

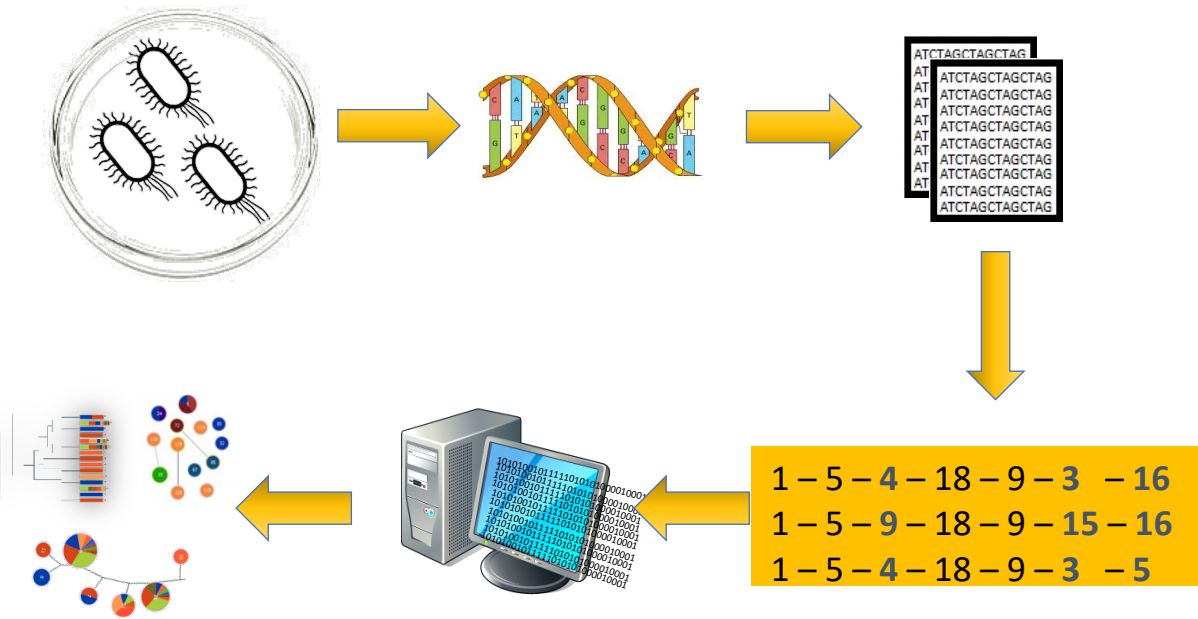
A JavaScript library for interactive data visualization in phylogenetics

Cátia Vaz, ISEL-IPL; INESC-ID

Inforum 2022

(co-authors: Ana Correia, Inês Sousa, Adriano Baptista, Francisco Filipe, Marta Ferreira, and Alexandre Francisco)

Phylogenetic inference



- **PHYLOGENETICS** is the study of the evolutionary history and relationships among individuals or groups of organisms
- The **evolutionary relationships** between different species or taxa are usually inferred through known phylogenetic analysis techniques
 - **Sequencing** the isolate data
 - **Assembly** the sequences, comparing the draft genome to a database of gene alleles
 - **Profiling** - Given the assembly results we can create an allelic profile characterizing the strain
 - Usually abstracted to categorical indexes.
- After profiling, the reconstruction of a **phylogenetic tree** can be done.

Phylogenetic data

- **Phylogenetic data** includes:
 - Genetic sequencing data
 - Profiling data
 - Phylogenetic inference data (ex: tree, networks)
 - Epidemiological data
- A **major challenge** consists on
 - Phylogenetic data integration
 - Phylogenetic data visualization

Phylogenetic tools

- **Phylogenetic tools** that allow to visualize phylogenetic trees together with metadata integration include:
 - PHYLOViZ [1,2],
 - PHYLOViZ Online [3]
 - GrapeTree [4]
 - Usually, these tools are standalone widgets that were not designed to serve as middleware.
-
- **JavaScript packages** for visualizing phylogenetic trees that can be integrated as components of other applications include:
 - PhyD3 [5]
 - phylotree [6]
 - Not designed to integrate and visualize epidemiological data.

