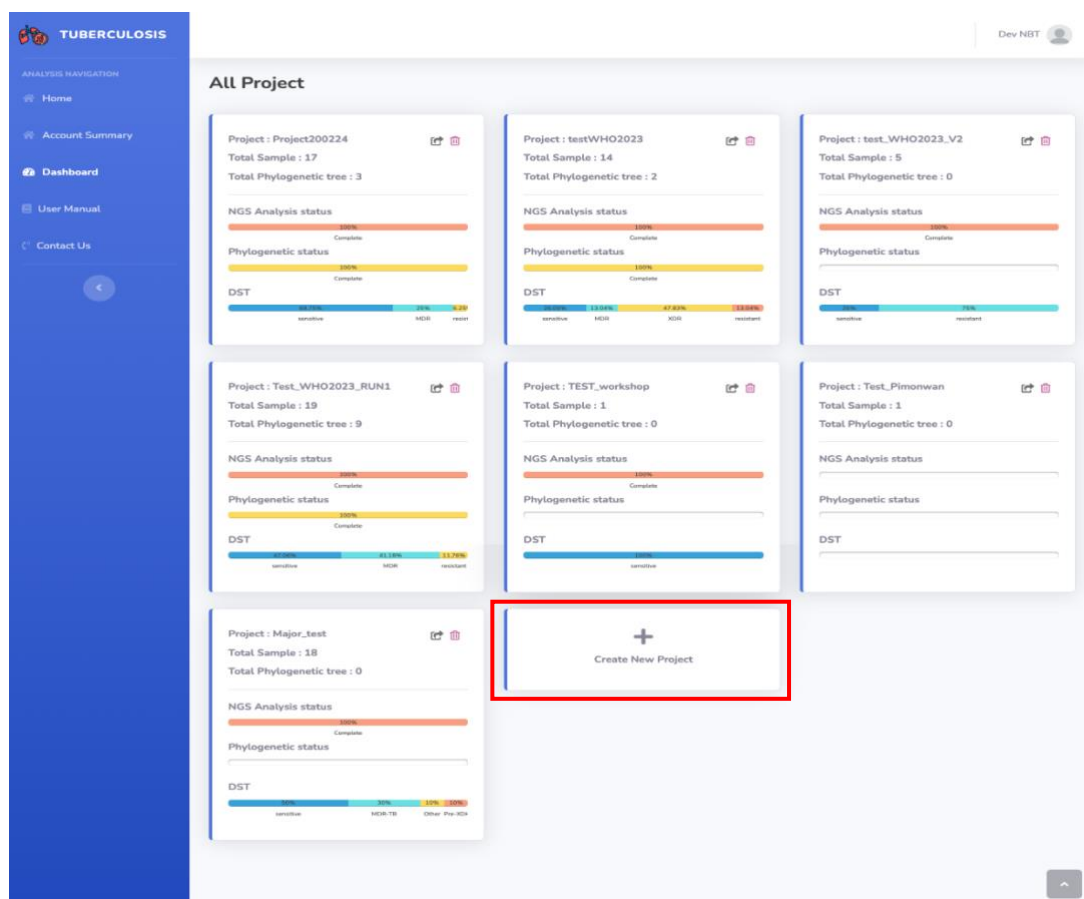
Cancel

Forgot Password?

4. After your account has been approved, log in using the email and password you registered with. The system will then take you to the **Account Summary** page as the first page. Click the **Dashboard** button to begin using the platform.



5. When you click the **Dashboard** button, the system will take you to the **All Projects** page. Then, click **Create New Project**.



Enter a project name in the Project Name field.

After naming the project, click Create Project.

6. when you click **Create Project**, the system will open the project you have created, with the following details:

1. Click Upload Data

2. Select the sequencing platform used for the data you want to

3. The user clicks Choose Files to select the sample files. The file names must follow a specific format, which depends on the sequencing platform you selected (multiple samples can be chosen):

- Illumina/BGI: Paired-end files must end with either `_1.fastq.gz` and `_2.fastq.gz`, or `_1.fq.gz` and `_2.fq.gz`.
- Nanopore: Single-end files must end with either `.fastq.gz` or `.fq.gz`.

