

News Release



April 16, 2019 | For immediate release

Delving deeper into individual genomic differences

Advanced technologies can characterize structural variants in human genomes more thoroughly and may improve the power for discovery in clinical genetic research.

Media Contact: The Jackson Laboratory: Sarah Laskowski, 860.837.2102, sarah.laskowski@jax.org

UW Medicine: Leila Gray, 206.685.0381, leilag@uw.edu



Iuro Motov/Getty

The researchers examined the genomes of three families, each composed of a mom, dad, and their child.

The most comprehensive view so far of the spectrum of genetic differences between individuals has been obtained using a suite of advanced genomic technologies. Routine sequencing and commonly used computer algorithms miss many of these variants.

A large, international team of researchers from the Human Genome Structural Variation Consortium conducted the study. Among the several leading contributors to the project were scientists from the Department of Genome Sciences at the University of Washington School of Medicine in Seattle.

Human genomes vary quite a bit from person to person. These differences include single nucleotide changes, which are like spelling mistakes in the DNA sequence. However, even more difference between people's genomes come from structural variants, which include additions, deletions and rearrangements of large segments of DNA.

According to the researchers, the incomplete identification of structural variants from whole-genome sequencing data has limited studies of human genetic diversity and disease association

A multiple-platform approach enabled the researchers to dive deeper than ever before to distinguish structural variants present in three families, information that could lead to determining what their functional consequences might be.

Genome sequencing has become much faster, more accurate and less expensive over the past decade. As a result, more human genomes are being sequenced, and knowledge about what those sequences actually mean for function and disease is growing rapidly.

"This work is important because it represents one of the first times that a human genome has been physically phased in order to discover and sequence-resolve structural variants. As a result, sensitivity is dramatically increased," said Evan Eichler, professor of genome sciences at the UW School of Medicine and one of the co-corresponding senior authors of the study.

It has also become clear that the more that is discovered, the more it becomes apparent how much is still unknown about human genomics. While a linear sequence of the DNA code looks tidy, genomes are actually dynamic entities that harbor considerable differences between individuals. These differences can alter traits contributing to both normal function and disease.

The genetic differences among people contribute to each person's individuality. These differences include millions of single nucleotide variants. A DNA molecule has four types of bases, abbreviated ACGT. In a certain position on their DNA, one person may have an A, another may have a C. There are also hundreds of thousands of structural variants. SVs include segments of DNA that are inserted into or deleted from the genome, segments that are duplicated, and segments that are inverted. SVs are more difficult to identify than are single nucleotide variants. That's why it has been unclear just how many SVs really exist in a person's genome.

Now a paper entitled, "Multi-platform discovery of haplotype-resolved structural variation in human genomes," published in *Nature Communications*, delves deeper into individual genomic differences than ever before. [Read the paper.](#)

The Human Genome Structural Variation Consortium work on this study was led by co-first authors Mark Chaisson, a former UW Medicine postdoctoral student in genome sciences who is now an assistant professor at the University of Southern California; Ashley Sanders of the BC Cancer Agency in Vancouver; and Xuefang Zhao, a Harvard postdoctoral fellow.

The co-senior authors are Paul Flicek, Ken Chen, Mark Gerstein, Pui-Yan Kwok, Peter Lansdorp, Gabor Marth, Jonathan Sebat, Xinghua Shi, Ali Bashir, Kai Ye, Ph.D., Scott Devine, Michael Talkowski, Ryan Mills, and Tobias Marschall

The co-corresponding senior authors, in addition to Eichler, are Jan Korbel of the European Molecular Biology Laboratory in Germany, and Charles Lee of [The Jackson Laboratory](#), an independent, non-profit genome science research institution.

The project team used a full suite of genomic technologies to extensively analyze the genomes of three family trios. Each trio consisted of a mother and father and their child. The genomic technologies include long-read, short-read, and strand-specific sequencing technologies, optical mapping and multiple computer algorithms for structural variant detection. The results present the most comprehensive catalog of SVs to date in the children's genomes. The cataloging included information on which set of parental chromosomes each SV was present.

"This study represents a conceptual advance in the field of structural variation," Eichler said. "Instead of treating a genome as a single entity, the ability to partition the sequence from the mom's and dad's chromosomes into haplotypes increases our power for genetic discovery."

He added, "I believe treating an individual genome as 6 Gbp (giga base pairs) as opposed to 3 Gbp is the future of clinical genetic research."

Overall, the researchers identified an average of 818,054 small insertions and deletions (genomic alterations that each affected less than 50 bases of DNA) and 27,622 SVs (genomic alterations that affected 50 bases or more of DNA) per genome.

Remarkably, they also found an average of 156 inversions per genome. Many of these inversions intersected with genomic regions associated with genetic disease syndromes. The researchers found that more than 100,000 variants per individual are actually missed by routine sequencing technologies and commonly used computer algorithms.

For example, 83% of the insertions identified were missed by standard short-read-calling algorithms. In fact, the true numbers of SVs in a given human genome appears to be three- to seven-fold more than most studies typically identify.

Hence, SVs constitute a large amount of genetic variation not commonly captured by current genome sequencing technologies and analytical methods. This implies that the contribution of SVs to human disease has not yet been well-quantified and the expanded SV repertoire can help identify new genetic associations to diseases and improved diagnostic yields in future genetic tests.

###

For details about UW Medicine, please visit <http://uwmedicine.org/about>.