phytotree

- **A JavaScript library:**
 - That allows simultaneously view the resulting trees from phylogenetic inference algorithms and their integration with epidemiological data
 - That can be easily integrated with phylogenetic inference tools
 - Designed to promote easier extensibility
 - With main modules loosely coupled
 - Reuse the d3.js library
- **phytotree** is being integrated in the new Phyloviz version (ongoing work).
- It is also available a **cross-platform JavaScript application** for demonstration purposes (using Electron)
 - <https://github.com/DIVA-IPL-Project/phytotree>

phytotree features

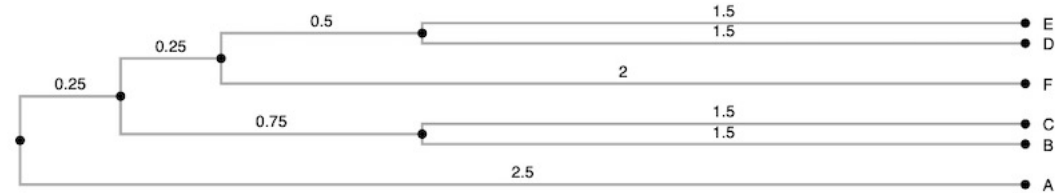
- With this library, users are able to:
 - visualize trees with a **dendrogram** or **radial layout**;
 - **integrate** epidemiological data;
 - **collapse** or **expand** nodes of the tree, as well as adding labels;
 - **plot statistics** based on queries over isolate data, with node charts;
 - **change node** color and size, **link** thickness and **labels** size;
 - **save** a simple **report** with the tree view and statistics;
 - **save** the current view and applied transformations (**study**);
 - **load** a saved **study**.

Phytotree main modules

Modules	responsible for:
Data Parser	<ul style="list-style-type: none">• reading the phylogenetic data• converting data to a JSON object created to represent a study
Visualization	<ul style="list-style-type: none">• drawing phylogenetic trees using radial or dendrogram layouts• performing graphical transformations over the trees
Storage	<ul style="list-style-type: none">• managing the study state, such as the actions of save and load a study• produce a report in PDF format

Data Parser

- The library receives as input 3 types of data:
 - a file with the **phylogenetic tree**, obtained from an inference method [7]
 - In newick or nexus format
 - a file in tabular format, with the **allele profiles**, obtained from the application of a typing method [8]
 - a file also in tabular format, containing the **isolate/epidemiological data**.



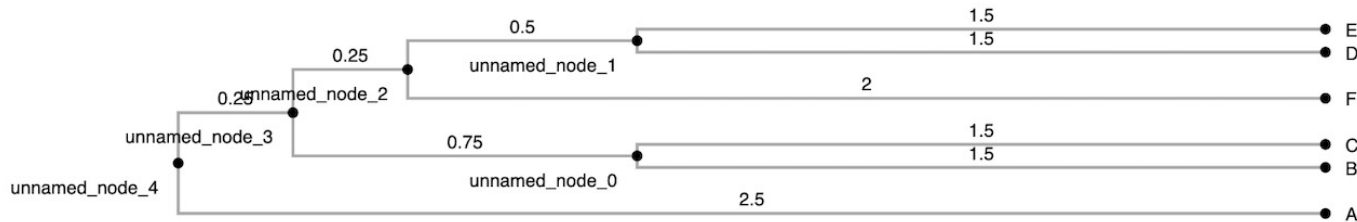
```
(A:2.5,((B:1.5,C:1.5)_:0.75,(F:2.0,(D:1.5,E:1.5)_:0.5)_:0.25)_:0.25)_;
```

ST	Gene_1	Gene_2	Gene_3	Gene_4	Gene_5
A	10	6	6	12	13
B	5	4	4	2	15
C	5	3	4	6	2
D	2	2	4	8	7
E	2	2	1	1	12
F	1	3	1	1	1

ST	Country	Year
A	China	1983
D	China	1996
C	China	1997
B	United-Kingdom	2010
E	China	2010
E	China	2010
E	China	2010
A	United-Kingdom	2012
A	United-Kingdom	2012

Data Parser

- Produces a json object, describing
 - the input data
 - The ongoing study
- This approach is easier for extending the library to new tree formats