4. Click the Submit button to upload the data.

- The specific suffix for paired fastq file can be defined in the advance options.

[Advance Options](#)

If the file names for analysis are not in the specified format, the user can define their own file suffix.

Define file suffix

Use this advance option to define your own specific suffix file. For example you have "sample_r1.fq.gz and sample_r2.fq.gz", your suffix should be "r1.fq.gz and r2.fq.gz" for forward and reverse respectively.

Forward file suffix

Reverse file suffix

Column 3:
District

Choose Platform for
uploaded data

Column 7:
Upload Status

Column 8:
Date of sample collection

Column 9:
Date of data upload

Show 10 entries Sequence Platform

BGI

Search:

☒	Sample Name	Platform	District	Province	Country	Analysis Status	Upload Status	Collected Date	Data Upload Date	Delete
☒	SRR19857396	BGI	-	Kanchanaburi	Thailand	Analysed	Upload Completed	2007-05-09	01-09-2025	
☒	SRR19857395	BGI	-	Bangkok	Thailand	Analysed	Upload Completed	2007-08-09	01-09-2025	
☒	SRR19857394	BGI	-	Roi Et	Thailand	Analysed	Upload Completed	2007-09-18	01-09-2025	
☒	SRR19857392	BGI	-	Kanchanaburi	Thailand	Analysed	Upload Completed	2007-11-07	01-09-2025	
☒	SRR19857393	BGI	-	Bangkok	Thailand	Analysed	Upload Completed	2007-09-22	01-09-2025	

Showing 1 to 5 of 5 entries

Previous 1 Next

Column 1:
Sample Name

Column 5:
Country

Upload Data

Edited metadata

Submit

Column 2:

Platform of the data chosen by the user

Column 4:
Province

Column 6: Analysis Status
(shows the status of data analysis)

5. Upload metadata via the "Edited metadata" button.

(Before uploading metadata, check the box of sample that you want to use for analysis.)

To upload metadata, use the Edited metadata. It will prompt you to download a template file in .xlsx format.

Upload Meta Data

Click the **Download template** button to get the required file.

We accept only ".xlsx" file that follow our template provided here: [Download](#)

****Cautions: The location name must be correctly type. You can check the correctness of spelling from this file: [Download](#)**

Choose File

Once you've filled in the location names in the template, click **Save**. Then, click **Choose File** to select the completed template and upload it.

Warning: To prevent incorrect location names (e.g., typos), you can't manually type the location. You must use the provided dropdown menu in the template to ensure the location name is correct.

Example template for Upload Metadata

A	B	C	D	E	F	G	H	I	J	K
sampleName	platform	district	province	country	Collected Date	History of tuberculosis	Previous treatments	Chest X-ray results	Current medications	Workplace
SRR31697453	Nanopore	-	Bangkok	Thailand	2023	-	-	-	-	-

If the user doesn't know the information, please use " - "

If you have uploaded data from multiple platforms, you can filter to select only the data you want to use for analysis.

You can also **search** for a specific sample in this table.

Data Store **NGS Analysis** Phylogenetic

Show 10 entries Sequence Platform BGI

	Sample Name %	Platform %	District %	Province %	Country %	Analysis Status %	Upload Status %	Collected Date %	Data Upload Date %	Delete
☒	SRR19857396	BGI	-	Kanchanaburi	Thailand	Analysised	Upload Completed	2007-05-09	01-09-2025	
☒	SRR19857395	BGI	-	Bangkok	Thailand	Analysised	Upload Completed	2007-08-09	01-09-2025	
☒	SRR19857394	BGI	-	Roi Et	Thailand	Analysised	Upload Completed	2007-09-18	01-09-2025	
☒	SRR19857392	BGI	-	Kanchanaburi	Thailand	Analysised	Upload Completed	2007-11-07	01-09-2025	
☒	SRR19857393	BGI	-	Bangkok	Thailand	Analysised	Upload Completed	2007-09-22	01-09-2025	

Showing 1 to 5 of 5 entries

Previous 1 Next

6. After uploading your data, please select the data you want to analyze.

7. Once you have selected the data you want to analyze, click **Submit** to start the analysis process.

When you click **Submit**, the system will take you to the **Submission Preparation** page, which includes the following details:

Analysis setting illumina® [Back](#)

Sample Select : SRR5266540 SRR34323143 SRR34323142 ERR718193

Drug Resistance mutation database

- ☒ TBDB (2024-05-17)
- ☐ TBDB2023
- ☐ WHO2021
- ☐ WHO2023
- ☐ Custom database

Lineage mutation database (Optional)

- ☒ TBprofiler (2024-03-15)
- ☐ Thai SNP DB
- ☐ Napier Barcode
- ☐ Custom database

Proceed

Users can select the desired drug resistance mutation database from five options:

- TBDB (2024-05-17)
- TBDB2023
- WHO2021
- WHO2023
- Custom database (users can upload their own file)

Users can select a Lineage mutation database (Optional) from four options:

- TBprofiler (2024-03-15)
- Thai SNP DB
- Napier Barcode
- Custom database (users can upload their own file)

8. Once the settings are configured, click the Submit button.

7. When you click Submit, the system will take you to the NGS Analysis page and wait for the processing to finish.

Data Store **NGS Analysis** **Phylogenetic**

Show 10 entries Search:

Lineage: All Sub-Lineage: All Drug profile: All Drug database: All Lineage database: All

Sample ID	Status	Lineage	Sub-lineage	DR type	Detail	DR db	Lin db	Date Analysed	Platform
ERR718197	Running	Processing...	Processing...	Processing...	50% HaplotypeCaller	TBprofiler	Thai SNP DB	29-11-2021	Illumina

Showing 1 to 1 of 1 entries

Previous **1** Next

Displays the status of the data analysis.

Once the data analysis is complete, this column will change to a **Full Report** button.

This section displays the platform for the **Drug Resistance mutation database** and the **Lineage mutation database** that the user selected.

8. Once processing is complete, the system will show the analysis results for your sample, including the lineage, sub-lineage, and drug resistance profile.

Sample ID	Analysis status	Lineage	Sub-lineage	Drug resistance	DR DB	Lineage DB	Analyzed date	Seq platform	
SRR19857396	Full Report	lineage2	lineage2.2.1	Pre-XDR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857395	Full Report	lineage1	lineage1.2.1.2.1	RR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857394	Full Report	lineage1	lineage1.1.1	MDR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857392	Full Report	lineage2	lineage2.2.1.1	RR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857393	Full Report	lineage1	lineage1.1.1	MDR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	

Specify what lineage the analyzed sample belongs to and what type of drug resistance it has.

9. Click the Full Report button on the NGS Analysis page to go to the Results page, where you can view the analysis report and download a PDF file with the following details:

Sample ID	Analysis status	Lineage	Sub-lineage	Drug resistance	DR DB	Lineage DB	Analyzed date	Seq platform	
SRR19857396	Full Report	lineage2	lineage2.2.1	Pre-XDR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857395	Full Report	lineage1	lineage1.2.1.2.1	RR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857394	Full Report	lineage1	lineage1.1.1	MDR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857392	Full Report	lineage2	lineage2.2.1.1	RR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	
SRR19857393	Full Report	lineage1	lineage1.1.1	MDR-TB	TBDB_2024-05-17	TBprofiler	01-09-2025	BGI	

