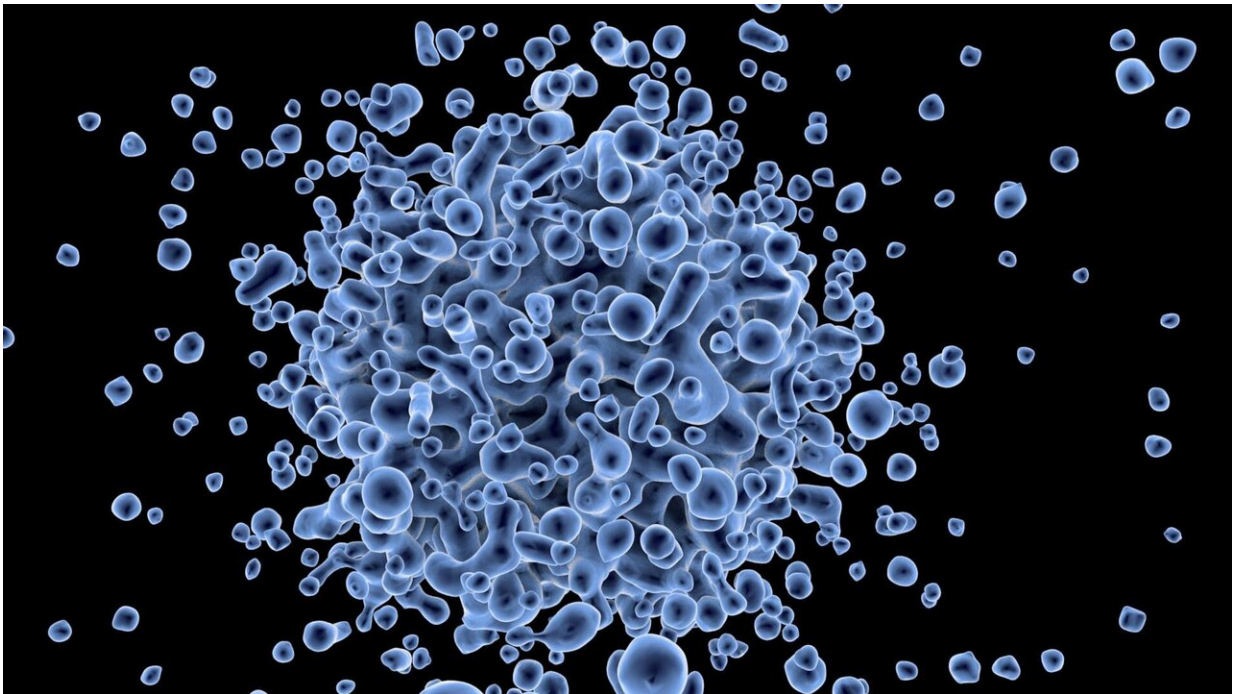


# Research team develops software that cuts time, cost from gene sequencing

December 3 2020

---



Credit: CC0 Public Domain

A team of Johns Hopkins University researchers has developed a new software that could revolutionize how DNA is sequenced, making it far faster and less expensive to map anything from yeast genomes to cancer genes.

The [software](#), detailed in a paper published in *Nature Biotechnology*, can

be used with portable sequencing devices to accelerate the ability to conduct genetic tests and deliver diagnoses outside of labs. The new technology targets, collects and sequences [specific genes](#) without sample preparation and without having to map surrounding genetic material like standard methods require.

"I think this will forever change how DNA sequencing is done," said Michael C. Schatz, a Bloomberg Distinguished Associate Professor of Computer Science and Biology and senior author of the paper.

The new process shrinks the time it takes to profile gene mutations, from 15 days or more to just three. That allows scientists to understand and diagnosis conditions almost immediately, while saving time and money by eliminating preparation and additional analysis.

"In cancer genomics there are a few dozen [genes](#) known to increase cancer risk, but with a standard sequencing run, you would have to sequence the whole genome just to read off those few genes," Schatz said, adding that adaptive sequencing allows researchers to "pick and choose which molecules we want to read and which can be skipped."

To provide a sense of how much this invention speeds up sequencing, Schatz relates it to finding a movie on Netflix. The older, standard method of sequencing would require someone to watch every second of every movie on Netflix to find what they want. Instead, adaptive sequencing eliminates hours of watching irrelevant content by quickly recognizing unwanted movies and skipping to the next entry.

The open-source software's algorithm was written by lead author Sam Kovaka, a Johns Hopkins doctoral student. Its acronymic name, UNCALLED, stands for a Utility for Nanopore Current Alignment to Large Expanses of DNA.

It took two years to code, develop and test the software, and another year to refine it enough to produce results worthy of publication, Kovaka said.

"UNCALLED allows for unprecedented flexibility in targeted sequencing," he added. "The fact that it's purely software-based means researchers can target any genomic region with no added cost compared to a normal sequencing run, and they can easily change targets just by running a different command."

The process identifies DNA molecules as they pass through tiny electrified holes, or "nanopores," inside devices called nanopore sequencers, which are smart phone-sized versions of the bulky machines used in labs. The software reads the data and checks it against a specified genome's reference sequence within a fraction of a second. Desired molecules are allowed to pass through the pore to be fully mapped. But if an undesirable molecule is detected, the software reverses the voltage in the nanopore, physically ejecting the molecule to make room for the next.

"It's like a nightclub doorman allowing desired guests on a list to enter while rejecting the rest with a Taser," Schatz explains.

The research team performed two demonstrations of UNCALLED.

The first showed that the software was able to enhance the sequencing of 148 genes known to increase cancer risk by quickly and accurately profiling all of their variants with just a single run through a portable sequencer. The software made it possible to catalog in real time dozens of complex structural mutations in the cancer genes that a standard run would have missed.

Then the team demonstrated how the software could selectively

sequence certain species collected from an environment, such as microbes living on skin or those in pond water. By rejecting molecules from known microbes (such as *E. coli*), the software was able to efficiently sequence the remaining [molecules](#), which revealed a less-understood yeast genome.

UNCALLED can operate on standard hardware used for nanopore sequencing without requiring special reagents or accelerators. The selection of genes or genomes to sequence is controlled entirely in the software and can be changed at any time.

**More information:** Sam Kovaka et al, Targeted nanopore sequencing by real-time mapping of raw electrical signal with UNCALLED, *Nature Biotechnology* (2020). [DOI: 10.1038/s41587-020-0731-9](https://doi.org/10.1038/s41587-020-0731-9)

Provided by Johns Hopkins University

Citation: Research team develops software that cuts time, cost from gene sequencing (2020, December 3) retrieved 24 April 2024 from <https://phys.org/news/2020-12-team-software-gene-sequencing.html>

<p>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.</p>
--