

# PhytoBacExplorer Workshop

## Interactive Session Walkthrough

Theresa Wacker

20/02/24

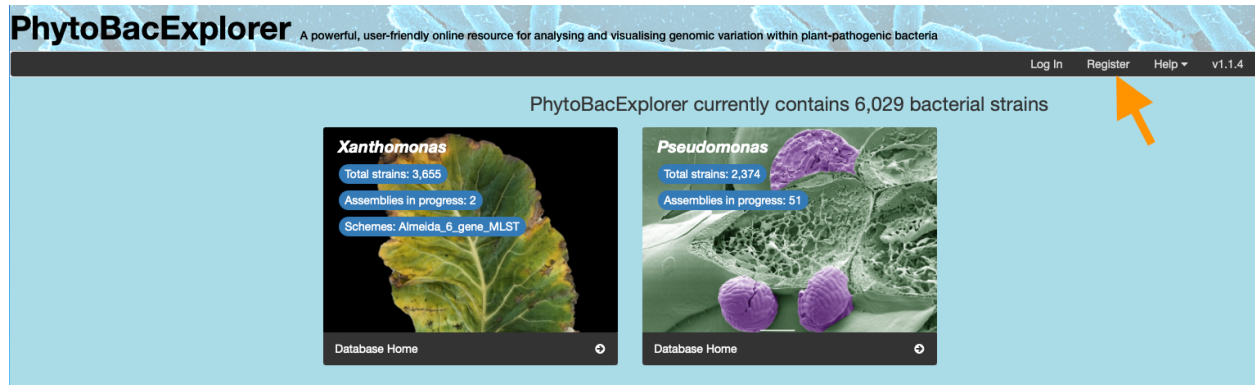
## Contents

<b>1</b>	<b>First steps on PhytoBacExplorer</b>	<b>2</b>
1.1	Register! . . . . .	2
1.2	Explore landing page. . . . .	3
<b>2</b>	<b>Interactive exercise: download target and neighbour assemblies for diagnostic primer generation for <i>Xanthomonas vasicola</i> pv. <i>vasculorum</i></b>	<b>3</b>
2.1	Define and load targets and neighbours (=off-targets) . . . . .	3
2.2	Open "Search Strains dialogue" . . . . .	4
2.3	Select assemblies and generate GrapeTree . . . . .	5
2.4	Annotate your GrapeTree . . . . .	6
2.5	Extracting the relevant STs of our target pathovar . . . . .	7
<b>3</b>	<b>Further Reading:</b>	<b>9</b>
3.1	on PhytoBacExplorer/ Enterobase: . . . . .	9
3.2	on MLSTs: . . . . .	9
<b>4</b>	<b>Questions?</b>	<b>9</b>

# 1 First steps on PhytoBacExplorer

## 1.1 Register!

Register on PhytoBacExplorer (<https://phytobacexplorer.warwick.ac.uk/>).

