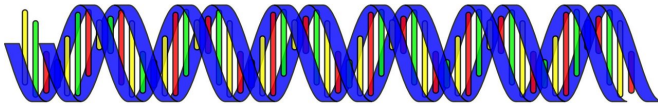


Reference genome comparison finds exome variant discrepancies in 206 genes

14 June 2021



Credit: CC0 Public Domain

In the two decades since the Human Genome Project mapped the entire human genome, improvements in technology have helped in developing updated reference genomes used for sequencing. But while the GRCh38 (hg38) human reference genome was released more than seven years ago, the older GRCh37 (hg19) reference remains widely used by most research and clinical laboratories. In a new study published in the *American Journal of Human Genetics*, researchers at the Human Genome Sequencing Center at Baylor College of Medicine identify genetic variant discrepancies between the two references, creating guidance for laboratories to take advantage of an improved human reference genome.

"There's a big push to update genomic sequencing resources to use the hg38 reference because the belief is that hg38 is a significant improvement over hg19," said Moez Dawood, co-first author of the study and student in the Medical Scientist Training Program at Baylor. "We wanted to identify the differences in sequencing readouts between the two references for labs that are still using hg19."

The Baylor researchers analyzed exome sequencing samples from more than 1,500

participants in the Baylor-Hopkins Center for Mendelian Genomics program. They found 206 [genes](#) with discordant variants between hg19 and hg38, including eight genes implicated in Mendelian diseases and 53 associated with common disease phenotypes. They found 73% of the discordant variants were clustered within sections of the [genome](#) with known assembly problems that the researchers called DISCORDANT REFERENCE PATCHES (DISCREPS).

"This study isn't a theoretical comparison of the two references; we looked at exome data from study participants and examined the impact of using the updated reference on Mendelian genes and pathogenic variants," said Dr. Aniko Sabo, a senior author of the study and assistant professor at the Human Genome Sequencing Center. "We wanted to provide the list of 206 genes enriched with discordant variants and bring this issue to the attention of the labs working on these genes."

"For variant interpretation in the 206 genes enriched for discordant variants, reference assembly differences should be accounted for in the analysis, especially when lifting over variant coordinates from one reference to the other," said Dr. He Li, co-first author of the study and a postdoctoral associate at Baylor at the time of research.

Transitioning from using the hg19 reference to the hg38 reference takes significant time and resources. Through this large-scale study of sequencing data, the researchers aim to ease the burden on labs considering the transition. The study quantifies the benefits and drawbacks of the new reference and validates its utility in a lab setting.

"It's one thing to make a better reference. It's quite another to integrate it into useful practice," said Dr. Richard Gibbs, senior author of the study, director of the Human Genome Sequencing Center and

Wofford Cain Chair and Professor of Molecular and Human Genetics at Baylor. "Some labs have been hesitant to use the new reference, but this study provides reassurance and guidance for those who are considering moving over."

More information: He Li et al, Exome variant discrepancies due to reference genome differences, *The American Journal of Human Genetics* (2021). [DOI: 10.1016/j.ajhg.2021.05.011](https://doi.org/10.1016/j.ajhg.2021.05.011)

Provided by Baylor College of Medicine

APA citation: Reference genome comparison finds exome variant discrepancies in 206 genes (2021, June 14) retrieved 25 September 2021 from <https://medicalxpress.com/news/2021-06-genome-comparison-exome-variant-discrepancies.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.