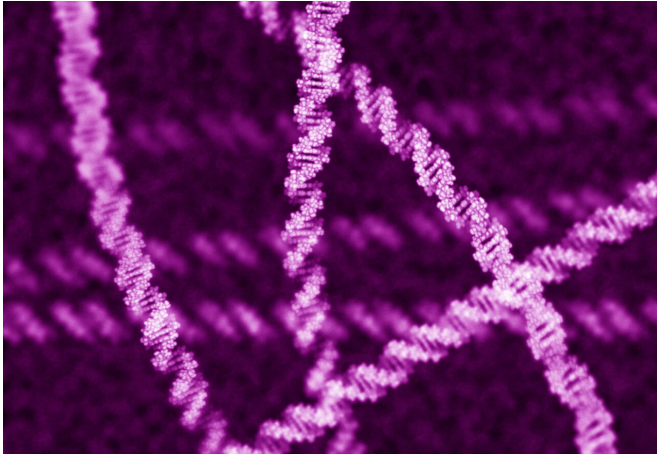


New database enhances genomics research collaboration

4 November 2019



Sharing datasets that reveal the function of genomic variants in health and disease has become easier, with the launch of a new, open-source database called MaveDB. Credit: Walter and Eliza Hall Institute, Australia

Sharing datasets that reveal the function of genomic variants in health and disease has become easier, with the launch of a new, open-source database developed by Australian and North American researchers.

The MaveDB database is a repository for data from experiments—called multiplex assays of variant effect (MAVEs) - that systematically measure the impact of thousands of individual sequence variants on a gene's function. These experiments can provide valuable information about how proteins produced by that gene function, how variants in that gene may contribute to disease, and how to engineer synthetic versions of naturally occurring proteins that are more effective than the original protein.

MaveDB is the first publicly accessible database for this data. Its development was led by Dr. Alan Rubin from the Walter and Eliza Hall Institute,

Australia, Associate Professor Douglas Fowler from the University of Washington, US, and Professor Frederick Roth from the University of Toronto, Canada. MaveDB was described today in the journal *Genome Biology*.

Enhancing genomics research

MAVEs have revolutionised the ability of researchers to understand the function of [genes](#) and their roles in disease, Dr. Rubin said.

"In the past, researchers had to focus on a handful of changes in a gene to understand its function," he said. "It was too complex to generate the data from an exhaustive scan of variants of a gene that might be hundreds or thousands of bases in length.

"The development of MAVEs provided a way for researchers to experimentally measure every single genetic change in a gene with its functional consequence. These assays can handle tens of thousands of genetic variants, allowing researchers to home in on the relevant changes and place them in context."

Until now, MAVE data from experiments has existed in isolation, with data from individual studies uploaded to journal websites when [research papers](#) are published, or provided upon request to other researchers.

"This made it hard for researchers to access the data of other groups, or even know that a particular MAVE experiment had been done. So it potentially hindered collaborations and the progress of genomics research," Dr. Rubin said.

"MaveDB makes it easier for scientists to share their datasets in a single location, using a flexible format that is applicable to multiple research fields, and enables other scientists to easily access this data to enhance their research. We've also ensured MaveDB can 'talk' to other databases to add an

extra level of collaborative capacity. For the growing field of MAVE research this database is an important step towards open science and reproducibility by ensuring data is made available."

Data obtained from MAVEs has many applications, including understanding how a gene or protein functions, measuring the involvement of genetic variants in a disease, or understanding how a synthetic protein—such as those used in biotechnology—can be made more effective.

Visualising data

As well as establishing MaveDB, the team also developed data visualisation software, called MaveVis, that makes it easier for researchers to understand and interpret the results of MAVE experiments.

"MaveVis provides an immediate and consistent display for MAVE data, including valuable annotations such as protein structure information, that will accelerate collaborative research," Dr. Rubin said.

"We envision that as MaveDB becomes more widely used within the bioinformatics community, other applications will be added that provide new ways to visualise and interpret complex genomics data—leading to new discoveries that enhance biomedical research. This could underpin the development of new medicines, or the understanding of how a patient's genomic variants contribute to a disease."

More information: Daniel Esposito et al, MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect, *Genome Biology* (2019). [DOI: 10.1186/s13059-019-1845-6](https://doi.org/10.1186/s13059-019-1845-6)

Provided by Walter and Eliza Hall Institute

APA citation: New database enhances genomics research collaboration (2019, November 4) retrieved 21 June 2022 from <https://medicalxpress.com/news/2019-11-database-genomics-collaboration.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.