3. Click the IGV button to view the Genome browser.

Result

Fast QC ▾
IGV
Report ▾
Download

FastQC 1
FastQC 2

Summary

✓ Basic Statistics

✓ Per base sequence quality

✓ Per tile sequence quality

✓ Per sequence quality scores

✓ Per base sequence content

✓ Per sequence GC content

✓ Per base N content

⚠ Sequence Length Distribution

✓ Sequence Duplication Levels

✓ Overrepresented sequences

✓ Adapter Content

2. You can select the sequence to view.

5. You can download .bam and .vcf files from here.

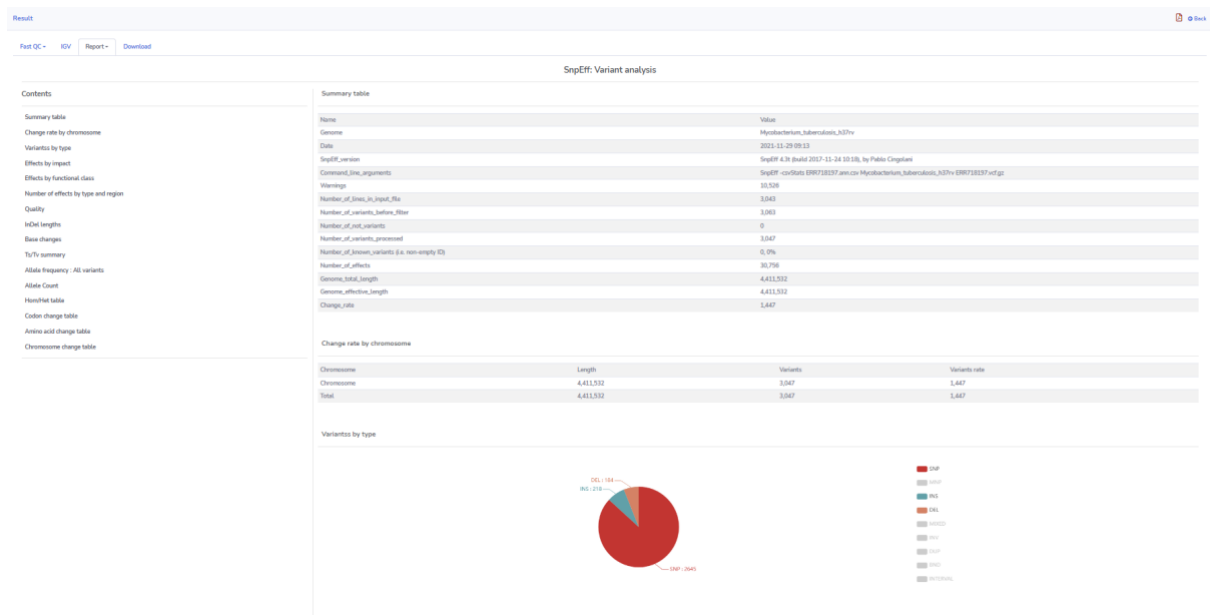
4. The user can select the type of Report they want.

There are three data types to choose from:

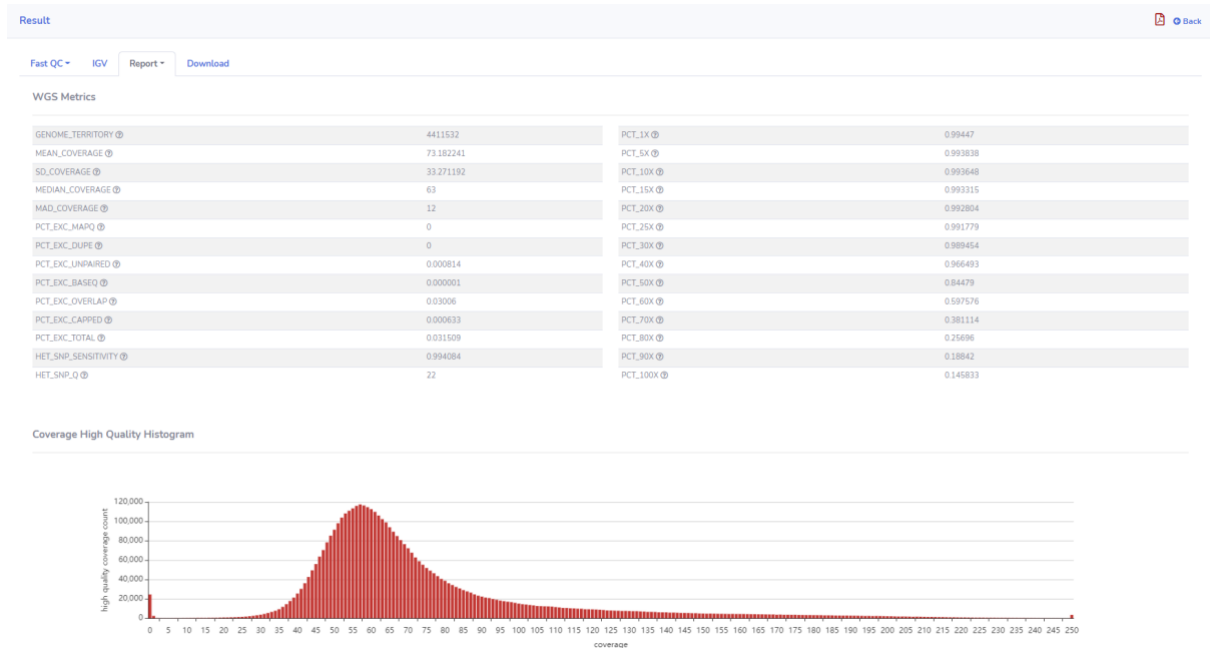
1. Lineage and Drug-Resistant Profile
2. Variations Profile (Summary data)
3. WGS Metrics



