

Making Discoveries While Becoming Proficient in Next-Generation Sequencing

Researchers identify novel cancer biomarkers that Sanger sequencing missed.

Introduction

Dr. Benedict Yan is head of the Molecular Diagnosis Centre (MDC) at the National University Hospital (NUH) in Singapore. Founded in 1998, MDC is the main molecular NUH laboratory and also provides molecular analysis services for research collaborators at local universities and research institutes. The MDC team uses various molecular analysis technologies, including Sanger and capillary electrophoresis (CE) sequencing, and real-time PCR (qPCR).

Dr. Yan's research focuses on hematologic cancers, specifically acute myeloid leukemia (AML). He believes that next-generation sequencing (NGS) could streamline AML diagnosis. He and his team have been conducting Sanger and NGS correlation studies over the last four years. In the process, their NGS-powered studies have identified novel genetic markers that could be useful in the development of diagnostics and therapeutics for AML and myeloid neoplasms. They've also gathered data for future validation of NGS in their molecular diagnostics laboratory.

iCommunity spoke with Dr. Yan about his current research and the potential value of NGS analysis approaches in identifying cancer mutations.

Q: What sparked your interest in becoming a pathologist?

Benedict Yan (BY): There are many different clinical specialties where the nature of the job doesn't allow for much scientific inquiry. As a pathologist, I work in a laboratory environment within a clinical setting, and this offers the opportunity to pursue scientific questions.

Q: What types of testing do you perform at MDC?

BY: Approximately 40–50% of our testing is in infectious disease. Specifically, the detection and quantification of viruses. The rest of our test menu is in oncology and inherited diseases such as thalassemia, Fragile X, and hemochromatosis.

Q: Do you partner with other research groups at the hospital?

BY: The lab is within an academic medical center and next to a university. There's also a whole ecosystem around this hospital that includes research institutes and individual clinician scientists.

We try our best to collaborate with as many people as possible. We publish about 10 papers a year with a diverse group of researchers that cover a wide spectrum of diseases and conditions.

Q: What is the focus of your own research studies?

BY: My focus is on leukemia, specifically AML. I'm also interested in studying myeloid neoplasms.

Q: What molecular biology technologies do you use at MDC?

BY: Initially, the lab was focused on pediatric diseases, such as Fragile X. We were using archaic technologies, such as Southern blot, which involved radioactivity. qPCR was, and remains, one of the staples in our work process. As we moved into cancer, we began performing DNA and RNA testing. We use conventional CE and Sanger sequencing, as well as techniques such as endpoint and digital droplet PCR.

Q: What technologies have you been using for your AML research?

BY: For our AML research, we've performed DNA sequencing using CE and Sanger, and used qPCR to study fusion transcripts.

Q: What is known about the AML genome?

BY: AML was the first cancer genome sequenced after the completion of the Human Genome Project. Its mutational landscape has been studied extensively and is well characterized concerning the biology and the clinical implications of these mutations.



Dr. Benedict Yan is head of the MDC-NUH in Singapore.

We have seen an evolution from 2001 when the World Health Organization (WHO) issued a taxonomy of AML,¹ to 2016 and the latest WHO guidelines for AML classification.² There are a greater number of AML categories that are clinically significant than was thought originally. We're now using conventional technologies (for

example, Sanger sequencing) to test four biomarkers in AML samples. Soon, we'll be adding several more biomarkers. It's hard to keep up using Sanger sequencing.

Q: What prompted you to consider NGS in your AML studies?

BY: We realized that the conventional methods for AML biomarker discovery and testing were too expensive and laborious. We decided to explore the potential of NGS. It's an efficient, cost-effective alternative to conventional sequencing methods. We started the AML project in late 2014, using a benchtop NGS platform.

Q: How are you implementing NGS in your laboratory?

BY: We use NGS to discover new AML variants and then perform Sanger sequencing to validate them. The amount of information that comes from NGS analysis is immense. It requires a whole different body of knowledge that traditional pathologists aren't familiar with. In our current AML studies, we are slowly gaining experience with NGS and constructing our knowledge.

"We realized that the conventional methods for AML biomarker discovery and testing were too expensive and laborious. We decided to explore the potential of NGS."

Q: What types of AML and multiple myeloma studies are you performing with NGS?

