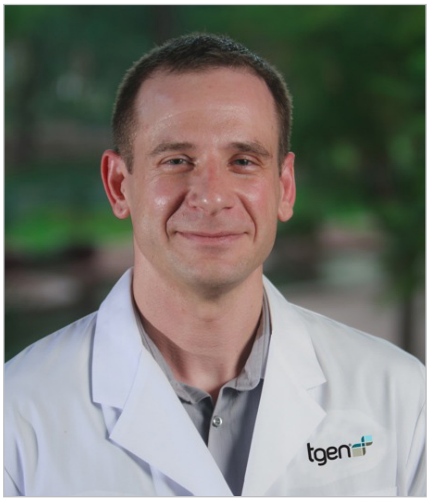
Single-Cell Transcriptome Sequencing, the Internet, and Memory

TGen researchers are combining crowd-sourcing and single-cell RNA sequencing on HiSeq® Systems to adequately power genetic studies of memory.



Matt Heuntelman, PhD, is Associate Professor in the Neurogenomics Division of TGen, and founder and Principal Scientist of the MindCrowd project.

Introduction

For hundreds of years, the brain has mystified and intrigued scientists. There is good reason for it. This three pound cluster of 100 billion neurons and a host of support cells is perhaps the only organ honed by natural selection that can contemplate itself. Neuroscience has come a long way in understanding how neurons and other brain cells do, and sometimes don't, work together to send signals and process information. Still, if the brain was a mystery novel and scientists the beleaguered detective working to crack the case, the answer of whodunit is still many chapters away from being solved.

That hasn't stopped Matt Huentelman, PhD, a neurogeneticist at the Translational Genomics Research Institute (TGen), a non-profit in Arizona that seeks to unravel the genetic components of common and complex diseases. Dr. Huentelman wants to understand the genetic and neurological variables that affect learning and memory, both in healthy humans and in those with diseases such as Alzheimer's disease. As part of this work, Dr. Huentelman has become an expert at successfully sequencing very low input samples and other difficult to sequence samples. He's become particularly adept at performing RNA sequencing of these sample types, and more recently has begun performing single-cell transcriptome sequencing to understand the role of certain brain cells in memory. Dr. Huentelman relies on Illumina HiSeq systems for nearly all of his work.

His latest project is MindCrowd, which seeks to use the power of crowdsourcing to get 1 million people around the world to take fun, short tests evaluating learning and memory. Then, in a later phase of the study, a subgroup of these users will be asked to complete more intensive questionnaires and tests, as well as submitting a sample of their DNA.

Identifying Genetic Variants Involved in Episodic Memory

Dr. Huentelman has never shied away from biology's tough questions. After receiving his degree in biochemistry from Ohio University, he completed his PhD in physiology and genomics at the University of Florida in the lab of Mohan Raizada, PhD. There, he focused on how the brain could control various aspects of the body's physiology, such as blood pressure. In 2004, Dr. Huentelman arrived at TGen as a postdoc and, two years later, started his own lab there. In 2006, Dr. Huentelman was part of a group at TGen that completed one of the first genome-wide association studies (GWAS) on human episodic memory.

Episodic memory is a long-term form of memory also described as autobiographical memory. This type of memory functions in the hippocampus and other structures of the temporal lobe. One way to think of it, Dr. Huentelman says, is to ask yourself where you were during a significant event in your life.

"It depends on what generation you're in, but for me, it's what were you doing when the Challenger explosion happened? Or what were you doing on 9/11?" Dr. Huentelman said. "When you're recalling that information you're using a form of episodic memory."

The goal of the 2006 GWAS study was to identify genetic variants that affected an individual's episodic memory. About 1000 individuals participated in the study, which allowed the researchers to link a single nucleotide polymorphism (SNP) in the kidney and brain expressed protein (KIBRA) gene to episodic memory performance[1]. KIBRA has been linked to memory and synaptic plasticity in subsequent studies.

> "It was clear that the standard approach of having research subjects visit our laboratory to complete the memory test face-to-face wasn't going to work. We realized that we might be able to conduct this study more effectively using the Internet."

The groundbreaking research was one of the first GWAS studies to look at human episodic memory. Yet, as time went on, Dr. Huentelman realized that the study was dramatically underpowered.

"We knew we needed more samples than that, and I was very interested in obtaining much larger sample numbers for the study," Dr. Huentelman said. "To do so, it was clear that the standard approach of having research subjects visit our laboratory to complete the memory test face-to-face wasn't going to work. We realized that we might be able to conduct this study more effectively using the Internet."

And thus MindCrowd was born. TGen has partnered with a pair of biologists at the University of Arizona to create short neuropsychological tests that anyone can take at www.mindcrowd.org. One test evaluates reaction time, while a second one assesses episodic memory.

"The tests are short and useful because we realized that if we asked people to come to our study and take a two-hour exam, we wouldn't have many participants," Dr. Huentelman said. "Those who would complete a longer exam would probably be outliers in the population anyway. I wouldn't sit down for two hours, and I'm a neuroscientist."

Nearly 50,000 people have already taken the 10 minute MindCrowd tests. Dr. Huentelman and his team are working to translate the tests into six different languages to broaden the base of Internet users who can participate. Dr. Huentelman hopes that he can persuade at least 1 million people to take these short, preliminary MindCrowd tests as part of Phase One of the study.

Discovering the Genetics Underlying Demographic Differences

The large number of initial users has allowed Dr. Huentelman to broaden the scope of his research. Instead of simply looking at the genetic contributors to learning and memory, he's now also examining the demographic variables that might affect memory. More than just where you live, demographics covers everything from lifestyle choices—such as the quality of your diet, how much you exercise, and what type of job you have—to things like handedness, ethnicity, and age.

Demographic analysis hasn't been done before on such large sample sets, Dr. Huentelman said. "Neuroscience, particularly the study of human brain performance on memory tests, hasn't done a good job of studying large numbers of demographic factors that might be associated with memory loss, he added."

Phase Two will involve a subset of these test takers, who will take more detailed tests at their leisure, as well as mail in a DNA sample. Dr. Huentelman wants to perform whole-exome sequencing on these samples to identify the genetic variants that might be associated with learning and memory. Based on his 2006 GWAS study and other research, Dr. Huentelman believes that rare genetic variants will play an important role in this relationship.

> "The HiSeq systems offer lots of advantages, including lower cost and higher speed of sequencing. They enable us to use molecular barcoding and run almost 96–100 samples in a single sequencing run for whole-exome sequencing."

"Our specific focus is to gather a large group of people who are well-characterized on at least one type of memory," Dr. Huentelman said. "We'll then perform genomic analysis of DNA from study members that are interesting based upon their phenotype and demographic profile."

Extracting DNA and sequencing user-acquired samples isn't always easy. Sample quality might be low, which can result in very low amounts of genetic material for scientists to sequence. Luckily, Dr. Huentelman has lots of experience working with difficult-to-sequence samples on HiSeq Systems.

"We've been using Illumina systems in my laboratory since we started sequencing in 2006," Dr. Huentelman said. "The HiSeq systems offer lots of advantages, including lower cost and higher speed of sequencing. They enable us to use molecular barcoding and run almost 96–100 samples in a single sequencing run for whole-exome sequencing. That's exciting and really important when you're sequencing hundreds of thousands of samples."

Understanding Memory, One Brain Cell at a Time

In addition to the MindCrowd project, Dr. Huentelman has been studying the brain using single-cell RNA sequencing. Historically, DNA sequencing has required large numbers of cells to obtain adequate genetic material. For a time, researchers believed that all cells in the body contained the same DNA, so single-cell sequencing didn't seem to be a useful tool for human samples. Recent discoveries of mosaicism in humans, especially in brain tissue, along with other advances have forced scientists to rethink the utility of single-cell sequencing to understand human biology. RNA sequencing in particular allows researchers to identify new cell types and understand how different genes are activated in response to a cell's microenvironment. Dr. Huentelman and other scientists feel this information will be important in understanding how the brain works.

> "We can look in the brain region and we might see 10 different types of neurons, just based on morphology. But is that the actual picture we see after we study the molecular genetics of each cell in the region? That's why we're really excited about single-cell sequencing."

The brain controls body and behavior using different functional units, circuits of brain cells that function in coordination with each other. Dr. Huentelman wants to dissect those circuits by capturing all of the information from the individual cells that are involved in particular behaviors. RNA sequencing gives him information about which genes are actively being transcribed in specific cells. Because not all cells in particular circuits are going to behave in the exact same way, Dr. Huentelman wants to know which cells contribute to the circuit and how they work together. It could be that 10 cells in one area of the brain are up-regulating genes A, B, and C, while 40 other cells are up-regulating genes X, Y, and Z, and so on. Dr. Huentelman believes that it's the diversity within the circuit that's really important for the biological response.

Single-cell sequencing is valuable because scientists still don't know all of the different types of brain cells that exist in certain regions of the brain. To identify those cells, scientists might need to sequence each individual cell and reconstruct those populations.

"Right now, we can look in the brain region and we might see 10 different types of neurons, just based on morphology," Dr. Huentelman said. "But is that the actual picture we see after we study the molecular genetics of each cell in the region? That's why we're really excited about single-cell sequencing."

Using laboratory animals like mice and rats, Dr. Huentelman and colleagues have been trying to molecularly tag the circuits in a particular brain region that are activated when the animal encodes a memory. With the help of the tags, he can identify those cells, take them out of the brain region using laser capture, and perform single-cell RNA sequencing. This will tell him what molecular approaches each cell type uses to encode memory. Here too, Dr. Huentelman's lab uses HiSeq systems and various Illumina library preparation kits.

"One of the reasons we like Illumina is because we have all this historical sequencing data that we can compare to," Dr. Huentelman said. "We use various library preparation kits to perform the work. For whole-genome sequencing, we use the TruSeq® PCR-Free DNA Library Prep kit, and TruSeq or Nextera® Rapid Capture Exome Library kits for whole-exome sequencing. For single-cell RNA sequencing, we use different library preps depending on what we're interested in, messenger RNAs or the whole transcriptome.

From Big Data to single cells, Dr. Huentelman's lab has embraced a range of neurogenomics technologies to help in understanding the human brain. Figuring out ways to deal with large numbers of variable quality genetic samples remains a large hurdle and will be crucial to move the field forward.

"Collecting data over the Internet is fine, but how do we perform genomics studies when sample quality and quantity are variable?" Dr. Huentelman asked. "We'll need to obtain samples from everybody, yet people often won't want to go in to see their doctor for a blood draw just to participate in a research study. We need to figure out if it's possible to conduct good genomics studies from a saliva sample, a urine sample, or a single drop of blood. If we could work through those issues, then we could have people collect their own samples at home and send them to us. That would enable us to combine Internet and genetic data to perform these types of large studies."

References

1. Papassotiropoulos A, Stephan DA, Huentelman MJ, et al. Common Kibra alleles are associated with human memory performance. Science 2006; 314: 475–478.

Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY.

Pub. No. 1070-2014-010 Current as of 16 February 2015