Example of a Genome Browser page, where you can view the position of various genes as needed.



Example of a Variations Profile, which shows a summary of the data.



Example of WGS Metrics, which provides a summary of the alignment data.

10. Click the **Phylogenetic** tab to create a phylogenetic tree with the following details:

Create Tree

1. Click the Create Tree button.

2. Name your tree in the **Tree name** field.

	Sample Name	Lineage	Sub-lineage	Drug profile	DR db	Lin db	District	Province	Country	Collected Date
☒	SRR19857396	lineage2	lineage2.2.1	Pre-XDR-TB	TBDB_2024-05-17	TBprofiler	-	Kanchanaburi	Thailand	2007-05-09
☒	SRR19857395	lineage1	lineage1.2.1.2.1	RR-TB	TBDB_2024-05-17	TBprofiler	-	Bangkok	Thailand	2007-08-09
☒	SRR19857394	lineage1	lineage1.1.1	MDR-TB	TBDB_2024-05-17	TBprofiler	-	Roi Et	Thailand	2007-09-18
☒	SRR19857393	lineage1	lineage1.1.1	MDR-TB	TBDB_2024-05-17	TBprofiler	-	Bangkok	Thailand	2007-09-22
☐	SRR19857392	lineage2	lineage2.2.1.1	RR-TB	TBDB_2024-05-17	TBprofiler	-	Kanchanaburi	Thailand	2007-11-07

Submit Phylogenetic data by selecting your desired platform (Illumina/BGI or Nanopore).

3. Check the boxes next to the data you want to use. **Note:** You can only select data that has already completed the NGS Analysis process.

4. Select the Phylogenetic Workflow type from the three available options:

- Phylogenetic tree + Nextstrain (You can select samples with collected date, country, and other metadata).
- Phylogenetic tree
- Placing tree (This is for creating a Phylogenetic tree to compare a user's samples with another tree that was created previously.)

5. Once you have made all your selections, click Submit to start creating the phylogenetic tree.

Note:

For Phylogenetic workflow options 1 and 2, you must select at least 4 samples. The system will not allow you to submit if you choose fewer than 4 samples.

11. For the Placing tree option, you can submit 1 or more samples. Once you click Submit, the system will begin processing with the following details:

Shows the number of samples selected by the user.

The name of the tree set by the user.

Shows the processing status.

Users who chose the **Phylogenetic tree** option will have a ✓ in this column.

Users who did not choose the **Placing tree** option will have a “-” in this column.

Specifies the date (day/month/year) the tree was created.

Tree ID	Status	Progress	Number of Sample	Back Bone Tree	Placing Tree	Date Create	Edit tree
tree 1	Running	100%	4	✓	-	2021-11-30	Previous 1 Next Create Tree
tree_test7	Complete	View Phylogenetic Tree Nextstrain Clonal	10	✗	-	2023-07-20	⚙️ 🗑️
placing02	Complete	View Phylogenetic Tree Nextstrain Clonal	1	✗	test_tree8	2023-07-19	⚙️ 🗑️
test_tree8	Complete	View Phylogenetic Tree Nextstrain Clonal	5	✗	-	2023-07-19	⚙️ 🗑️
placing	Complete	View Phylogenetic Tree Nextstrain Clonal	1	✗	Thai backbone tree	2023-07-19	⚙️ 🗑️
tree_test6	Complete	View Phylogenetic Tree Nextstrain Clonal	5	✗	-	2023-07-19	⚙️ 🗑️
test_tree5	Complete	View Phylogenetic Tree Nextstrain Clonal	6	✗	-	2023-07-19	⚙️ 🗑️
test_tree3	Complete	View Phylogenetic Tree Nextstrain Clonal	6	✗	-	2023-07-19	⚙️ 🗑️
test_tree2	Complete	View Phylogenetic Tree Nextstrain Clonal	5	✗	-	2023-07-19	⚙️ 🗑️
tree_test	Complete	View Phylogenetic Tree Nextstrain Clonal	4	✗	-	2023-07-19	⚙️ 🗑️

