

(1) Submission ID#1527439

Visual analytic developments on the PubMLST databases

Author(s)

Keith A. Jolley, PhD

Senior Researcher

Department of Biology, University of Oxford

Margaret J. Varga, PhD

Senior Researcher

Department of Biology, University of Oxford

Holly B. Bratcher, Doctor

Senior researcher and laboratory manager

University of Oxford

Martin C. J. Maiden, PhD

Professor

Department of Biology, University of Oxford

Background

Visual analytics in the form of dashboards and interactive graphical breakdowns allow users to rapidly assimilate and understand the nature of complex datasets. We have recently introduced front-end and query dashboards that show the breakdown of field values within complete databases and within user-defined and public projects.

Aim/Methods

The aims are to provide easily-digested overviews of PubMLST database contents and the dataset results returned from queries, to improve understanding and the user experience.

Results

Default dashboards can be defined for individual databases or projects and these are fully customizable with users able to: 1) modify the size and layout of individual visualizations and the selection of fields shown; 2) select how these are visualized with a choice of different charts including pie charts, doughnut charts, bar charts, treemaps, word clouds, lists, gauges, and geographical maps; 3) choose colour themes and individual colours for different elements. Separate dashboards are also displayed following an isolate database query that show the composition of isolates within a returned dataset. In addition, we have introduced high-resolution mapping that allows the values of individual geographic fields, such as town or city, to be linked to lookup tables of GPS coordinates so that isolation locations can be displayed, as well as exported to external tools such as Microreact. Finally, an interactive Data Explorer tool has been developed that links to front-end dashboards and allows users to investigate how different fields relate to each other, for example showing how

clonal complexes are distributed among capsule groups and countries. Links within the Data Explorer take users directly to a query page that returns datasets filtered by selected criteria.

Conclusions

The new functionality enhances the user experience of the PubMLST databases, providing improved understanding of dataset contents and guidance of how complex queries can be constructed.

Uploaded File(s)

Supplemental Document Upload

Default front-end dashboard on the PubMLST *Neisseria* database

