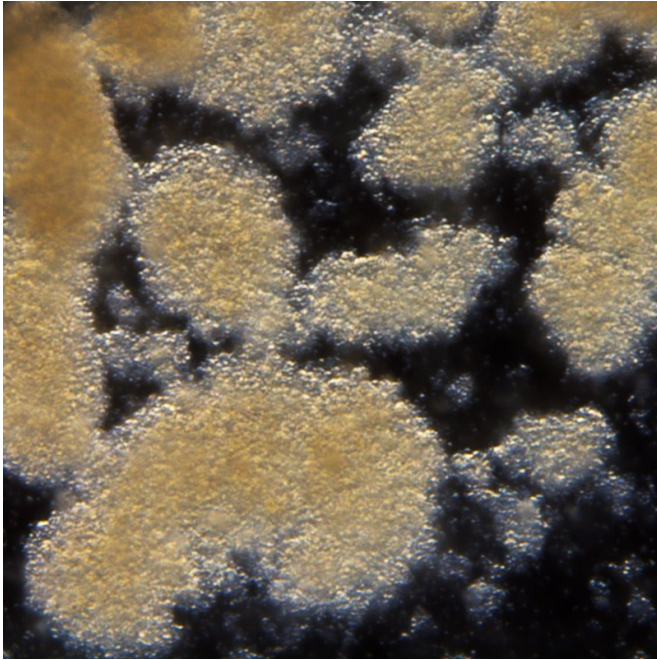


# Natural barcodes enable better cell tracking

24 April 2018, by Lindsay Brownell



The researchers used lines of human B cells like those above to test their SNP barcoding method, which accurately predicted the proportion of different donors' cells in a mixed cell pool. Credit: Wyss Institute at Harvard University

Each of us carries in our genomes about 10 million genetic variations called single nucleotide polymorphisms (SNPs), which represent a difference of just one letter in the genetic code. Every human's pattern of SNPs is unique and quite stable, as they are inherited from our parents and are rarely mutated, making them a kind of "natural barcode" that can identify the cells from any individual. A group of researchers from the Wyss Institute for Biologically Inspired Engineering at Harvard University and Harvard Medical School (HMS) has developed a new genetic analysis technique that harnesses these barcodes to create a faster, cheaper, and simpler way to track what happens to cells from different individuals when they are exposed to any kind of experimental condition, enabling large pools of cells from

multiple people to be analyzed for personalized medicine. The research is reported in *Genome Medicine*.

As the Big Data revolution in healthcare gallops apace, it is becoming possible and more attractive to perform experiments on [cells](#) from multiple people simultaneously, as differences in how the cells respond can indicate that genetic variances between the individuals are conferring some kind of effect. However, keeping track of which cells belong to which person throughout such a multiplexed experiment currently requires that a unique tag or barcode be added to each individual's cells, a time-consuming and costly process that frequently involves integrating a barcode (e.g., a unique DNA sequence) into each cell line separately so that they can identify the cells during testing. By taking advantage of all humans' unique SNP profiles, the Wyss/HMS team achieved the same cell tracking without the cumbersome labeling process.

While SNPs have been known to science for almost two decades, unlocking their utility as barcodes has proven extremely difficult. SNPs are distributed sparsely throughout the genome (approximately one SNP occurs in 1,000 base pairs), meaning that any one SNP can only distinguish between two individuals. Current, commonly used high-throughput sequencing technologies have sequencing read-lengths of less than 1,000 base pairs, making it nearly impossible to ascribe each of the sequencing reads to any particular person based on SNPs.

To overcome this problem, the team's new method combines genomic DNA extraction from a mixed pool of cells, whole-genome sequencing of the extracted DNA, and a computational algorithm that predicts the proportion of each individual within the pool based on the entire SNP allele profile of every known person's cells. Many of the [cell lines](#) publicly available for research already have whole-genome SNP allele profiles associated with them, and a given individual's profile can be determined with the

use of genotyping arrays or low-coverage whole-genome sequencing.

SNP allele profiles can be used to track cells' identities across any number of different experiments in which the pool of multiple cell samples is subjected to two or more different conditions (usually a "control" condition and an "experimental" condition), and then analyzed. Yingleong Chan, Ph.D., a Postdoctoral Fellow in the laboratory of George Church at the Wyss Institute and HMS, and his coworkers have developed an algorithm that predicts the proportions of each person's cells in the pool before and after the experiment, and compares them to determine which cells are expressed differently when exposed to the condition tested. "The change in the proportion of the individuals' cells in the experimental group when compared to the control group tells you what happened to those cells during the experiment, and whether cells from any particular person might have a genetic advantage," says Chan.

The researchers first tested their method by simulating a pool of cells and varying the number of samples, quantity of SNPs analyzed, and number of times that the pool was sequenced. They found that, over several iterations, the algorithm converged to a fixed estimated proportion for each SNP profile in the pool that closely matched the simulated proportions. The algorithm was able to accurately estimate the proportions of pools of up to 1,000 different individuals by analyzing 500,000 SNPs, and could handle samples of even more cell lines if either the number of SNPs analyzed or the depth of sequencing were increased.

Next, the researchers tested their algorithm on actual human B-lymphocytes whose genomes had been sequenced as part of the [Harvard Personal Genome Project](#), and found that it accurately predicted the proportion of the individuals within a pool of 50 different cell lines. "There are numerous experiments that this technique could be applied to," says Chan. "You can test a cancer drug against different cell lines from different people, see whether a particular patient's cell line responded well to the drug, and then use that drug for a targeted approach to treatment. We've effectively

built a discovery tool to enable personalized medicine."

The authors point out that their method will not work on samples where the different cell types come from the same person, because the SNP profiles would be identical, but it holds great promise for multiplexed testing of genetic variation among many human samples.

"Testing the effects of drugs on multiple cancer cell lines is one application that can be implemented immediately," says co-corresponding author George Church, Ph.D., who is a Founding Core Faculty member of the Wyss Institute, a Professor of Genetics at HMS, and Professor of Health Sciences and Technology at Harvard and MIT. "You can test a lot more people at once, which not only gives you more data, but translates into significant time and cost savings."

"This new technology harnesses the very core of what makes us who we are - the unique variations in our DNA - and crafts it into a tool that can accelerate discovery by obviating the need for analyzing individual responses in multiple parallel, time consuming, and expensive experiments. It also opens up an entirely new approach to personalized medicine," says Wyss Founding Director Donald Ingber, M.D., Ph.D., who is also the Judah Folkman Professor of Vascular Biology at HMS and the Vascular Biology Program at Boston Children's Hospital, as well as Professor of Bioengineering at the Harvard John A. Paulson School of Engineering and Applied Sciences.

**More information:** Yingleong Chan et al, Enabling multiplexed testing of pooled donor cells through whole-genome sequencing, *Genome Medicine* (2018). [DOI: 10.1186/s13073-018-0541-6](https://doi.org/10.1186/s13073-018-0541-6)

Provided by Harvard University

APA citation: Natural barcodes enable better cell tracking (2018, April 24) retrieved 25 June 2019 from <https://medicalxpress.com/news/2018-04-natural-barcodes-enable-cell-tracking.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.*