illumina®

Illumina Case Study: Dr. Christopher E. Mason, Ph.D.

Investigating Genome Regulatory and Functional Elements with Whole-Genome Sequencing

Christopher E. Mason, Ph.D., uses the Illumina Genome Network (IGN) to quickly and reliably examine multiple layers of human biology.

Introduction

Christopher E. Mason, Ph.D., is an Assistant Professor of Computational Genomics in the Department of Physiology and Biophysics and at the Institute for Computational Biomedicine at Weill Cornell Medical College. He leverages advances in genome technology and our current understanding of the genome to examine the critical structures of the human genome and how they change in disease, with particular emphasis in neurogenetics and cancer biology. "The challenge that we face is determining the molecular recipe for the synthesis of an entire human being," says Dr. Mason. "The instructions are present in a single cell—in the embryo. The hard and critical question is, 'How do you go from one cell to ten trillion cells?'"

The Mason Lab uses Illumina products and the IGN service in their studies of human genome disease states and evolution. "We didn't know all the letters of the genome until 2001, and only until the last few years have we started to get a handle on how many of these genes are actually transcriptionally active, potentially regulatory, or functional," explains Dr. Mason. "We've known for decades that it's more than just the coding regions of the gene. But we can only now start to tease some of those out."

Using a Multimodal Approach to Understanding Regulation and Variation

"Because the biology is complex, we use a multimodal approach that includes Illumina MiSeq® for DNA sequencing; HiSeq® for whole-genome sequencing; and TruSeq® for RNA sequencing, transcriptional profiling, and epitranscriptomic site detection[1] to examine disease and answer evolutionary questions," says Dr. Mason. "There's DNA, RNA, and proteins, and at every level there is regulation and modification. There's post-transcriptional and post-translation, pre-transcriptional, and pre-transcriptome modification. There's RNA editing that can change the information content of what the gene is supposed to look like." Important biochemical or biological change within an organism could be anywhere from the DNA all the way out to the metabolite. "Since you can't know *a priori* where the most important function or mechanism is at play within a cell, we try to examine multiple layers of the biology simultaneously." Using this approach, Dr. Mason can confirm mutations observed in the DNA or gene expression with RNA sequencing or DNA methylation sequencing.

Christopher E. Mason, Ph.D., is an Assistant Professor in the Department of Physiology and Biophysics and the Institute for Computational Biomedicine at Weill Cornell Medical College.

"Having the deep sequencing coverage with RNA-seq, we've discovered, and have been validating, tens of thousands of novel genes."

Refining Methods to Better Understand the Genome

"We know that there's a lot more than just the coding sequences in 2% of the genome. The ENCODE data[2] and our own work[3] provide evidence that the regulatory pieces are, without question, regions we have to examine," advocates Dr. Mason. "The GENCODE gene annotation set shows the presence of more non-coding RNAs than coding RNAs, making exome sequencing an incomplete method for understanding the human genome…I would prefer to perform whole-genome sequencing and spend three times as much money to get 50 times as much information, because if you find something interesting you have to go back and sequence it. It's a worthwhile investment in the future compatibility and utility of the data." As the costs of whole-genome sequencing continue to decrease, Dr. Mason suspects that the exome will eventually fade away. "But targeted amplicon resequencing, or TruSeq sequencing, will still have a place because at some point you will find something interesting and your question will be, 'In how many of the thousand samples that I have in my freezer do I see these same 10 mutations?'"

IGN Service for Faster and Easier Sequencing Projects

To accomplish some of his research goals, Dr. Mason has turned to the Illumina Genome Network (IGN) service. "IGN is easy to use and saves time on library preparation as well as informatics. I send in DNA and get back a fully processed and annotated genome in the form of sorted, completed BAM files and variant calls," explains Dr. Mason. When the Mason lab compared the variant calls with the GATK pipeline they saw 90–91% concordance of the variant calls. "The data we're getting back is comparable to what we'd get if we did our own analysis," states Dr. Mason. "Also, we've been able to go down to 500 nanograms of DNA input for whole-genome sequencing*, which is very nice because for clinical samples that's often all we have." Dr. Mason adds, "the Illumina IGN service is fast. We sent DNA and 11 days later we got a full, complete processed genome, which was cool."

Opening New Windows on Biology

"Illumina technology opens new windows into understanding the underlying biology. Having the deep sequencing coverage with RNA-Seq, we've discovered, and have been validating, tens of thousands of novel genes. We've been watching the frequency of mutations change within the same patient over time. Illumina's newer technology and deep sequencing lets us see things we could never see before and ask questions that we couldn't even imagine that we could ask," concludes Dr. Mason. "IGN's been great and we've already seen some interesting things with the data. I'd be happy to recommend it."

Learn more about IGN at www.illumina.com/ign

Summary

Overview

Christopher E. Mason, Ph.D., Weill Cornell Medical College, studies disease states and evolution to in an effort to understand the regulatory and functional elements of the human genome.

Challenge

Biology is complex, with modifications occurring at the DNA, RNA, and protein level, making it difficult to identify mutations.

Solution

Dr. Mason uses a combination of Illumina HiSeq and MiSeq sequencers and IGN services to efficiently manage his many sequencing projects.

Benefits

Illumina's IGN service offers a fast and easy method for outsourced sequencing, providing meticulous quality control and fully processed and annotated data.

"IGN is easy to use and saves time on library preparation as well as informatics...we've already seen some interesting things with the data."

References:

1. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, et al. (2012) Comprehensive analysis of mRNA methylation reveals pervasive adenosine methylation in 3' UTRs. Cell 149(7): 1635-46.
2. www.nature.com/encode/#/threads
3. Mason CE, Shu FJ, Wang C, Session RM, Kallen RG, et al. (2010) Location analysis for the estrogen receptor-a reveals binding to diverse ERE sequences and widespread binding within repetitive DNA elements. Nuc Acids Res PMID: 20047966.

* The low-sample input IGN program is currently in early access only.

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. 770-2012-059 Current as of 19 November 2014

illumina®