

# Thousands of unknown DNA changes in the developing brain revealed by machine learning

24 September 2018



Jerold Chun, M.D., Ph.D., professor and senior vice president of Neuroscience Drug Discovery at SBP  
Credit: Sanford Burnham Prebys Medical Discovery Institute

Unlike most cells in the rest of our body, the DNA (the genome) in each of our brain cells is not the same: it varies from cell to cell, caused by somatic changes. This could explain many mysteries—from the cause of Alzheimer's disease and autism to how our personality develops. But much remains unknown, including when these changes arise, their size and locations, and whether they are random or regulated. DNA technologies used to study these "copy number variations" (CNVs) in single brain cells have been limited to longer DNA sequences—those above one million base pairs.

Now, scientists at Sanford Burnham Prebys Medical Discovery Institute (SBP) have developed new single-cell analysis approaches wedded to machine learning, allowing the detection of CNVs smaller than one million [base pairs](#). This approach

has revealed thousands of previously unknown DNA changes that arise during prenatal life in the developing mouse brain. The researchers also found that these changes peaked during a key stage of brain development, implicating their creation as a regulated, deliberate process. Further research aims to elucidate the purpose and regulatory mechanisms of these CNVs. The study published today in the *Proceedings of the National Academy of Sciences (PNAS)*.

"This study fills critical holes in our understanding of copy number variations in the brain and provides important clues for further study," says Jerold Chun, M.D., Ph.D., senior author of the paper and professor and [senior vice president](#) of Neuroscience Drug Discovery at SBP. "We show that a great number of CNVs in single brain cells arise before birth as the brain begins to form and are later incorporated into the mature brain, indicating they are foundational to the brain's cellular diversity and development. We also found that these changes developmentally peak, implicating a regulatory mechanism at work."

Single-cell sequencing—a technique that allows the DNA of an individual cell to be studied—is powerful but lacks replication since each single cell is destroyed by the sequencing process. To overcome this limitation, the scientists used immune cells—that recombine DNA in stereotyped, reproducible ways—to train machine learning algorithms to more accurately recognize a genuine CNV.

"Other researchers have simply ignored all signals of small alterations because they are more likely to be incorrect," says Suzanne Rohrback, Ph.D., first author of the paper, former graduate student researcher in the Chun laboratory and current scientist at Illumina, Inc. "But characterizing what a

real change looks like allowed us to remove more than 90 percent of false positives without sacrificing the shorter CNVs, allowing the most comprehensive examination of CNVs in the developing brain."

The scientists then applied this approach to single cells during neurogenesis—when cells of the outer layer of the brain (cerebral cortex) are born. This part of the brain controls many functions, including movement and sensory information (such as how we hear and see) and consciousness.

Using this method, the researchers uncovered thousands of previously unrecognized CNVs. More than half were less than one million base pairs in length. DNA deletions were more common than duplications. And most important, while the changes were randomly distributed throughout the genome, they peaked at a predictable time: halfway through neurogenesis of the developing brain.

"These findings demonstrate that the fetal brain is patterned by a complex mosaic of myriad CNVs before birth, diversifying the individual [cells](#) that make up our brains," says Chun. "Our [brain's](#) tabula rasa may receive its first writing, which could remain with us for life, through this process."

**More information:** Suzanne Rohrback et al., "Submegabase copy number variations arise during cerebral cortical neurogenesis as revealed by a single-cell sequencing approach," *PNAS* (2018). [www.pnas.org/cgi/doi/10.1073/pnas.1812702115](http://www.pnas.org/cgi/doi/10.1073/pnas.1812702115)

Provided by Sanford Burnham Prebys Medical  
Discovery Institute

APA citation: Thousands of unknown DNA changes in the developing brain revealed by machine learning (2018, September 24) retrieved 23 May 2019 from <https://medicalxpress.com/news/2018-09-thousands-unknown-dna-brain-revealed.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.*