GenapSys™ Sequencing Platform: Oncology Research Applications

- **Germline and somatic mutation calling with allele detection as low as 1%**
- **Compatibility with existing hybrid capture and amplicon based library prep methods**
- **Demonstrated with cell line DNA, tumor FFPE, fresh frozen, and blood samples**
- **High accuracy NGS platform with >1.2 Gb of data and low price per run**

Introduction

Next-Generation Sequencing (NGS) technologies have made rapid strides in the throughput and accuracy of DNA sequencing in recent years. These advances have revolutionized biomedical and clinical research, especially in oncology. NGS cancer assays are used to determine cancer predisposition, detect cancer early, identify key tumor mutations, monitor treatment response and make progress towards personalized therapies. GenapSys™ has developed a novel, scalable, low cost, and high accuracy NGS platform that is an ideal solution for a diverse range of oncology research applications. Here, we demonstrate the performance of the GenapSys Sequencing Platform for identifying germline and somatic mutations across different sample types relevant for oncology research. Specifically, we demonstrate the performance of hybrid-capture libraries (exome and pan-cancer panels) on the following sample types: reference DNA standards, and DNA extracted from patient tumor FFPE, fresh frozen tumor tissue and blood samples.

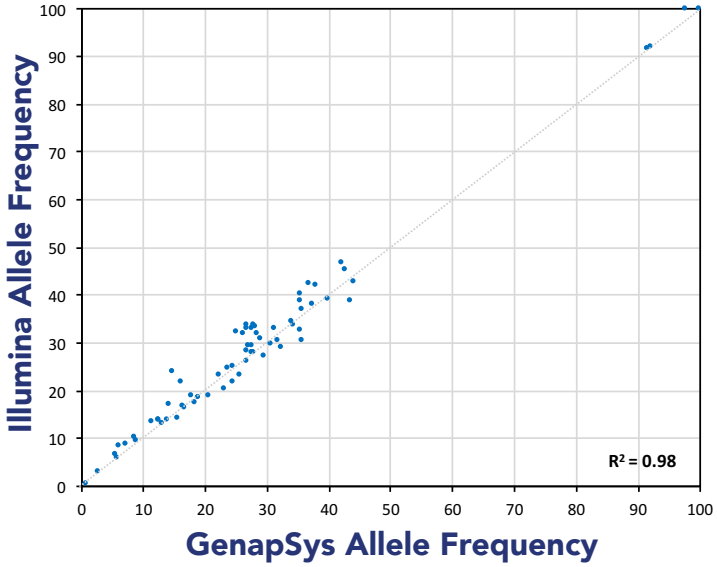

Fig. 1 Correlation of allele frequencies measured from sequencing data generated by the GenapSys and Illumina sequencing platforms. Sequencing libraries were generated from the OncoSpan Reference DNA standard enriched using a pan cancer hybrid capture panel.

Technology

The GenapSys sequencing technology is based on the accurate detection of electrical impedance changes resulting from nucleotide base incorporation 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 (1M), medium (16M), and high (144M) sensor throughputs can be run on the same GenapSys Sequencer, 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 greater than 99% raw accuracy and an average read length of 150 bp.

Streamlined workflow for oncology applications on the GenapSys Sequencing Platform

Library prep

Generate sequencing library using hybrid capture panels or multiplex PCR amplicons.

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.

Methods

GenapSys collaborated with independent third parties who provided DNA from patient tumor FFPE, fresh frozen tumor, and blood samples. The OncoSpan Reference Standard (HD827) was ordered from Horizon Discovery, and is composed of DNA from multiple cell lines. Genomic libraries were generated following mechanical fragmentation, adapter ligation, size selection with a median insert size of ~200 bp and PCR. Hybrid capture-based enrichment was done using the IDT Exome Research Panel (39 Mb region, 19,396 genes) for the whole exome sequencing (WES) library, or using the IDT Pan Cancer Panel v1.5 (800 kb region, 127 genes) for the cancer panel library. Post-capture PCR was performed to generate the final libraries. Individual library molecules were clonally amplified onto beads, and enriched amplified beads were loaded onto the chip for sequencing on the GenapSys Sequencer. Illumina® cancer panel libraries were generated and sequenced on an Illumina NextSeq™ instrument, to compare sequencing performance. Illumina WES data was generated by the independent collaborator, using an Agilent® SureSelect Human All Exon library.

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™, which has been further trained on GenapSys sequencing data, while Illumina data was analyzed using the standard DeepVariant model. Variant calling of the OncoSpan reference standard and patient samples was done using Vardict for both GenapSys and Illumina data.

