# Riding the Wave: An Endocrinologist's Perspective on NGS Technologies

Dr. Rory Clifton Bligh and his team at the University of Sydney are using the MiSeq® System and TruSeq® Custom Amplicon Assay to discover causative variants of heritable endocrine cancer.

## Introduction

Multiple endocrine neoplasia (MEN) affects about 1 in 30,000 people.[1] These rare heritable syndromes involve tumors (benign or malignant) in at least 2 endocrine glands. If cancerous, they can be life threatening. Researchers have been making strides in understanding the genetics underlying these and other heritable endocrine tumor syndromes. In the last 5 years, next-generation sequencing (NGS) systems such as the MiSeq System have enabled clinical researchers to perform multiplexed sequencing of 20–30 genes per sample uncovering numerous candidate genes.

NGS technologies are fueling the research of Rory Clifton-Bligh, MBBS, PhD, who began studying the molecular mechanisms of thyroid growth and function during his doctoral studies at the University of Cambridge. In addition to a busy clinical practice, Dr. Clifton-Bligh is an Associate Professor of Medicine at the University of Sydney. His research team is using the MiSeq System to understand the genetics of endocrine neoplasms and metabolic bone disease mechanisms. So far, they've uncovered "more findings than we could possibly assess by traditional methods alone," according to Dr. Clifton-Bligh.

iCommunity spoke with Dr. Clifton-Bligh about how his team relies on Illumina technologies to discover causal variants and potential diagnostic and therapeutic targets for hereditary endocrine disorders.

Q: What are your lab's research goals?
Rory Clifton-Bligh (RCB): Our primary research goal is to understand genetics as a physiological basis of neuroendocrine cancer. This encompasses heritable cancers and heritable tumor syndromes, including multiple endocrine neoplasia types 1 and 2 (MEN1 and MEN2), hereditary paraganglioma-pheochromocytoma syndromes (PGL/PCC), and some of the hereditary pituitary disorders. Our lab also studies tumor biology of nonheritable tumors, such as thyroid cancer and adrenal cancer. We are part of the cancer division of the Kolling Institute of Medical Research, where we have access to a large bank of sporadic and familial endocrine tumor samples that have been stored over many decades and have reasonable clinical annotation.

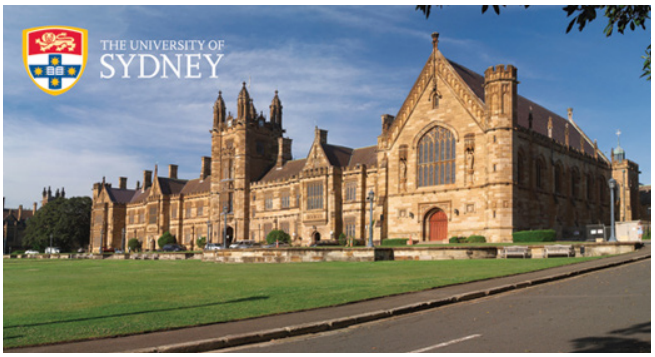
Q: What can you learn by studying the hereditable basis of endocrine disease?
RCB: Studying the hereditable basis of endocrine disease enables us to understand the genomic disorganization of these tumors at their most basic level. For instance, pheochromocytoma and paraganglioma tumors are relatively uncommon, but are the most highly heritable neoplasms known. In fact, up to 40% of people presenting with these tumors carry a germline variant imparting hereditary predisposition.

By knowing what the initial genetic insult is, we can follow the path forward to tumor development and understand the steps along the way. We've already identified useful targets for diagnosis and treatment of endocrine disorders.

Q: How have NGS and other technologies contributed to these advances?
RCB: It's become an incredibly fast moving field, with more and more genes discovered by NGS. In the last 5 years alone, the major advances in the new age of tumor biology have been fueled by NGS, including RNA-Seq, as well as DNA methylation analysis, CNV array, and mass spectrometry technologies. We feel like we're surfing a wave of discovery at the moment with the hereditary endocrine neoplasms. It's a wild ride and very exciting.

Q: What does NGS enable you to discover that you couldn't see using older technologies?
RCB: NGS is enabling us to discover variants in several ways. Our conditions are genetically heterogeneous, so we need to assess multiple genes in every sample. This time consuming process increases the probability of a false negative result. With NGS, we have the ability to sequence large numbers of genes in multiple samples on a single assay simultaneously.

We are also interested in studying whether gene inactivation in some endocrine neoplasia syndromes might occur through postzygotic mosaicism. To this end, NGS brings the advantage of being able to




**The University of Sydney was founded in 1850 and was the first university in Australia. It now has a student population of more than 50,000.**

For Research Use Only. Not for use in diagnostic procedures.

perform ultradep sequencing of 600x and higher, using the MiSeq System and TruSeq Custom Amplicon Assay.

We have access to the Kolling Institute's large collection that spans 20–30 years of heritable endocrine tumor samples from afflicted families. Previous technology approaches often failed to provide an answer to what was occurring genetically in these samples. We're steadily discovering causative variants using NGS to sequence samples from this tissue bank. That's been exciting.

> "In the last 5 years alone, the major advances in the new age of tumor biology have been fueled by NGS, including RNA-Seq, as well as DNA methylation analysis, CNV array, and mass spectrometry technologies."

Q: How is the MiSeq System enabling your studies?
RCB: We're using TruSeq Custom Amplicon to develop custom amplicon panels of 20 or more genes for NGS on the MiSeq System. The MiSeq System is fast, efficient, and enables us to perform ultradeep sequencing that is critical for the success of our research.

