

Cost-Effective Sequencing Enables Researchers to Scale Cancer Study

Garvan researchers use the HiSeq® 2500 System to identify cancer-associated variants in the germline.

Introduction

Professor David Thomas, FRACP, PhD, and his colleagues at the Garvan Institute's Kinghorn Cancer Centre felt that the germline was being overlooked in cancer studies. They wondered whether germline DNA might contribute to an improved understanding of the genetic determinants of cancer, particularly in early onset cancers such as sarcoma.

"There are 11,000 genomes or exomes in the International Cancer Genome Consortium and the Cancer Genome Atlas that have been sequenced and put into accessible databases," Professor Thomas stated. "However, a tiny fraction of those have attached germline genomes. The germline has become the forgotten genome in cancer research."

Cost-Effective Sequencing

They embarked on a ground-breaking sarcoma study. Instead of looking directly at tumors, they analyzed the germline DNA of an 1162 proband cohort of subjects with sarcoma, using a targeted exon panel of 72 genes associated with cancer risk.

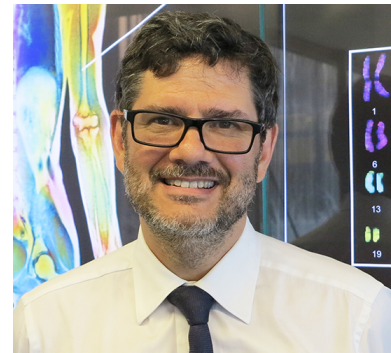
"The study was the first of its kind in a rare disease, and getting as much as we could from every dollar we spent was an important part of our strategy," Professor Thomas said. They chose the HiSeq® 2500 System to perform targeted exon sequencing.

"The HiSeq 2500 System was perfectly suited for the study and made sequencing incredibly affordable," Professor Thomas stated. "The cost of sequencing the panel was about \$200 AUD per sequence, which meant that we had the capacity to do this on a scale of more than 1000 cases."

Speed and Efficiency

Ten batches of sequencing were performed over a three-year period. "For the later batches, the sequencing was completed in three days using the high-output mode of the HiSeq 2500 System." Professor Thomas added. "We were 96-plexing in a single lane and ending up with ample read depth. All the sequencing was completed about three years ago. We've been analyzing it since that time."

The team performed rare variant burden analysis to identify which genes among the 72 on the panel were enriched. In addition to *TP53*, they identified strongly enriched genes that they didn't expect to be associated with sarcoma, including *ATM*, *ATC*, and *ERCC2*.



Professor David Thomas, FRACP, PhD is Director of the Kinghorn Cancer Center and Head of the Cancer Division of the Garvan Institute in New South Wales.

"The HiSeq 2500 System was perfectly suited for the study and made sequencing incredibly affordable."

"Among people with sarcoma, there were 61 individuals who had pathogenic variation in genes like *APC*, *MLH1*, *MSH2*, and *MSH6*, which are associated with bowel cancer," Professor Thomas said. "We also identified 28 individuals in the sarcoma cohort who have *BRCA1* or *BRCA2* mutations, which is well known in breast cancer medicine. The importance of these genes is that we already have established programs for managing risk. I think that's a very important outcome."

Whole-Genome Sequencing Follow-Up

The team will soon perform whole-genome sequencing (WGS) of the same cohort to identify additional genes involved in sarcoma risk. "That is incredibly exciting because we can start to ask broader and deeper questions than we could when we were looking at a panel of known genes," Professor Thomas stated. "With WGS, we can look for pathogenic variation across the totality of that pathway and ask whether we see a signal from every gene, or just the ones we picked for the panel. Looking at the whole genome also allows us to discover new genes that have not previously been associated with cancer."

WGS might enable them to fill in the heritability data that is missing from what they detected with targeted exon sequencing. "It will be interesting to see how many more biomarkers we can discover that are associated with the early onset of sarcoma," Professor Thomas added. "The whole human genome sequencing will be performed in our HiSeq X Ten facility at the Garvan Institute, which we're very excited about."

Read more about this study:

Exploring the Forgotten Genome of Cancer Research, www.illumina.com/science/customer-stories/icommunity-customer-interviews-case-studies/exploring-the-forgotten-genome-of-cancer-research.html

Learn more about the Illumina system used in this study:

HiSeq 2500 System, www.illumina.com/systems/sequencing-platforms/hiseq-2500.html

Summary

Overview

Garvan Institute researchers are conducting sarcoma studies to identify rare cancer-associated variants in the germline.

Challenge

Performing cost-effective sequencing to stretch their study budget.

Solution

Dr. David Thomas and his team chose the HiSeq 2500 System to perform sequencing using a targeted exon panel of 72 genes associated with cancer risk.

Benefits

The low cost of sequencing on the HiSeq 2500 System enabled the team to scale the study to >1000 cases. They identified pathogenic mono- and polygenic variation in known and novel cancer genes. A follow-on WGS study in the same cohort will be conducted using the HiSeq X Ten System.

"For the later batches, the sequencing was completed in three days using the high-output mode. We were 96-plexing in a single lane and ended up with ample read depth."