Results

The OncoSpan Reference Standard (HD827) is a cell line-derived sample containing 386 variants across 152 key cancer genes, with variant allele frequencies (AF) characterized using ddPCR or WES. We focused on the variants covered by the Pan Cancer Panel, and demonstrated high concordance of the allele frequencies obtained following variant calling of sequencing data from the GenapSys platform and Illumina. (R^2 = 0.98) (Fig. 1). In particular, we detected low frequency variants with allele frequencies as low as 1% using both hybrid-capture Pan Cancer Panels (Fig. 1) and multiplex PCR-based amplicon panels, such as the AmpliSeq™ Cancer Hotspot Panel v2 (data not shown).

The GenapSys sequencing performance was also tested on a range of clinically relevant sample types, including DNA extracted from tumor FFPE, fresh frozen tumor and blood, and across different hybrid capture libraries (pan cancer and whole exome). Whole exome sequencing was performed on libraries generated from DNA extracted from an FFPE sample, and the SNV calls show high concordance between GenapSys results and Illumina results. In particular, the SNV calls show a precision score of 0.997, sensitivity of 0.917, and a F1-score of 0.955 (Fig. 2). Some of the non-overlapping variant calls may be attributable to the use of different hybrid capture panels (IDT xGen® Exome and Agilent SureSelect Human All Exon). We sequenced Pan Cancer Panel libraries generated from matched fresh frozen tumor tissue and blood samples, and our results show strong overlap between the SNV calls on the GenapSys and Illumina platforms, for both fresh frozen and blood samples (Fig. 3 A, B). Additionally, comparison of SNV calls between fresh frozen and blood samples on each platform, allows us to identify potential germline and somatic SNVs (Fig. 3 C,D). The GenapSys and Illumina results show 147 and 152 potential germline SNVs respectively, with 144 shared SNVs. The GenapSys analysis shows 3 potential somatic SNVs above 2%, of which Illumina captured 1 SNV at >2% and the other two SNVs were detected but not reported by the Illumina analysis, since they were below the 2% threshold.

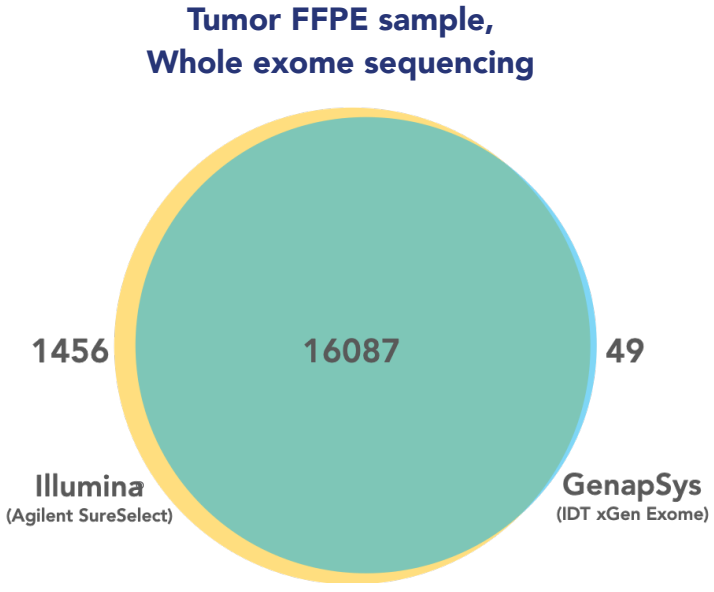

Fig. 2 Comparison of single nucleotide variant (SNV) calls with the GenapSys and Illumina platforms, based on whole exome sequencing of a tumor FFPE sample. The GenapSys library was generated using the IDT Exome Research Panel and the Illumina library was generated using an Agilent SureSelect Human All Exon panel. SNV calls in the regions of overlap are reported here, and show high concordance with a precision score of 0.997, sensitivity of 0.917, and a F1-score of 0.955.

