Sequencing Unlocks the Secrets Behind Neurological and Psychological Diseases

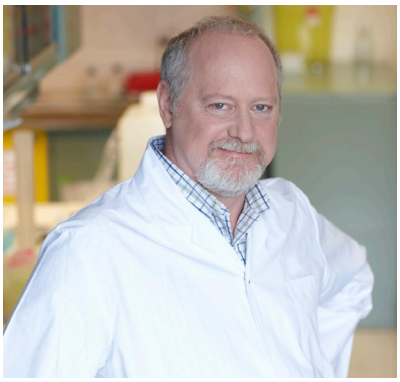
Dr. Guy Rouleau uses Illumina sequencing to pinpoint the *de novo* mutations that appear to be responsible for several neurodevelopmental diseases.

Introduction

Guy Rouleau, MD, PhD, has dedicated his career to understanding the genetic basis for many of today's most vexing neurological and psychiatric diseases, including amyotrophic lateral sclerosis (ALS), stroke, epilepsy, autism, schizophrenia, and bipolar disorder. Among Dr. Rouleau's most notable achievements are his landmark contributions to the identification of more than 20 disease-causing genes and his discovery of new mutational mechanisms.

Over the course of his career, Dr. Rouleau has used a number of genetic analysis technologies to uncover these genes and understand the roles they play in neurological and neurodevelopmental disorders. As Director of the Montreal Neurological Institute and Hospital (known as the "Neuro") at McGill University and McGill University Health Centre, the pioneering clinician-scientist is using Illumina sequencing systems to break new ground.

iCommunity spoke with Dr. Rouleau about his fascination with unlocking the mysteries of the human brain and how Illumina next-generation sequencing (NGS) systems are making his research possible.



Guy Rouleau, MD, PhD, FRCP(C), OQ is Director of the Montreal Neurological Institute and Hospital. He is also Chair of the Department of Neurology and Neurosurgery at McGill Univeristy and Director of the Réseau de Médecine Génétique Appliquée (FRSQ).

Q: What inspired you to study brain diseases?
Guy Rouleau (GR): As a student, I believed that the brain was the next frontier and I thought it would be a great subject to study. We have a very poor understanding of how the brain works. It's complicated anatomically and functionally. The mystery behind how the brain functions fascinated me initially and it still intrigues me today. That's why I became a clinical neurologist.

Q: What neurological and neurodevelopmental diseases are you studying?
GR: I've spent several years studying ALS, also known as Lou Gehrig's disease. We don't know what causes ALS, but we do know that there are some familial forms of it. I became interested in identifying genes that explain some of these familial cases and in how these gene mutations predispose someone to developing the disease.

I am also studying several neurodevelopmental diseases, such as schizophrenia, autism, and intellectual disability. I'm interested in understanding why these diseases are so common and why they have such a strong genetic basis. The fact that it's difficult to find the genetic basis of these diseases is what led me to explore a *de novo* mutation hypothesis. The notion is that a new mutation occurs in the sperm or the egg that forms in the individual. As the brain structure forms, a mutation in any of a large number of genes, would predispose someone to developing one of these neurodevelopmental diseases. That hypothesis was already accepted for intellectual stability, and our work did a lot to suggest that it's also an important mechanism in autism and schizophrenia. Pinpointing the mechanism enabled us to identify a number of genes that predispose someone to these diseases.

We are also looking at somatic mutations as a cause of a neurological disease. While somatic mutations have been investigated in cancer, they're not very well explored in neurological disease.

Q: How have the tools you've used to answer these research questions changed over time?
GR: I've always tried to be ahead of the curve when it comes to new technological developments, and to be able to use the new technologies to ask new research questions. I was one of the first to use PCR in the lab where I did my graduate studies. I transitioned from using restriction fragment length polymorphism (RFLP) to micro-satellite analysis, and later to the use of arrays to study single nucleotide polymorphisms.

As the technology has evolved, so has our ability to ask new research questions. Very early on, we jumped onto the NGS bandwagon. A nice example is our study of *de novo* mutations, where you have to sequence billions of bases. That's not practical to do with anything other than NGS.

Q: What type of sequencing studies are you performing today?
GR: We're performing trio studies of —children with disease and their parents to look for *de novo* mutations involved in predisposition to certain diseases. We're also using sequencing to study more refined phenotypes in schizophrenia and autism.