The processing status will show as **Complete**.

Once the system has finished processing, the **Progress** column will display three buttons for viewing the reports:

- View Phylogenetic tree
- Nextstrain
- Clonal

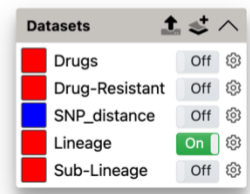
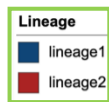
When the system finishes processing, the Status field will show "Complete."

11.1 When you click the **View Phylogenetic tree** button, the system will take you to the Phylogenetic tree page.

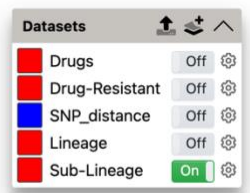
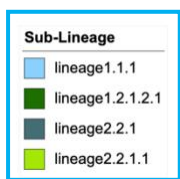
The screenshot shows the ITOL interface with a phylogenetic tree. An orange box highlights the tree's appearance, with a text box stating: "The user can modify the Phylogenetic tree's appearance as needed using the tools provided by the system." Another orange box highlights the 'Control panel' on the right, which includes tabs for Basic, Advanced, Datasets, and Export. The 'Advanced' tab is selected, showing options for Mode (Circular, Rectangular, Unrooted), Rotation (0°), Branch lengths (Use, Ignore), Invert tree (Yes, No), Slanted (Yes, No), and Labels (Display, Hide). The 'Label options' section shows Font (Arial), Font style (20 px), Position (Aligned), Alignment (Left), and Rotation (On). A 'Reset tree' button is also present. A green box highlights the 'Datasets' panel on the right, which lists various data types with checkboxes: Drugs (Off), Drug-Resistant (Off), SNP_distance (Off), Lineage (Off), and Sub-Lineage (Off). A text box states: "This data is an annotation for the sample data. The user can click to enable this data."

09/09/68

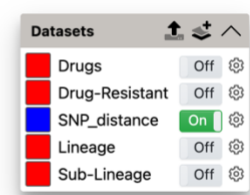
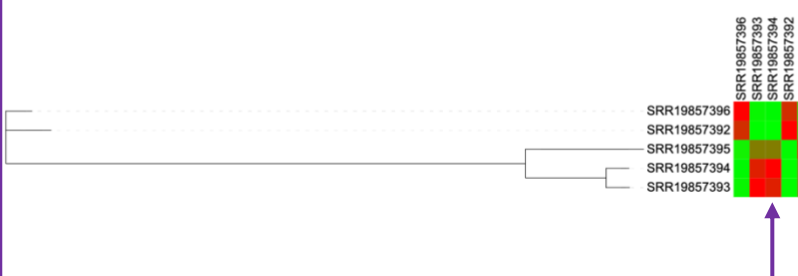
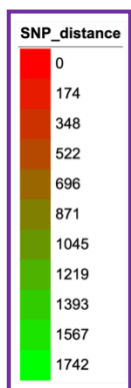
The screenshot shows the ITOL interface with a phylogenetic tree. An orange box highlights the drug resistance data, with a text box stating: "Displays drug resistance data. If the pathogen has resistance to a specific drug, the corresponding circle will turn black." A red box highlights the 'Datasets' panel on the right, which lists various data types with checkboxes: Drugs (On), Drug-Resistant (Off), SNP_distance (Off), Lineage (Off), and Sub-Lineage (Off). A text box states: "Displays drug resistance data using a color-coded bar. The data will be shown as Sensitive, Drug-resistant, MDR, and XDR." A legend titled 'Drug-Resistant' shows color-coded bars: Sensitive (green), RR-TB (light blue), HR-TB (dark blue), MDR-TB (purple), Pre-XDR-TB (orange), XDR-TB (red), and Other (black). The phylogenetic tree shows the same five samples as the previous screenshot, with color-coded bars indicating their drug resistance status.