Fig. 3: A **Tumor: Fresh frozen**

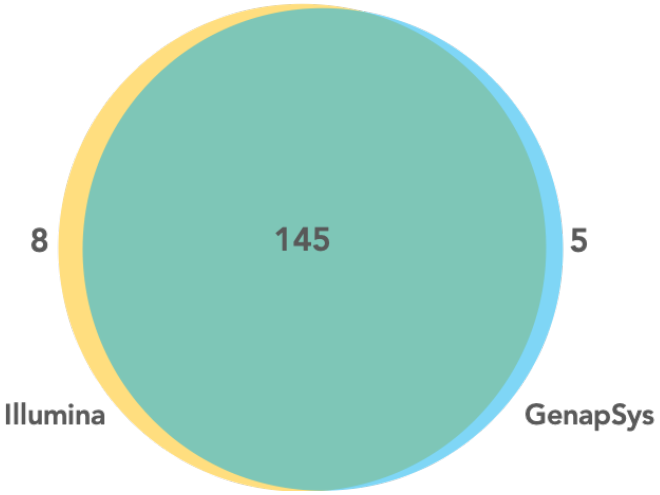

Fig. 3: B **Blood**

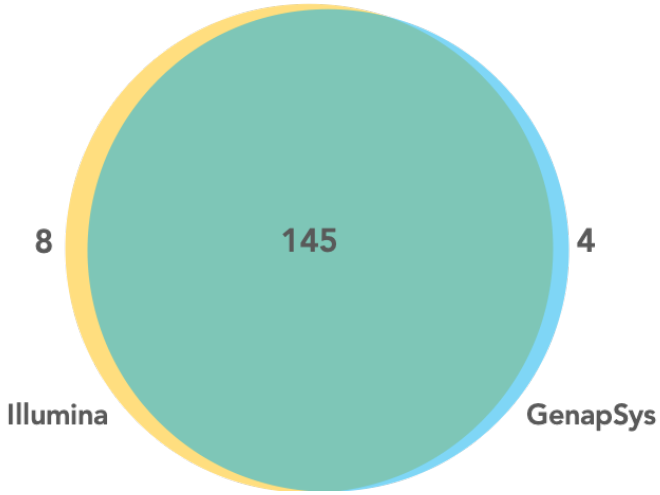

Fig. 3: C **Illumina**

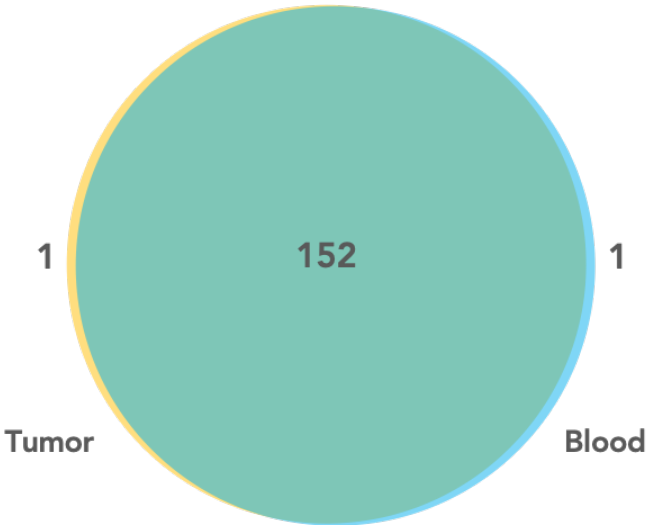

Fig. 3: D **GenapSys**

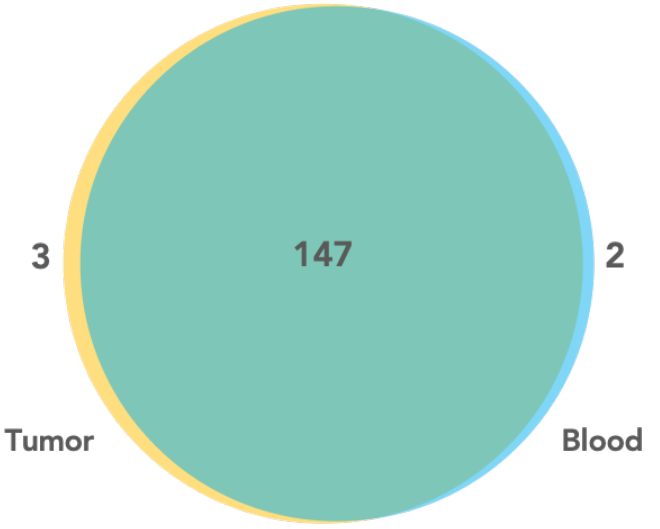

Fig. 3 Comparison of single nucleotide variant (SNV) calls with the GenapSys and Illumina platforms, based on Pan Cancer Panel sequencing of matched fresh frozen tumor and blood samples. (A) and (B) show high concordance of SNV calls based on GenapSys and Illumina NextSeq sequencing of the tumor and blood samples, respectively. Comparison of the fresh frozen tumor and blood sample SNV calls (AF > 2%) shows potential germline and somatic mutations for Illumina and GenapSys in (C) and (D). GenapSys data shows 147 potential germline variants, of which 144 are shared with Illumina calls. GenapSys data also shows 3 potential somatic SNVs, of which 1 is shared with Illumina and the other two SNVs are present in Illumina data, but fall below the 2% threshold.

Conclusion

The GenapSys sequencing technology offers high accuracy sequencing, thus making it ideal for applications ranging from germline variant calling to somatic mutation detection in cancer. We have demonstrated high concordance of variant allele frequencies and the detection of low frequency mutations down to 1%, using a well characterized reference DNA standard. We also show strong correlation in the SNV calls between GenapSys and an established sequencing technology, for whole exome and pan cancer libraries generated from DNA extracted from patient FFPE, fresh frozen tumor, and blood samples. Thus, the GenapSys sequencing platform offers a high accuracy, low-cost, and scalable solution for diverse oncology research applications.

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

100-2019-003-A