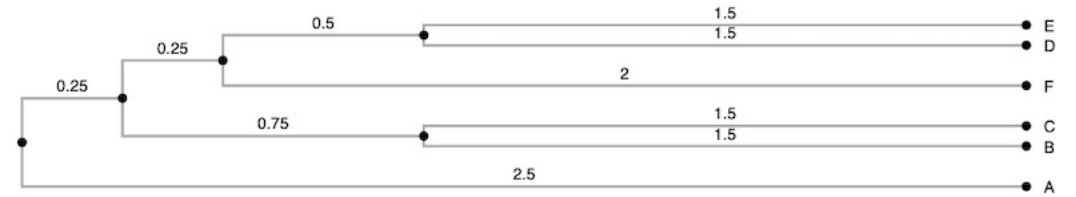
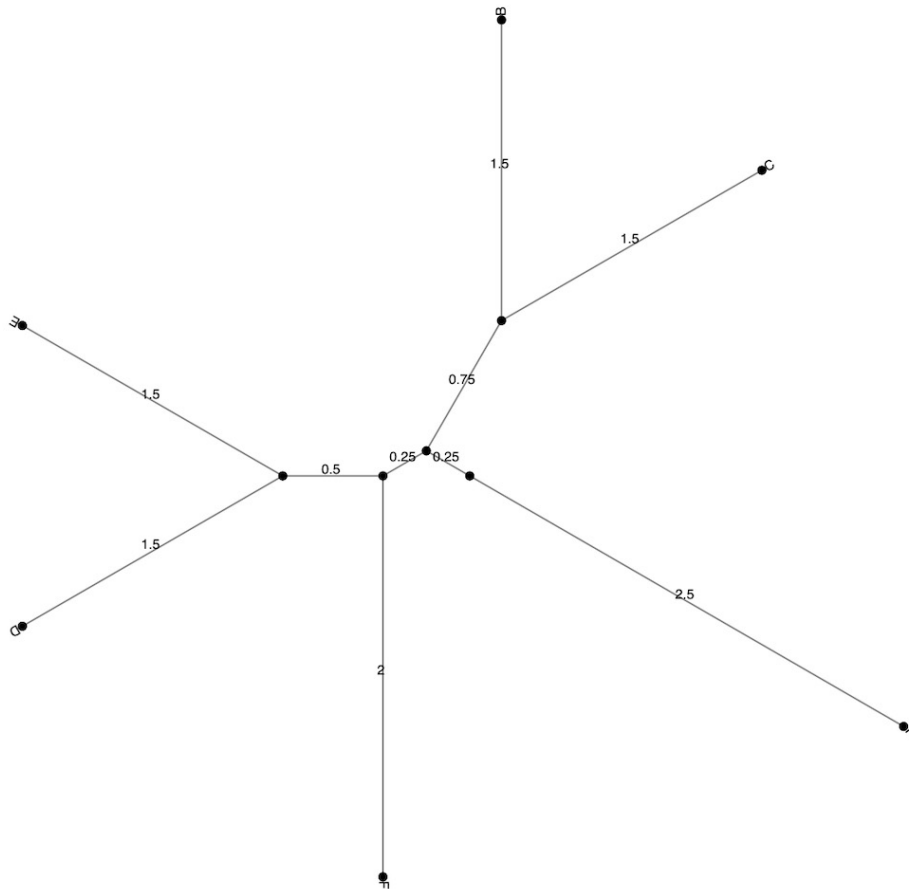


```