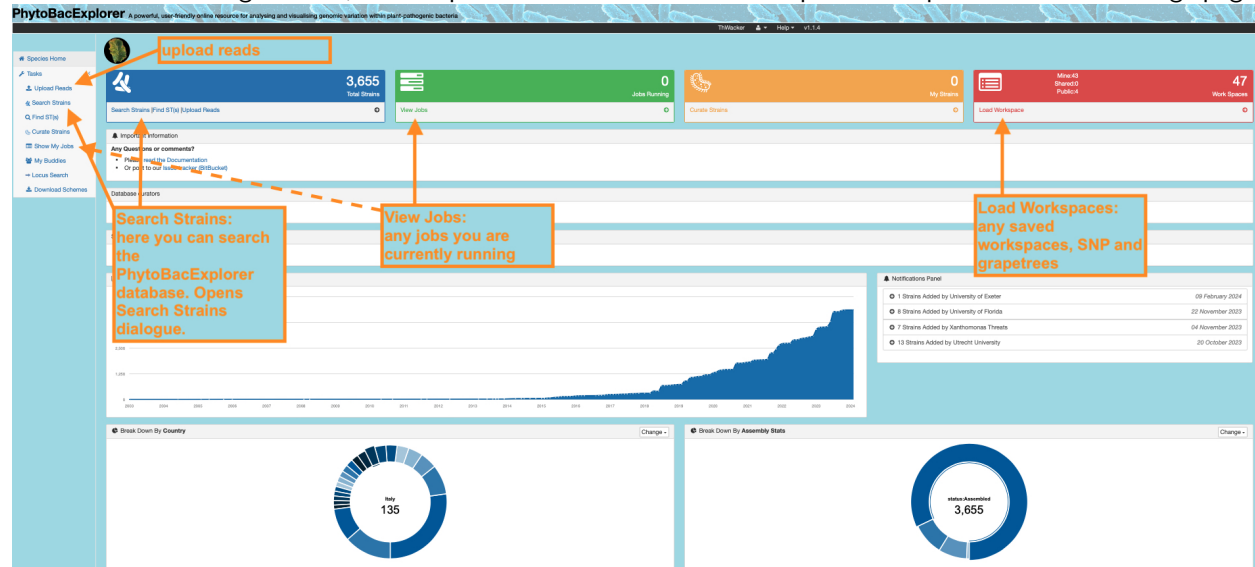


This opens the registration dialogue

The registration dialogue form is titled "Register". It includes a instruction: "Please read the [EnteroBase terms and conditions](#) before registering." The form contains the following fields: Username, Firstname, Lastname, Email, Department, Institution, City, Country, Password, and Confirm password. Below these fields is a checkbox for "I agree to the EnteroBase terms and conditions". At the bottom right of the form is a "Register" button.

## 1.2 Explore landing page.

After successful registration, we explore some of the most important aspects of the landing page:



## 2 Interactive exercise: download target and neighbour assemblies for diagnostic primer generation for *Xanthomonas vasicola* pv. *vasculorum*

### 2.1 Define and load targets and neighbours (=off-targets)

As an example we will load *Xanthomonas vasicola*. You can choose any other target and neighbour (= off-targets) species.

In the example, our target is *vasculorum*. Our neighbours are all other *Xv* assemblies that are not *vasculorum* on PhytoBacExplorer.

## 2.2 Open “Search Strains dialogue”

Search all Strains of Xanthomonas Help

Predefined Search: All Strains My Strains Latest 200 Save Current Query Load Query 92strainsOR

☐ Ignore Legacy Data ☐ Only Editable Strains ☐ Show Failed Assemblies ☐ Show Sub Strains

Strain Metadata ☒ AND ☐ OR Experimental Data

Field	Operator	Value
Species	contains	Xanthomonas vasicola

AND OR Clear

Cancel Submit

Other options for **Field** in the Strain Metadata tab include:

- **Uberstrain:** Certain records are duplicated in that there are many entries for what is essentially the same strain. This can skew analysis because your analysis may produce false clusters, which in reality are just the same strain. Thus in Enterobase such entries are merged and a single Uberstrain is created.
- **Name**
- **Species**
- **Pathvar**
- **Race**
- **Barcode:** not the assembly barcode that will be the name of the downloaded assemblies
- **Data Source:** Accession No. etc.
- **Project:** Bio Project ID etc.

Other **Operator** options include:

- **contains:** all entries where the Field contains the Value
- **in:** clicking into the Value field will open a dialogue which allows you to insert lists of values separated by comma or whitespace.
- **not contains:** all entries where the Field does **not** contain the Value
- **equals:** all entries where Field is exactly what is defined in Value
- **not equals:** all entries where Field is **not** exactly what is defined in Value

Below that, the AND button will allow you to search the same or different Fields that are contain, not contain etc. values. Entries returned will have to fulfill **all** defined criteria

This means that when you search for two different species using AND no entries will be returned because there are no entries that are two different species at the same time

Alternatively, you can use the OR button, which will return any entrie that has a Field of a certain Value **or** other Fields and Valuess.

For instance you, you can search for entries that are *fragariae* or *campestris*, which will return all assemblies that are *fragariae* and all that are *campestris*

## 2.3 Select assemblies and generate GrapeTree

We first select all assemblies:

**PhytoBacExplorer** A powerful, user-friendly online resource for analysing and visualising genomic variation within plant-pathogenic bacteria

Thruwacker v1.1.4

Species Home | Data | View | Workspace | Experiment | Rows Total:80 Filtered:80

Tasks | Upload Reads | Search Strains | Find ST(s) | Curate Strains | Show My Jobs | My Buddies | Locus Search | Download Schemes

Strain	Name	Lab Contact	Species	Pathovar	Race	Comment	Date Enter	Release Da	Barcode	Data Source	Source	Collection	Location	ST	Status	Coverage
■XAN_AA1	NCCPB 890	University of Exeter	Xanthomonas vasculosa	vasculorum			2023-04-25	2021-08-27	XAN_AA0780AA	ORR1564780	ORR1564780	1960	South Africa	67	Assembly	141.7
■XAN_AA1	Xc1601	Kansas State Univ.	Xanthomonas vasculosa	vasculorum			2023-04-25	2020-08-17	XAN_AA1343AA	ORR1245917	ORR1245917	25/9/2016	United States	178	Assembly	150
■XAN_AA1	BCC247	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1376AA	ORR11608302	ORR11608302	2004	Ethiopia	141	Assembly	99.4
■XAN_AA1	BCC210	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1377AA	ORR11608311	ORR11608311	2004	Ethiopia	141	Assembly	82.1
■XAN_AA1	BCC280	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1378AA	ORR11608312	ORR11608312	2004	Ethiopia	141	Assembly	120.4
■XAN_AA1	BCC248	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1379AA	ORR11608313	ORR11608313	2004	Ethiopia	141	Assembly	84.4
■XAN_AA1	BCC265	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1380AA	ORR11608314	ORR11608314	2004	Ethiopia	141	Assembly	75.7
■XAN_AA1	BCC278	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1381AA	ORR11608315	ORR11608315	2004	Ethiopia	141	Assembly	103.3
■XAN_AA1	BCC281	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1382AA	ORR11608316	ORR11608316	2004	Ethiopia	141	Assembly	86
■XAN_AA1	BCC274	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1383AA	ORR11608317	ORR11608317	2004	Ethiopia	141	Assembly	154.6
■XAN_AA1	E52	University of Exeter	Xanthomonas vasculosa	muscorum			2023-04-25	2020-04-25	XAN_AA1384AA	ORR11608318	ORR11608318	2015	Ethiopia	141	Assembly	101
■XAN_AA1	BCC282	University of Exeter	Xanthomonas vasculosa	muscorum			2023-04-25	2020-04-25	XAN_AA1385AA	ORR11608319	ORR11608319	2004	Ethiopia	141	Assembly	218.8
■XAN_AA1	SY109C	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1386AA	ORR11608320	ORR11608320	2014	Uganda	141	Assembly	37.6
■XAN_AA1	BCC246	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1387AA	ORR11608321	ORR11608321	2004	Ethiopia	141	Assembly	579.4
■XAN_AA1	R5P	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1388AA	ORR11608322	ORR11608322	2015	Uganda	141	Assembly	124.9
■XAN_AA1	BCC250	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1389AA	ORR11608323	ORR11608323	2004	Ethiopia	141	Assembly	409
■XAN_AA1	4W4	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1390AA	ORR11608324	ORR11608324	2016	Uganda	141	Assembly	161
■XAN_AA1	SAM119	University of Exeter	Xanthomonas vasculosa				2023-04-25	2018-11-26	XAN_AA1393AA	OCF_003017	OCF_003017	1988	South Africa	138	Assembly	250

then, in the top right corner we change Experimental Data to “Almeida\_6\_gene\_MLST” and click the Grape Tree symbol:

**PhytoBacExplorer** A powerful, user-friendly online resource for analysing and visualising genomic variation within plant-pathogenic bacteria

Thruwacker v1.1.4

Species Home | Data | View | Workspace | Experiment | Rows Total:80 Filtered:80

Tasks | Upload Reads | Search Strains | Find ST(s) | Curate Strains | Show My Jobs | My Buddies | Locus Search | Download Schemes