Displays lineage data using a color-coded bar, allowing you to compare the sample's lineage by its color.

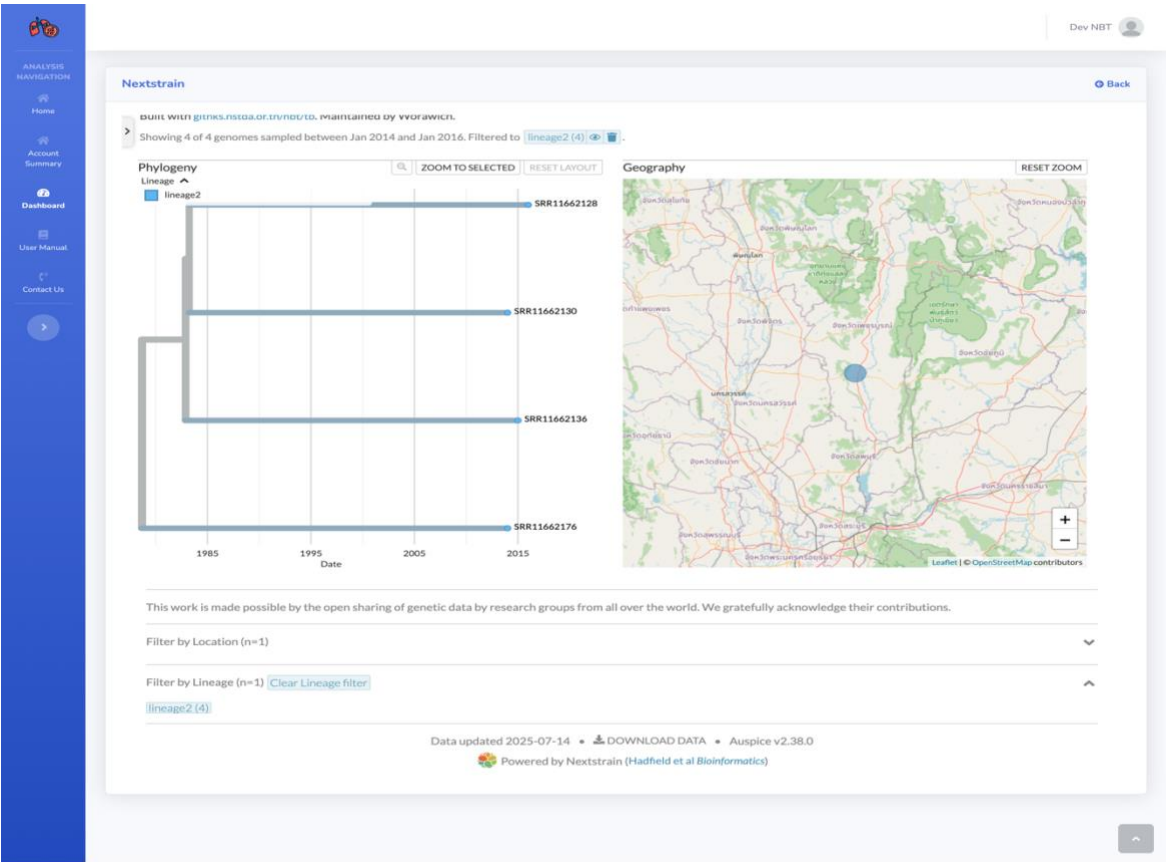


Displays sub-lineage data using a color-coded bar, allowing you to compare the sample's sub-lineage by its color.



SNP-distance: Displays data as a heat map.