BY: For the last 10 years, we have banked peripheral blood and bone marrow of AML subjects. We've just completed an NGS study of a biomarker panel using these samples, and are validating our findings analytically and clinically. For analytical validation, we're comparing NGS findings with data from Sanger sequencing, which is the current gold standard.

The body of knowledge for AML biomarkers has been developed with Western populations. However, our Singapore population is diverse, consisting of Chinese, Indians, our own indigenous populations, and expatriates living here. We are trying to build a database of biomarkers that are relevant for our diverse population.

We're also using NGS to identify the presence of novel gene fusions in multiple myeloma.

Q: What are the results of these studies so far?

BY: We've found that NGS is very accurate. For example, we used the NGS system to analyze samples for *CEBPA* mutations, an AML biomarker. *CEBPA* is high in GC content. Using NGS, we identified the same mutations as Sanger sequencing, as well as a few more.³

In our multiple myeloma study, we've identified several novel fusions. To confirm whether these fusions were genuine, we designed primers against them and performed reverse transcriptase PCR (RT-PCR). We were delighted that our RT-PCR

and NGS data matched and that the fusions were genuine. We're conducting studies to see if these novel fusions have clinical impact.^{4,5}

"We've found that NGS is very accurate. For example, we used the NGS system to analyze samples for *CEBPA* mutations, an AML biomarker. *CEBPA* is high in GC content. Using NGS, we identified the same mutations as Sanger sequencing, as well as a few more."

Q: Is there potential for other NGS-based approaches, such as single-cell RNA sequencing, to be used in cancer research?

BY: Single-cell genomics holds much promise in cancer. Currently, we use flow cytometry for single-cell analysis of hematological malignancies to determine whether there's residual disease. Approximately 20,000 cells can be analyzed by flow cytometry to assess about 30 markers, depending on the antibody panel used.

Single-cell RNA sequencing could allow us to look at thousands of markers potentially and I'm excited by the promise that this technology holds. The biggest challenge would be optimizing the protocol. Working with RNA is not straightforward currently. We would need to develop a process to obtain high-quality RNA, and then identify a platform and develop a protocol to ensure reproducibility.

Q: What is the future of NGS technologies in your research?

BY: Our end goal is to reduce cost while delivering high-quality results with NGS. NGS holds the promise to deliver insights that conventional technologies could never offer. We're hoping to perform NGS studies for other cancers, as well as infectious diseases such as HIV, inherited diseases, and thalassemia. I'd also like to investigate the use of NGS to enable single-cell genomics for the clinic.

Read more about this study:

Sanger/NGS correlation studies,
www.illumina.com/science/customer-stories/icomunity-customer-interviews-case-studies/yan-nuh-interview-targseq-casestudy.html

Read more about the Illumina systems and products used in this study:

MiSeq™ System, www.illumina.com/systems/sequencing-platforms/miseq.html

Nextera™ XT DNA Library Prep Kit, www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-xt-dna.html

TruSight™ Myeloid Sequencing Panel, www.illumina.com/products/by-type/clinical-research-products/trusight-myeloid.html

TruSight RNA Pan Cancer Panel, www.illumina.com/products/by-type/clinical-research-products/trusight-rna-pan-cancer.html

NextSeq™ 550 System, www.illumina.com/systems/array-scanners/nextseq-550.html

References

1. Jaffe ES, Harris NL, Stein JW, et al. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. *Ann Oncol.* 2002;3:490–491.
2. Leonard JP, Martin P, Roboz GJ. Practical Implications of the 2016 Revision of the World Health Organization Classification of Lymphoid and Myeloid Neoplasms and Acute Leukemia. *J Clin Oncol.* 2017;23:2708–2715.
3. Yan B, Hu Y, Ng C, et al. Coverage analysis in a targeted amplicon-based next-generation sequencing panel for myeloid neoplasms. *J Clin Pathol.* 2016;69:801–804.
4. Tan M, Ng IKS, Chen Z, et al. Clinical implications of *DNMT3A* mutations in a Southeast Asian cohort of acute myeloid leukaemia patients. *J Clin Pathol.* 2017; 70:669–676.
5. Ng IK, Ng C, Low JJ, et al. Identifying large indels in targeted next generation sequencing assays for myeloid neoplasms: a cautionary tale of the *ZRSR1* pseudogene. *J Clin Pathol.* 2017; pii:clinpath-2017-204440. doi: 10.1136/clinpath-2017-204440. Epub ahead of print.