

Coronavirus data analyses published on open-source platform Galaxy

February 27 2020



Credit: CC0 Public Domain

Dr. Wolfgang Maier and Dr. Björn Grüning from the University of Freiburg, together with researchers from universities in Belgium, Australia and the U.S., have reviewed previously available data on

sequences of the novel coronavirus and published their analyses on [the open source platform Galaxy](#). The two Freiburg bioinformaticians hope that this will facilitate the exchange of data between authorities, institutes and laboratories dealing with the virus. The Freiburg researchers have documented their approach and results on the bioRxiv portal.

The Galaxy platform is suitable for big data analysis in life sciences. Public servers provide scientists with [free access](#) to analysis tools and reproducible evaluation procedures. Maier, Grüning and their colleagues have used Galaxy to re-analyze all publicly available COVID-19 genome data for their study. Previous publications often lacked transparency with regard to data analysis, explains Grüning. For example, only one of four studies on the COVID-19 genome published at the beginning of February contained clear information on the raw data used, says Grüning. "And the analyses were also not well documented and not reproducible." As a result, it was not possible to understand or verify the respective statements.

Within a few days, the team was able to apply identical workflows to each of the available sequences and make them publicly accessible via Galaxy. As a result, researchers worldwide now have access to the network of Galaxy servers in Europe, the U.S. and Australia, not only for the evaluation of the data, but also as the scientific infrastructure for their own work with COVID-19 data. This means that scientists will be able to analyze new COVID-19 datasets on public servers within hours after their release through the same workflows used to analyze the current data.

The researchers agree that there is currently a lack of data exchange in research on COVID-19, says Maier. This should change with the publications on Galaxy. "Global cooperation, which is necessary to deal with public health emergencies such as the COVID-19 outbreak,

ultimately requires unrestricted access to data, analytical tools and computational infrastructure."

The Galaxy project was initiated at Penn State University in the U.S. and further developed at the University of Freiburg in the Collaborative Research Centre "Medical Epigenetics" and as part of the German Network for Bioinformatics Infrastructure (de.NBI). The European server is located in the IT Services department at the University of Freiburg and is designed as a community project. The data is freely accessible online. Scientists who wish to use the server do not need to have any programming skills. All analyses can be set up through a graphical user interface. The team at the University of Freiburg led by Prof. Dr. Rolf Backofen from the Department of Computer Science is responsible for Galaxy's further development.

More information: No more business as usual: agile and effective responses to emerging pathogen threats require open data and open analytics, *bioRxiv* (2020). [DOI: 10.1101/2020.02.21.959973](https://doi.org/10.1101/2020.02.21.959973)

Provided by Albert Ludwigs University of Freiburg

Citation: Coronavirus data analyses published on open-source platform Galaxy (2020, February 27) retrieved 3 May 2024 from <https://medicalxpress.com/news/2020-02-coronavirus-analyses-published-open-source-platform.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--