Strain	Name	Lab Contact	Species	Pathovar	Race	Comment	Date Enter	Release Da	Barcode	Data Source	Source	Collection	Location	ST	Status	Coverage
■XAN_AA1	NCCPB 890	University of Exeter	Xanthomonas vasculosa	vasculorum			2023-04-25	2021-08-27	XAN_AA0780AA	ORR1564780	ORR1564780	1960	South Africa	67	Assembly	141.7
■XAN_AA1	Xc1601	Kansas State Univ.	Xanthomonas vasculosa	vasculorum			2023-04-25	2020-08-17	XAN_AA1343AA	ORR1245917	ORR1245917	25/9/2016	United States	178	Assembly	150
■XAN_AA1	BCC247	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1376AA	ORR11608302	ORR11608302	2004	Ethiopia	141	Assembly	99.4
■XAN_AA1	BCC210	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1377AA	ORR11608311	ORR11608311	2004	Ethiopia	141	Assembly	82.1
■XAN_AA1	BCC280	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1378AA	ORR11608312	ORR11608312	2004	Ethiopia	141	Assembly	120.4
■XAN_AA1	BCC248	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1379AA	ORR11608313	ORR11608313	2004	Ethiopia	141	Assembly	84.4
■XAN_AA1	BCC265	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1380AA	ORR11608314	ORR11608314	2004	Ethiopia	141	Assembly	75.7
■XAN_AA1	BCC278	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1381AA	ORR11608315	ORR11608315	2004	Ethiopia	141	Assembly	103.3
■XAN_AA1	BCC281	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1382AA	ORR11608316	ORR11608316	2004	Ethiopia	141	Assembly	86
■XAN_AA1	BCC274	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1383AA	ORR11608317	ORR11608317	2004	Ethiopia	141	Assembly	154.6
■XAN_AA1	E52	University of Exeter	Xanthomonas vasculosa	muscorum			2023-04-25	2020-04-25	XAN_AA1384AA	ORR11608318	ORR11608318	2015	Ethiopia	141	Assembly	101
■XAN_AA1	BCC282	University of Exeter	Xanthomonas vasculosa	muscorum			2023-04-25	2020-04-25	XAN_AA1385AA	ORR11608319	ORR11608319	2004	Ethiopia	141	Assembly	218.8
■XAN_AA1	SY109C	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1386AA	ORR11608320	ORR11608320	2014	Uganda	141	Assembly	37.6
■XAN_AA1	BCC246	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1387AA	ORR11608321	ORR11608321	2004	Ethiopia	141	Assembly	579.4
■XAN_AA1	R5P	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1388AA	ORR11608322	ORR11608322	2015	Uganda	141	Assembly	124.9
■XAN_AA1	BCC250	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1389AA	ORR11608323	ORR11608323	2004	Ethiopia	141	Assembly	409
■XAN_AA1	4W4	University of Exeter	Xanthomonas vasculosa				2023-04-25	2020-04-25	XAN_AA1390AA	ORR11608324	ORR11608324	2016	Uganda	141	Assembly	161
■XAN_AA1	SAM119	University of Exeter	Xanthomonas vasculosa				2023-04-25	2018-11-26	XAN_AA1393AA	OCF_003017	OCF_003017	1988	South Africa	138	Assembly	250

The Create GrapeTree dialogue opens and we can give the tree a name, as well as chose between minimum spanning tree algorithms or neighbourhood joining algorithms. Both generate trees based on MLST allelic profiles<sup>1</sup>. Since we do not have missing data (all entries have defined STs), I chose a simple NJ algorithm (Ninja NJ).