We're performing a lot of exome sequencing of small families where there's a high prevalence of bipolar disorder, restless leg syndrome, or ALS. In those cases, there are too few samples to adequately perform linkage analyses. Exome sequencing gives us a better chance to identify predisposition variants for these diseases.

Q: How do you collaborate with researchers within and outside of McGill University?
GR: We've created a biobank of about 70,000 unique DNA samples collected over the past few decades from multiple sources, but mostly through collaborations. The biobank offers easy access to any researcher who wants to use it, enabling me to collaborate broadly with different groups all over the world.

I'm working with a large group in France for schizophrenia studies, and with groups in France, Portugal, England, and the U.S. on some ALS studies. I'm collaborating with colleagues at McGill on clinical research studies using the variants we've discovered to identify ALS, schizophrenia, and autism cases.

> "Illumina systems have completely transformed our research. What we're doing now we couldn't even imagine five years ago."

Q: Where is the sequencing of these samples performed?
GR: The bulk of our sequencing is performed at the Genome Quebec and McGill Genome Center in Montreal, which has a MiSeq® and 16 HiSeq® Systems. I've also used the Illumina Genome Network (IGN) service for some of our genome sequencing.

Both groups are easy to work with. We decide what we want sequenced and the questions we want to answer. We send our samples to the genome center or IGN. They perform the sequencing, give us back the data, and we deal with the data analysis. In our lab, we have quite a bit of sequence—about 2,700 exomes sequenced and about 500–600 genomes. Through NGS, we've already identified quite a few different disease genes in essential tremor, ALS, and in a number of rare Mendelian disorders.

> "Through NGS, we've already identified quite a few different disease genes in essential tremor, ALS, and in a number of rare Mendelian disorders."

Q: Has all of your sequencing been performed using Illumina technologies?
GR: At the beginning, we used both ABI (Applied Biosystems) and Illumina sequencing. It turned out that the Illumina sequence data was easier to interpret and worked better for us. I'd say 98% of everything we've sequenced to date is with Illumina sequencing systems and we've had samples sequenced on all the different generations of Illumina systems. Illumina has really been the workhorse – the major source of sequencing for our lab.

Illumina systems have completely transformed our research. What we're doing now we couldn't even imagine five years ago. The ability to be able to quickly sequence massive amounts of DNA has changed the kind of questions we can ask.

Q: What is superior about Illumina sequencing technology?
GR: When you get an exome or a genome sequenced, you have basically billions of fragments. The quality of the sequence we obtain from Illumina systems is better, with fewer errors and artifacts, decreasing the noise in analysis. You also need efficient, high-quality bioinformatic tools to extract the information that you need. There are a wide range of bioinformatics tools available to analyze Illumina sequencing data. They are easy to use and very efficient.

Q: What kinds of data sets do you create for your research?
GR: We refine the biobank data into two data sets. One is more clinically oriented and is refined into phenotypes and subphenotypes. The other is a research data set refined into different kinds of mutations, such as copy number (CNV), missense, intronic, and noncoding regulatory variants. The art is further refining the two data sets, putting them together, and seeing how we can derive new information from them.

Q: What sequencing applications do you use the most?
GR: It entirely depends on what we think is the genetic architecture of the entity we're looking at. We use different methods to address different questions. For example, there's one project where we looked at the environmental effects

on the *de novo* mutation rate. For that study, we performed whole-genome sequencing to increase the number of variants. In another project, we used exome sequencing to identify disease-causing genes with *de novo* mutations. To see if the phenotype is expressed through differences in gene expression levels, we used RNA-Seq. For each disease and each question, there's a different approach and sometimes modifications in those approaches.

> "The quality of the sequence we obtain from Illumina systems is better, with fewer errors and artifacts, decreasing the noise in analysis."

Q: What are the next steps in your research?
GR: We're working on several avenues, so I'll talk about just two of them. For the first, we'll need methods and/or technologies that enable us to look at different kinds of mutations. The methods we use now are good for certain kinds of mutations, such as CNVs, but other types of mutations might be present. The question is how can we improve the detection of the mutations that we're currently missing? I suspect the answer will be technological, but it might also involve a new bioinformatics approach.

We also want to explore the correlation between genotype and phenotype. We're finding a lot of gene variants in the disease subgroups and want to link them to more refined, very detailed disease phenotypes. If we look at the various subphenotypes, we'd like to see if some kinds of variants are associated with specific segments of the phenotypic spectrum.

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. 1070-2014-005 Current as of 24 April 2015

illumina®