TOOLBOX

Database compares protein function across dozens of mouse strains

BY **BRIANNA ABBOTT**

23 NOVEMBER 2018

A new resource details the exomes, or protein-coding portions of the genome, of 36 popular mouse strains[1].

The database, called **Mousepost**, could help researchers compare the strains and select the best one for their study — or interpret their results on one they previously chose.

Other databases, such as the Mouse Genome Informatics **Gene Expression Database** and the **Mouse Genomes Project**, also catalog the genomes of various mouse strains. But searching through them is time-consuming. Mousepost connects the gene, mutation type and change in protein function in a way that is easy to navigate.

The team chose the C57BL/6J mouse — a strain widely used in autism research — as their reference[2]. They compared RNA transcripts of its more than 20,000 genes with those of 35 other strains in the database.

Three types of mutations emerged: Some mutations result in shortened proteins, whereas others produce elongated proteins or affect single amino acids[2]. The team used a software tool called **PROVEAN** to generate a score that predicts whether a mutation changes the protein's function.

Users can search for a particular strain and compare that strain with C57BL/6J. Mousepost highlights any mutations and specifies each mutation's type, location and PROVEAN score. If a mutation changes a protein's length, Mousepost compares the length of the amino acid chain with that of the reference strain.

Researchers also can search for a specific gene or chromosome among all the mouse strains or compare two strains directly to find all of the differences between them. Each listed mutation includes links to **UCSC Genome Browser**, **e!Ensembl** and **PubMed** for related studies on that mutation and more detailed transcript information. The researchers also compared C57BL/6J with the other strains to see whether the reference has any mutations that deviate from the majority[3].

The researchers launched the database in September and described it in *Trends in Genetics.* They plan to expand it by adding more mouse strains. The team also aims to include **copy number variants** — large deletions or duplications in the genome — and noncoding parts of the genome, such as microRNAs.

REFERENCES:

1. Timmermans S. and C. Libert *Trends Genet.* Epub ahead of print (2018) **PubMed**
2. Timmermans S. *et al. Proc. Natl. Acad. Sci. USA* **114**, 9158-9163 (2017) **PubMed**
3. Timmermans S. and C. Libert. *JCI Insight* Epub ahead of print (2018) **PubMed**