**Create GrapeTree**

Name: Almeida\_6\_gene\_MLST\_Tree\_2024-10-10

Algorithm: NINJA NJ

Selected Only: ☐

Scheme: Almeida\_6\_gene\_MLST

Workspace: None

No. Strains (With STs): 80

Number of Nodes (STs): 9

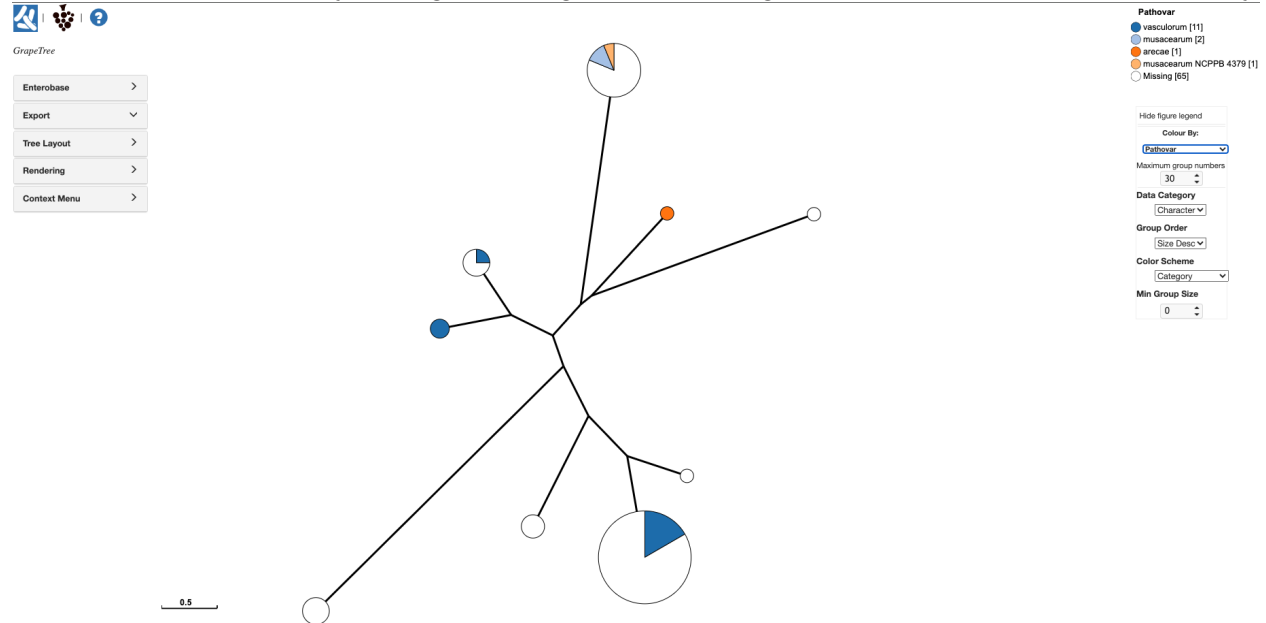
Submit

<sup>1</sup>compare <https://enterobase.readthedocs.io/en/latest/grapetree/tutorial-2.html?highlight=grape%20tree> and Z Zhou, NF Alikhan, MJ Sergeant, N Luhmann, C Vaz, AP Francisco, JA Carrico, M Achtman (2018) “GrapeTree: Visualization of core genomic relationships among 100,000 bacterial pathogens”, Genome Res. doi: <https://doi.org/10.1101/gr.232397.117>

Suggestion: if you want to, try out multiple different algorithms and compare topology.

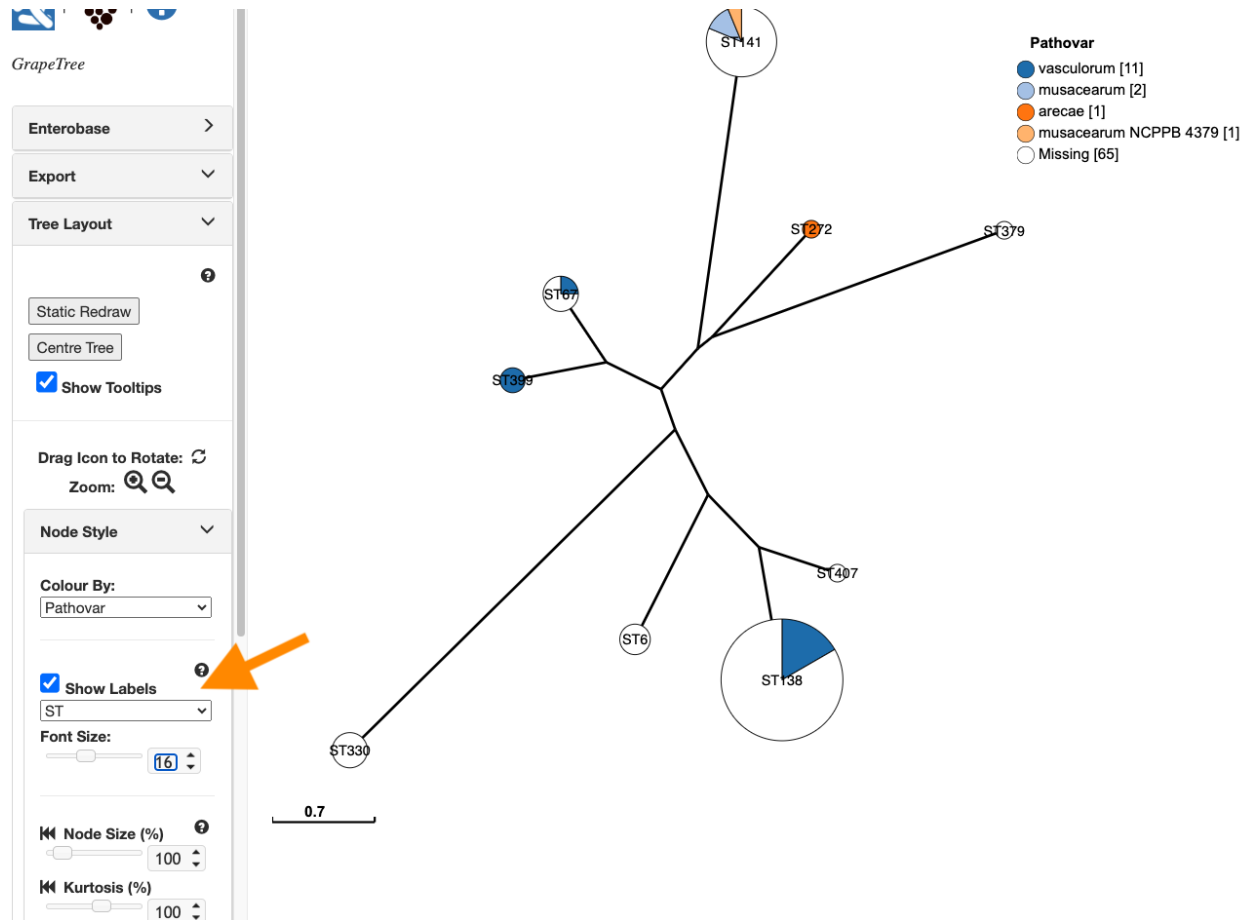
## 2.4 Annotate your GrapeTree

We can annotate our tree by clicking on the legend and choosing which metadata we want to colour it by:



Suggestion: try colouring by other metadata, like country, too.

We can also annotate the nodes and show their labels. We are interested in the **STs**:



## 2.5 Extracting the relevant STs of our target pathovar

Our **target STs** in the Example are **67, 138 and 399**.

We can press **Shift** and click on the nodes to select them. Then we can click **Load Selected** in the left panel.





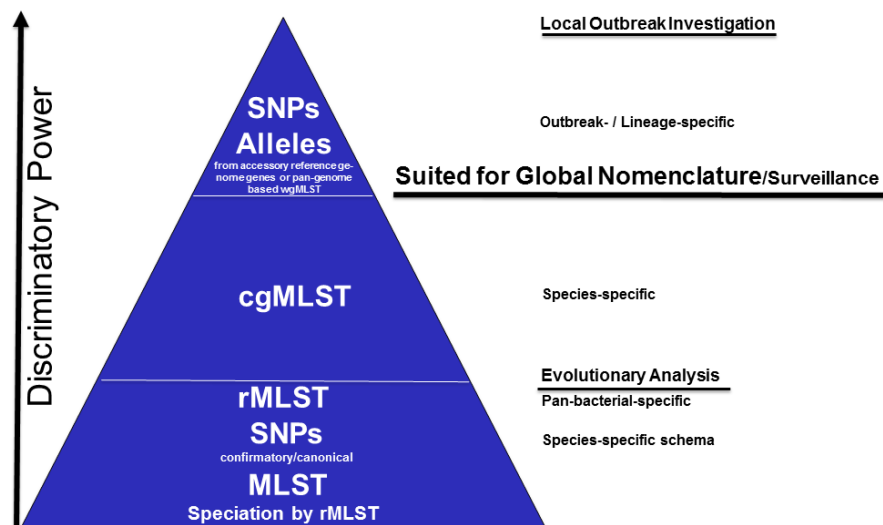
### 3 Further Reading:

#### 3.1 on PhytoBacExplorer/ Enterobase:

- Enterobase Documentation: <https://enterobase.readthedocs.io/en/latest/features/main-search-page.html>
- GrapeTree Documentation: <https://enterobase.readthedocs.io/en/latest/grapetree/grapetree-about.html>
- Enterobase Publication: <https://doi.org/10.1101/gr.251678.119>

#### 3.2 on MLSTs:

- MLST scheme based on *Almeida et al.*: doi:10.1094/PHYTO-100-3-0208
- general explanation of MLSTs:
  - Ridom Bioinformatics: <https://www.ridom.de/seqsphere/cgmlst/> (picture copyright belongs to Ridom



- Bioinformatics)
- PubMLST: <https://pubmlst.org/multilocus-sequence-typing>

### 4 Questions?

- Laura Baxter: [laura.baxter@warwick.ac.uk](mailto:laura.baxter@warwick.ac.uk)
- Sascha Ott: [s.ott@warwick.ac.uk](mailto:s.ott@warwick.ac.uk)
- Theresa Wacker: [t.wacker2@exeter.ac.uk](mailto:t.wacker2@exeter.ac.uk)
- Regular drop-in session via Zoom on most Tuesdays between 9 am and 12 noon (UK time): <https://zoom.us/j/8852150277>