

CRG standardises COVID-19 data analysis to aid international research efforts

March 27 2020



Credit: CC0 Public Domain

Researchers from the Centre for Genomic Regulation (CRG) have launched a new database to advance the international research efforts studying COVID-19.

The publicly-available, free-to-use resource (covid.crg.eu) can be used by researchers from around the world to study how different variations of the virus grow, mutate and make proteins.

"Scientists are working round the clock to understand SARS-CoV-2, the virus causing COVID-19, so that we can find its weak spots and beat it. A huge amount of scientific data is being published around the world," says Eva Novoa, a researcher at the CRG in Barcelona.

"However, some of the technologies we use to study SARS-CoV-2, such as nanopore RNA sequencing, are so new that the results of one paper aren't comparable to another due to the patchwork of different standards and methodologies used. We are taking all this data and analyzing it so that it meets a more universally comparable standard. This will help researchers more quickly and accurately spot the strengths and weaknesses of the coronavirus."

To understand how the coronavirus grows, mutates and replicates, scientists have to sequence the RNA of COVID-19. The RNA sequence reveals crucial information about the proteins the virus makes to invade [human cells](#) and replicate, which in turn informs governments on the infectiousness and severity of the pandemic.

Traditional sequencing tools can take a long time to provide results. In recent years, sequencing data in real time has become a reality thanks to the use of nanopore sequencing technologies, revolutionizing genomics research and disease outbreak monitoring. Nanopore sequencing provides scientists and clinicians with immediate access to the DNA and RNA sequence information of any living cell in real-time, enabling a rapid response against the threat of a pandemic.

However, the raw data produced by nanopore sequencing is highly complex. Scientists and clinicians currently lack systematic guidelines

for the reproducible analysis of the data, limiting the vast potential of the nascent technology.

To standardize the analysis of publicly available SARS-CoV-2 [nanopore sequencing](#) data, researchers at the Centre for Genomic Regulation (CRG) in Barcelona are using MasterOfPores, a computer program developed by the group of Eva Novoa and CRG Bioinformatics Unit. The software was first described last week in *Frontiers in Genetics*.

"The internet and an increasing culture of open science, data sharing and preprints have transformed the research landscape. Infrastructure that would take months to set up to research an emerging virus can now be done in just a few days owing to novel scientific computing approaches," says Julia Ponomarenko, Head of the Bioinformatics Unit at the CRG.

MasterOfPores can be executed on any Unix-compatible OS on a computer, cluster or cloud without the need of installing any additional software or dependencies, and is freely available in Github. The publicly-available, free-to-use resource has currently analysed 3TB of SARS-CoV-2 [nanopore](#) RNA sequencing data. The CRG researchers will continue to update the resource with new data as soon as it becomes available.

More information: covid.crg.eu

Provided by Center for Genomic Regulation

Citation: CRG standardises COVID-19 data analysis to aid international research efforts (2020, March 27) retrieved 9 May 2024 from <https://medicalxpress.com/news/2020-03-crg-standardises-covid-analysis-aid.html>

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.