High Accuracy NGS Platform based on Electrical Impedance Detection: Applications for Oncology Research

Saurabh Paliwal, Meysam Rezaei Barmi, Ali Nabi, Xavier Gomes, Mohammad Fallahi, Maryam Jouzi, Seth Stern, Paul Kenney, Kosar B. Parizi, Tyson A. Clark, Hamid Rategh, Subra Sankar, and Hesaam Esfandyarpour

GenapSys, Inc. 200 Cardinal Way, Redwood City, CA 94063

Abstract

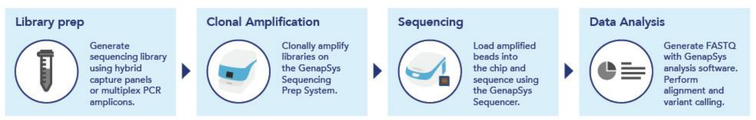
Next Generation Sequencing (NGS) technologies have made rapid strides in the throughput and accuracy of DNA sequencing in recent years. These advances have revolutionized biomedical and clinical research, especially in oncology. NGS cancer panels are used to determine cancer predisposition, detect early cancer, identify tumor mutations, and develop personalized therapies. Here, GenapSys™ presents a novel, scalable, low cost, and high accuracy sequencing platform, and demonstrates its applications to oncology research.

The GenapSys Sequencing Platform is based on accurate detection of electrical impedance changes resulting from single base incorporations during sequencing-by-synthesis. We show that impedance changes measure a steady state dNTP incorporation signal, leading to higher accuracy. The core of the technology is a CMOS-based electronic chip that enables scalability and low instrument and consumable costs. Chips with 1M, 16M and 144M sensors can be run on the same GenapSys Sequencer, giving a lab flexibility in NGS assay design and sample multiplexing. We demonstrate that a single run with a 16M sensor chip generates 1.2 - 2 Gb of data, with greater than 99% raw accuracy and average read length of 150 bp.

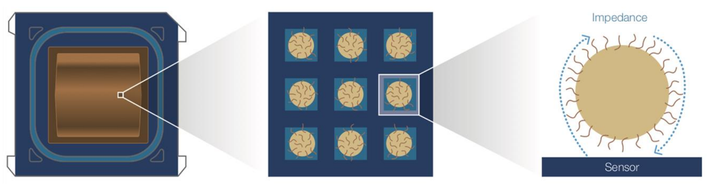
We tested hybrid-capture and amplicon-based cancer panels on a range of DNA sources, including oncology reference standards derived from cell line DNA, as well as clinical tumor FFPE, fresh frozen tissue and blood sample DNA. Reference DNA standards from Horizon Discovery included the Quantitative Multiplex (HD701) and Oncospan (HD827). For hybrid-capture libraries, we tested the IDT xGen Pan Cancer Panel v1.5 (800 Kb target region, 127 genes) and the IDT xGen Exome Research panel (39 Mb target region, 19,396 genes). We detected low frequency mutations in the range of 1%-24.5% across multiple standards with the cancer panel, with mean coverage of >600x in a single run. Whole exome sequencing of clinical FFPE and blood samples showed high concordance (F1 score > 95%) of SNV mutation calling with Illumina technology. For amplicon panels, we used the Ion AmpliSeq Cancer Hotspot Panel v2 (207 amplicon pairs, 50 genes) with Horizon reference standards. We demonstrate detection of low frequency mutations (>1%) and a high correlation (R squared > 0.99) with expected allele frequencies.

GenapSys Sequencing Workflow

A. Sequencing Workflow Overview



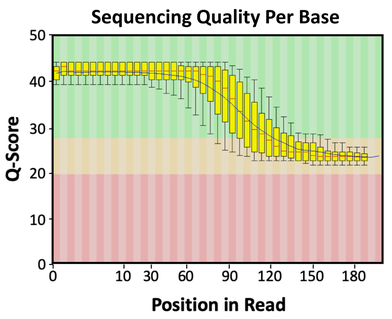
B. GenapSys Sequencing Chip



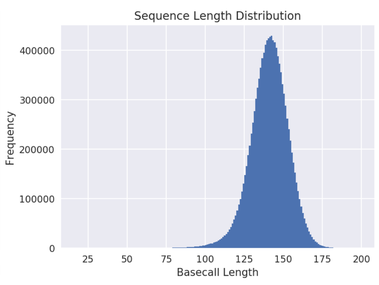
A. Schematic of the GenapSys Sequencing workflow. Hybrid capture panel or multiplex PCR amplicon libraries are generated using standard industry techniques. They are clonally amplified on the automated GenapSys Sequencing Prep System. Amplified beads are sequenced on the GenapSys Sequencer, and the data analysis pipeline generates a FASTQ file and variant calls
B. The GenapSys Sequencing Chip is an array of electronic CMOS sensors. Each sensor is loaded with a clonally amplified bead, and detects impedance changes resulting from nucleotide incorporation during sequencing-by-synthesis