```

Visualization



- Dendrogram view

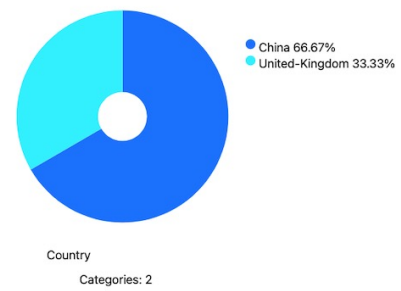
- Radial view

Visualization

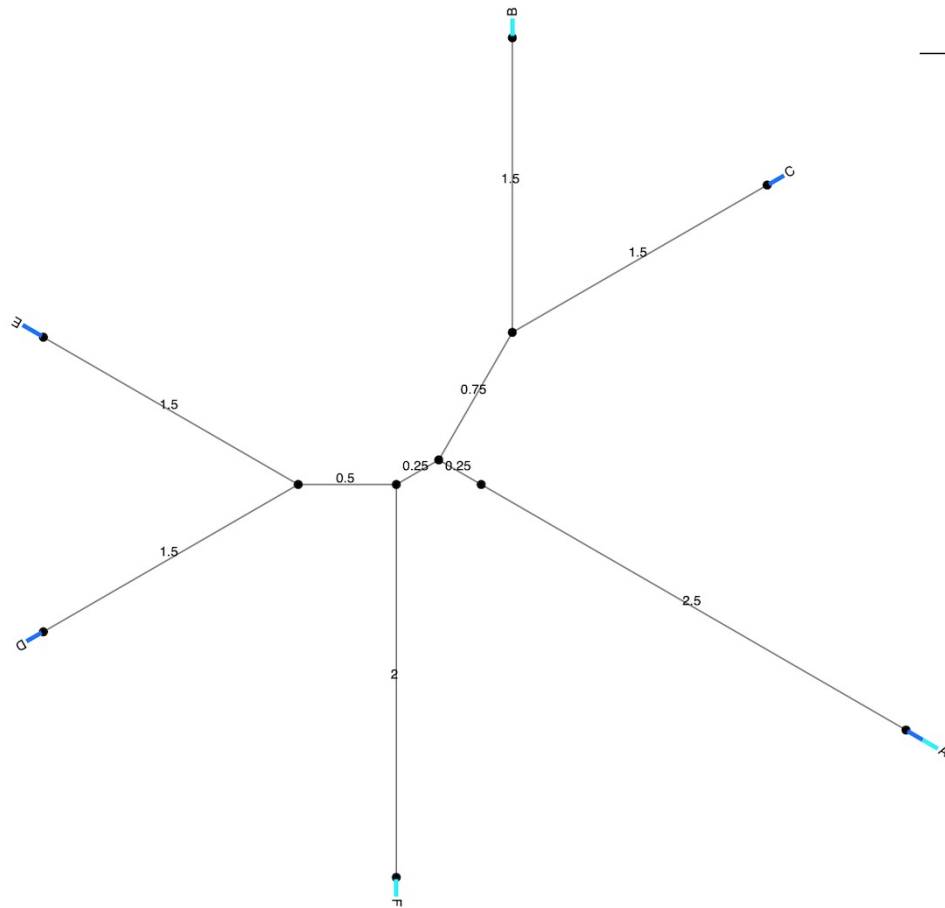
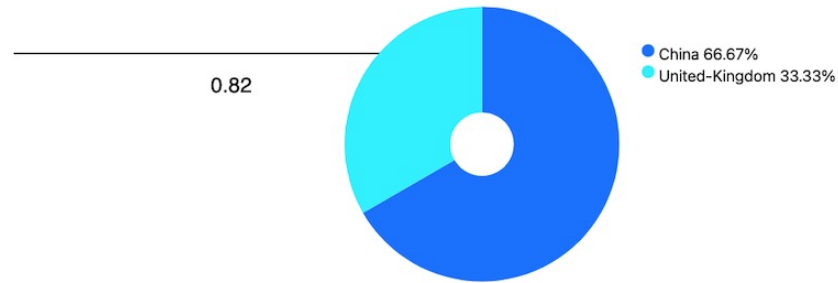
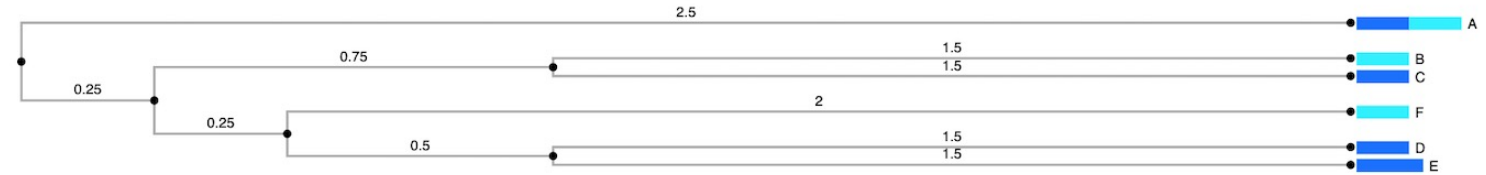
Tree Profile Data Isolate Data		
ST	Country	Year
B	United-Kingdom	2010
C	China	1997
D	China	1996
E	China	2010
E	China	2010
E	China	2010
F	United-Kingdom	2012
A	China	1983
A	United-Kingdom	2012

[Link to Dendrogram](#)

[Link to Radial](#)



Visualization



Supported graphical transformations over the trees:

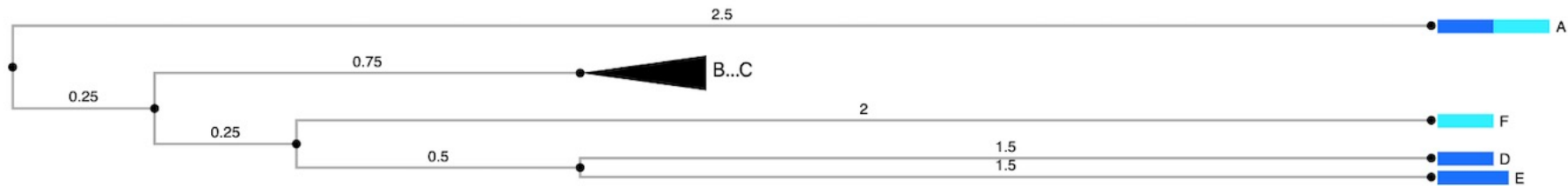
1. scaling the tree
2. Aligning the leaves of the dendrogram
3. Add labels to nodes and links
4. Isolate/Epidemiological integration
 - Color can be changed
 - The bar chart is proportional to the number of isolates with that profile

Visualization

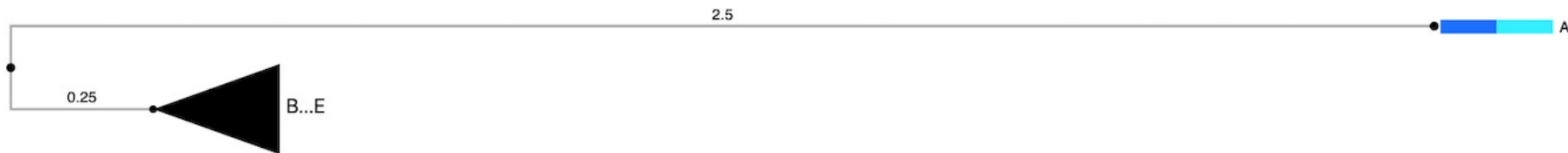
Supported graphical transformations over the trees:

5. expand/collapse subtrees

- Relevant for very large trees
- It is performed a DFS changing the visibility of each node.
- The size of the triangle is proportional to the contained nodes.



Collapsing nodes of subtree that contains B and C



Collapsing nodes of subtree that contains nodes B, C, F, D, and E

Storage

- allow to save the current study
 - With a JSON format
 - the transformations applied to the current selected view are also saved
- Allow to load the current study
- Report the study in PDF format, using the jsPDF library

```
{
  "type": "dendrogram",
  "data": {
    "tree": [
      {
        "source": null,
        "target": "unnamed_node_4",
        "value": 0
      },
      {
        "source": "unnamed_node_4",
        "target": "A",
        "value": 2.5
      }
    ],
    "barChart": {
      "width": "44.45378125959109",
      "height": "11",
      "fill":
    }
  }
}
```

(...)

Final Remarks

- phytotree is a freely available and easy to use JavaScript library for the analysis of phylogenetic data, including epidemiological data.
- The library development was focused on modularity, reusability and efficiency, giving the users the possibility to easily integrate on their web applications or even to extend it
- Future work will include new visualization layouts and adding more statistical information.

References

- [1] Francisco, A.P., Vaz, C., Monteiro, P.T., Melo-Cristino, J., Ramirez, M., Carriço, J.A.: *Phyloviz: phylogenetic inference and data visualization for sequence based typing methods*. BMC bioinformatics 13(1), 1–10 (2012)
- [2] Nascimento, M., Sousa, A., Ramirez, M., Francisco, A.P., Carriço, J.A., Vaz, C.: *Phyloviz 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods*. Bioinformatics 33(1), 128–129 (2017)
- [3] Ribeiro-Gonçalves, B., Francisco, A.P., Vaz, C., Ramirez, M., Carriço, J.A.: *Phyloviz online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees*. Nucleic acids research 44(W1), W246– W251 (2016)
- [4] Zhou, Z., Alikhan, N.F., Sergeant, M.J., Luhmann, N., Vaz, C., Francisco, A.P., Carriço, J.A., Achtman, M.: *Grapetree: visualization of core genomic relationships among 100,000 bacterial pathogens*. Genome research 28(9), 1395–1404 (2018)
- [5] Kreft, L., Botzki, A., Coppens, F., Vandepoele, K., VanBel, M.: *Phyd3: a phylo-genetic tree viewer with extended phyloxml support for functional genomics data visualization*. Bioinformatics 33(18), 2946–2947 (2017)
- [6] Shank, S.D., Weaver, S., Kosakovsky Pond, S.L.: *phylotree.js-a javascript library for application development and interactive data visualization in phylogenetics*. BMC bioinformatics 19(1), 1–5 (2018)
- [7] Vaz, C., Nascimento, M., Carriço, J.A., Rocher, T., Francisco, A.P.: *Distance-based phylogenetic inference from typing data: a unifying view*. Briefings in Bioinformatics 22(3), bbaa147 (2021)
- [8] Vaz, C., Francisco, A.P., Silva, M., Jolley, K.A., Bray, J.E., Pouseele, H., Roth-ganger, J., Ramirez, M., Carriço, J.A.: *Typon: the microbial typing ontology*. Journal of Biomedical Semantics 5(1), 1–11 (2014)