GenapSys™ Sequencing Platform: Variant Calling

- **Highly accurate sequencing data yields high confidence variant calling**
- **Excellent concordance for variant calling relative to established sequencing technologies**
- **Low price per run and low price per sample**
- **More than 1.2 Gb of highly accurate DNA sequence data per run**

Introduction

Genetic medicine relies upon the ability to confidently identify differences between a patient's genome and the reference. Next-Generation Sequencing (NGS) has become the tool of choice for discovering and identifying such variants. Variant calling can be used for carrier screening, detecting rare diseases, and provides the foundation for personalized medicine. GenapSys™ has developed a novel, scalable, low cost NGS platform that is capable of generating accurate sequence information needed for variant calling applications. Using targeted sequencing of a human genomic DNA sample, we demonstrate the capabilities of the GenapSys Sequencing Platform for variant calling and show high levels of concordance with another sequencing technology.

Technology

The GenapSys NGS technology is based on detection of electrical impedance changes resulting from single base incorporations during sequencing-by-synthesis. The core of the technology is a CMOS-based electronic chip that enables scalability and low instrument and consumable costs. Chips with low, medium, and high sensor throughputs can be run on the same GenapSys instrument, giving the user flexibility in NGS assay design and sample multiplexing. A typical single run with a medium throughput 16M sensor chip generates more than 1.2 Gb of data with an average read length of 150 bp. Sequencing accuracy is key to identifying variants and the GenapSys platform is capable of generating sequencing reads with accuracy >99%.

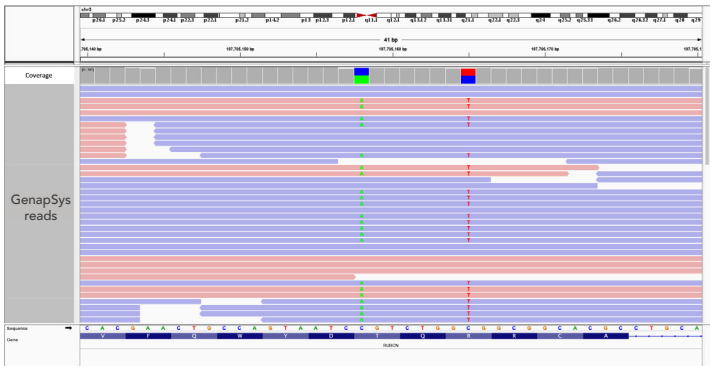

Fig. 1 IGV visualization of two heterozygous SNPs in the RUBCN gene.

Experimental Methods

GenapSys collaborated with an independent third-party that supplied human genomic DNA from patient samples. Genomic libraries were generated by random shearing, end repair, A-tailing, and ligation of adapters. Libraries were size selected to achieve a median insert size of ~200 bp. Libraries from the patient samples were enriched via probe-based capture using the IDT xGen® Inherited Diseases Panel, which spans 11.1 MB of the human genome (116,355 probes) and targets a range of variants that have been associated with inherited disease disorders. Individual library molecules were clonally amplified onto beads and beads were loaded onto the chip for sequencing on the GenapSys Sequencer. The sample enriched libraries were also subjected to Illumina® sequencing.

Streamlined targeted sequencing workflow on the GenapSys Sequencing Platform

Library prep

Generate sequencing libraries. Size select. Enrich using capture panels.

Clonal Amplification

Clonally amplify libraries on the GenapSys Sequencing Prep System.

Sequencing

Load amplified beads into the chip and sequence using the GenapSys Sequencer.

Data Analysis

Generate FASTQ with GenapSys analysis software. Perform alignment and variant calling.

Results

Sequencing data was processed using the GenapSys base-calling pipeline. FASTQ sequences were aligned to the hg38 reference genome using BWA-MEM. Germline variant calling of the patient samples was performed using Google DeepVariant™ that had been further trained on GenapSys sequencing data. Similar bioinformatics processing was performed on libraries sequenced with Illumina. An example of an identified heterozygous variant visualized in the IGV genome browser is shown in Figure 1.

Variant calling was carried out independently on sequencing data generated on the GenapSys and Illumina sequencing platforms. As a benchmark, we compared the variant calling results from the two platforms. A summary of the results is shown in the Venn diagram in Figure 2. Results demonstrate high concordance for variant calling between data generated using the GenapSys and Illumina platforms (F1-score of 0.972, sensitivity of 0.949, and precision of 0.995).

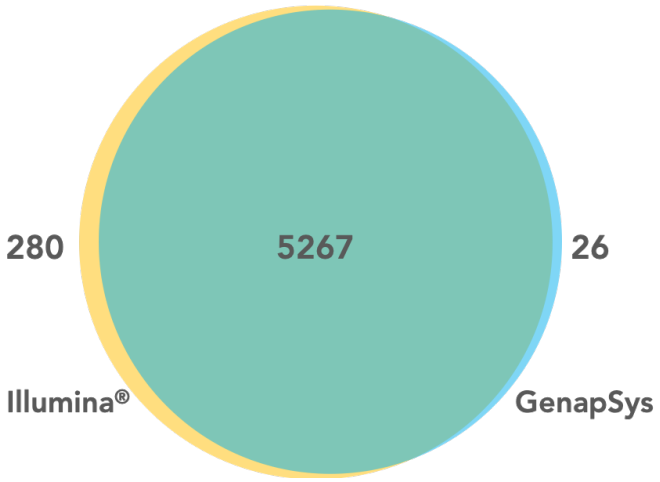

Fig. 2 Concordance of germline single nucleotide variants called from GenapSys or Illumina sequencing data. Sequenced libraries were generated from patient samples enriched using an Inherited Disease hybrid capture panel.

Conclusion

The GenapSys sequencing platform is capable of generating the highly accurate DNA sequence data that is critical for identification of variants. Variant calls are highly concordant with other sequencing technologies. The GenapSys data was all generated on a small, low-cost instrument with the footprint of an iPad. The GenapSys system will put the power of variant calling into the hands of the individual researcher.

To learn more about the GenapSys Sequencing Platform, visit GenapSys.com

GENAPSYS, INC.
200 Cardinal Way, 3rd Floor, Redwood City, CA 94063 • (800) 796-7710 • GenapSys.com

GenapSys
Sequencing without compromise

100-2019-002-A