We had experience using other sequencing systems for our studies, but the MiSeq System has been the best fit for our lab. When we obtained it 4 years ago, we were concerned about volatility and potential contamination issues and were looking for a system that, from a design perspective, would limit the potential of cross-contamination between tissue samples. We felt that the MiSeq System performed the best in this regard.

> "TruSeq Custom Amplicon Assay works well and we're comfortable with the chemistry, the readout, and the bioinformatics processing."

Q: Have you used DesignStudio™ to prepare your TruSeq Custom Amplicon panels?
RCB: DesignStudio is easy to use in preparing our custom amplicon panels and more importantly, it works well and is robust. It has good coverage of most of the genes that we've looked at so far.

Q: What do you like about the TruSeq Custom Amplicon Assay?
RCB: TruSeq Custom Amplicon Assay works well and we're comfortable with the chemistry, the readout, and the bioinformatics processing. Creating custom amplicon assays is an iterative process, and we've created various custom designs, adding new genes, and sometimes adjusting the design slightly. We have limited experience with the TruSeq Amplicon-Cancer panel, mostly because we are obtaining such good results from our own custom designs. The TruSeq Amplicon-Cancer panel was updated recently and is a method that we could look at using again in the future.

Q: What bioinformatics challenges do you face in performing these studies?
RCB: Overall, the MiSeq System has been extraordinary in capturing the low-hanging fruit. The trap for beginners in amplicon sequencing is forgetting that not every amplicon is covered evenly for every sample. To account for this, our bioinformatics processing includes an algorithm that allows us to see, for any individual sample, what was actually covered in each individual run.

We're still working on how best to use NGS and bioinformatics processing to detect larger deletions. We have been able to detect deletions of up to 26 bp and maybe 1% of cases are genetically inactivated through larger deletions—either whole-exon or whole-gene deletions. Our bioinformatics processing is not yet up to the standard of detecting those large deletions, for which we still rely on multiplex ligation-dependent probe amplification (MLPA).

> "The MiSeq System is fast, efficient, and enables us to perform ultradeep sequencing that is critical for the success of our research."

Q: What do you use for your bioinformatics processing?
RCB: We're not a bioinformatics laboratory, so we're choosing software tools that are easy to use. We're using MiSeq Reporter for alignment and variant calling, and then annotate the variants using Annovar. We've got a few inhouse algorithms that we use, however MiSeq Reporter has been very robust for our purpose. The Integrative Genomics Viewer (IGV) is also user-friendly and we use it to visualize the data amplicon by amplicon.

Q: How many samples are you sequencing at one time?
RCB: We run 48–96 samples per run. We don't run less than that because it's not cost effective. We're slightly cautious, because every sample has a high probability of containing a pathogenic variant and we have to interpret every sample on its merits. Given that our custom amplicon panels consist of 20–30 genes, we sequence 48–96 samples per run at more than 1000×. As our collaborators frequently tell us, we're often using a Porsche to do the local shopping. We've got overkill for depth of coverage, but it has suited our needs.

Q: Do you sequence FFPE samples with the MiSeq System?
RCB: We have sequenced FFPE DNA samples with success in our research studies. The MiSeq System is a promising tool in sequencing FFPE samples to obtain good coverage of a larger range of genes in single assay runs.

For Research Use Only. Not for use in diagnostic procedures.

Q: How quickly did your laboratory incorporate the MiSeq System into its workflow?
RCB: We went from no NGS knowledge to running the MiSeq System competently within 2–3 sequencing runs. Ease of use is an advantage of the MiSeq System and one of the reasons we chose it. We like the chemistry and it's straight forward.

> "That's a year's worth of sequencing work compressed into a week by using the MiSeq System."

Q: What is your turnaround time for sequencing amplicon panels with the MiSeq System?
RCB: It takes us about 2 days to prep libraries and load the system; the sequencing, alignment, and variant calling then takes ~27 hours (all performed on the MiSeq System). Annotation of variant calls can be performed in 1–2 hours. Interpretation of annotated output generally takes us about a week.

Therefore, with the MiSeq System, our comprehensive panel takes about 10 days of processing time to interrogate 20 different genes in 48–96 samples. We perform a large amount of Sanger sequencing and it would take us more than a year to sequence that many samples one by one. That's a year's worth of sequencing work compressed into a week by using the MiSeq System. Such a leap forward in productivity is something you rarely see in science. It's amazing.

## References

1. Genetics of Endocrine and Neuroendocrine Neoplasias—for health professionals (PDQ). National Cancer Institute. www.cancer.gov/types/thyroid/hp/medullary-thyroid-genetics-pdq. Accessed October 15, 2015.

**Learn more about the Illumina products and systems mentioned in this article:**

MiSeq System, www.illumina.com/systems/miseq.html

TruSeq Custom Amplicon Assay, www.illumina.com/products/truseq_custom_amplicon.html

DesignStudio, www.illumina.com/informatics/research/experimental-design/designstudio.html

TruSeq Amplicon-Cancer Panel, www.illumina.com/products/truseq_amplicon_cancer_panel.html

MiSeq Reporter Software, www.illumina.com/systems/miseq/software/miseq-reporter.html

Integrative Genomics Viewer, www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/integrative-genomics-viewer-1886885.html

For Research Use Only. Not for use in diagnostic procedures.

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

**For Research Use Only. Not for use in diagnostic procedures.**

© 2016 Illumina, Inc. All rights reserved. Illumina, DesignStudio, HiSeqX, MiSeq, TruSeq, and the pumpkin orange color are trademarks of Illumina, Inc. and/or its affiliate(s) in the U.S. and/or other countries. Pub. No. 770-2015-025-B Current as of 07 November 2019

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGACG
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA