Database of inbred mouse proteins responsible for strain discrepancies created

August 8 2017, by Bob Yirka

Credit: CC0 Public Domain

(Phys.org)—A trio of researchers affiliated with several institutions in Belgium has developed a bioinformatics tool that allows for comparing protein-coding genes of 36 inbred mouse strains to the C57BL/6J strain. In their paper published in *Proceedings of the National Academy of Sciences*, Steven Timmermans, Marc Van Montagu and Claude Libert describe the tool, the database they created using it and applications for other researchers.

Scientists use mice to test out ideas, whether behavioral or physical, but as the researchers with this new effort note, not all strains of mice of a given species are the same. A gene mutation in one common strain, for example, can lead to colon polyps, while the same mutation in another strain does not. A failure to account for such characteristics can, quite obviously, lead to problems when researchers choose a strain of mice to conduct experiments. To overcome this problem, Timmermans, Van Montagu and Libert developed a tool to identify such differences between a standard strain and 36 others. Then, they used the tool to catalogue known differences in a database that could be accessed by other researchers.

The researchers chose C57BL/6J as the comparative strain because it was the first to be sequenced. The other 36 strains that appear in the database have since also been sequenced. The tool they developed was essentially a technique for using available information about a sequenced strain, looking at the coding regions, computationally mutating them and then translating the DNA codons into their respective amino acid sequences—that allowed for predicting which mutations in which genes would impact which resultant protein. Those results were entered into the database along with other known information obtained through other studies.

As the team notes, much of the information in the database has been publicly available for some time, but unfortunately for researchers, not in a single place—they had to do a quite a lot of work to ensure they were not choosing a strain inappropriate for their planned work. Now, all they have to do is search a single database, the results of which will only improve as more data is added when new discoveries are found. The hope is that the database will ease and therefore improve experimental design, resulting in better results.

More information: * Database:

bioit2.irc.ugent.be/prx/mousepost/Home.php

Abstract

Mouse inbred strains remain essential in science. We have analyzed the publicly available genome sequences of 36 popular inbred strains and provide lists for each strain of protein-coding genes that acquired sequence variations that cause premature STOP codons, loss of STOP codons and single nucleotide polymorphisms, and short in-frame insertions and deletions. Our data give an overview of predicted defective proteins, including predicted impact scores, of all these strains compared with the reference mouse genome of C57BL/6J. These data can also be retrieved via a searchable website (mousepost.be) and allow a global, better interpretation of genetic background effects and a source of naturally defective alleles in these 36 sequenced classical and high-priority mouse inbred strains.

© 2017 Phys.org