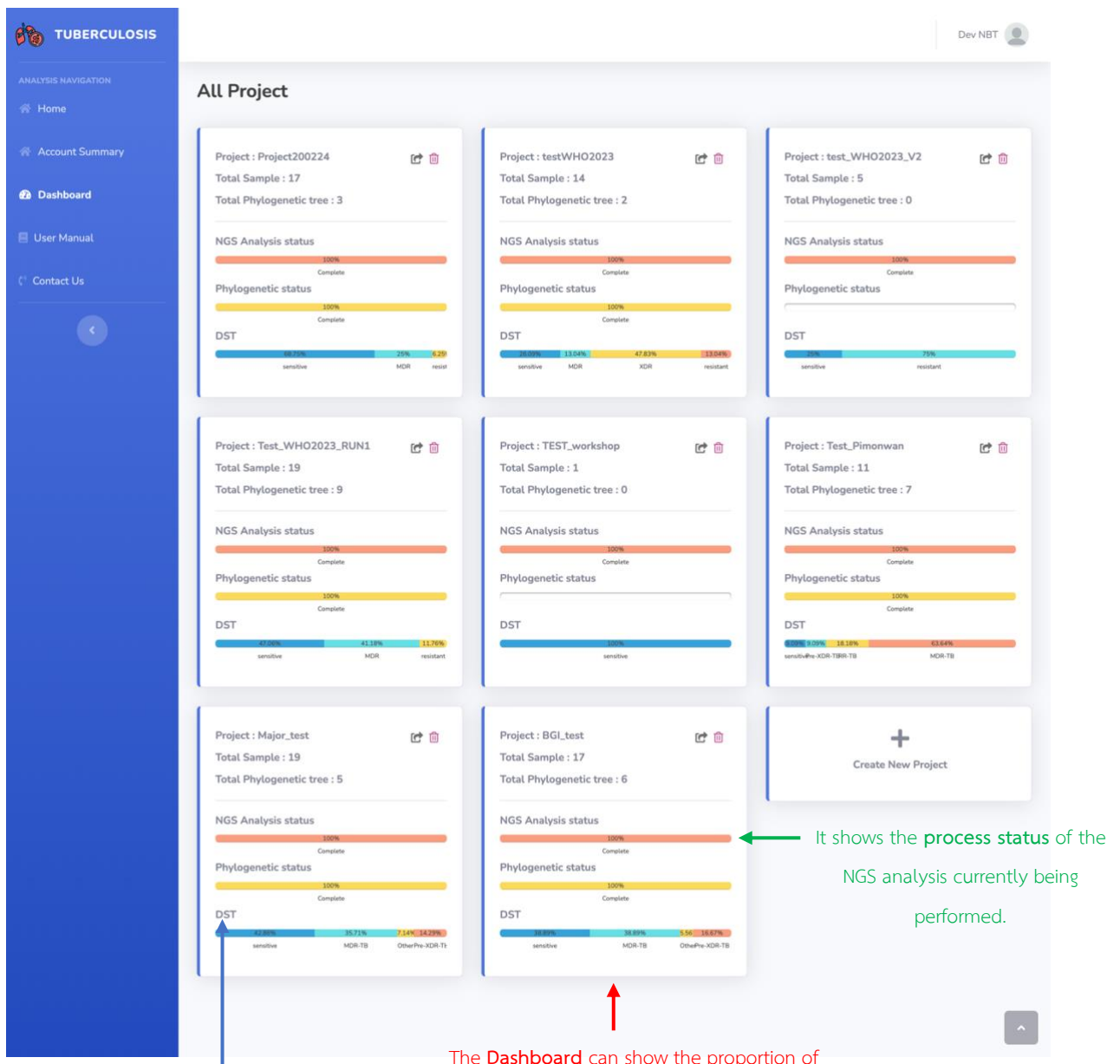
11.2 When you click the **Nextstrain** button, the system will take you to the **Phylogeography** page



11.3 When you click the **View Clonal** button, the system will take you to the **Clonal** page, which displays a **Heatmap**.



12. Click the **Back** button to return to the **Dashboard**.



DST (Drug Susceptibility Testing) data is based on the results from the runs that were able to be processed.

The **Dashboard** can show the proportion of different types of data within the user's project.

13. The **Account Summary** page displays a complete overview of all data.