Sequencing Performance

C. Accuracy



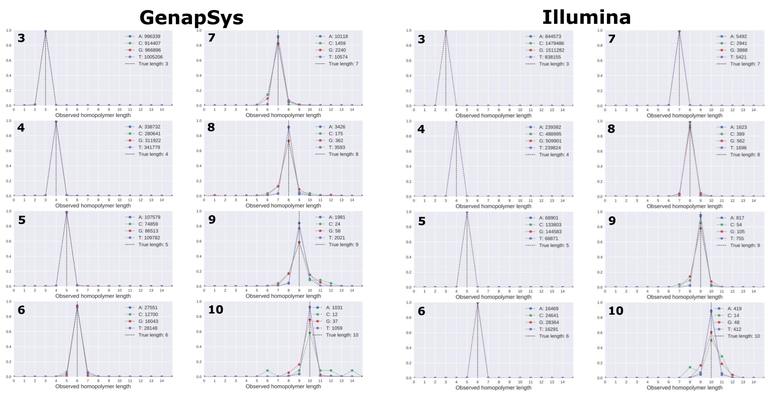
D. Read Length



***E. coli* Raw Accuracy* = 99.85%**

Substitution Errors	Deletion Errors	Insertion Errors
0.016%	0.089%	0.044%

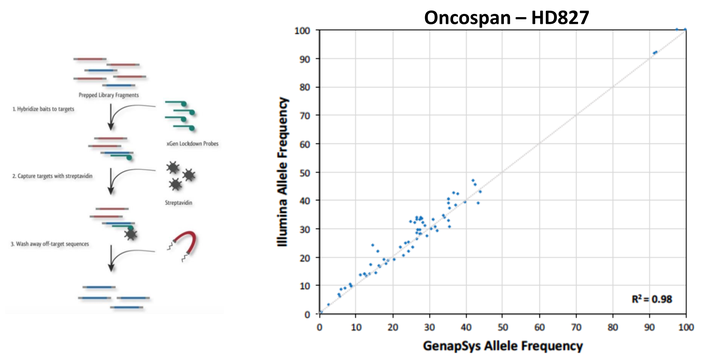
E. Homopolymer Performance



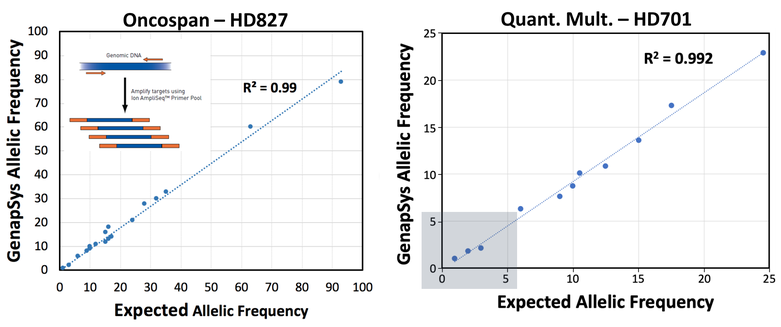
C, D Sequencing metrics for a typical sequencing run. Plots for an E. coli run show Sequencing quality per base vs. the Position in Read (C), the histogram of read lengths in (D) and breakdown of error rates (Substitution, Deletion, and Insertion Errors), giving rise to ~99.85% raw accuracy.
E. Whole exome sequencing of a NA12878 library was performed on the GenapSys and Illumina NextSeq sequencers. Base calls from homopolymer regions (3-10 nt) were compared to the hg38 reference for each individual nucleotide, to characterize the Homopolymer calling performance

Library Preparation Approaches: Hybrid Capture & Amplicon PCR

F. Hybrid Capture Pan Cancer Panel



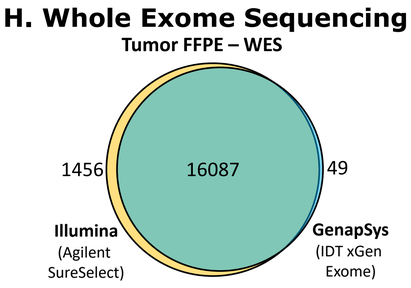
G. Multiplex Amplicon PCR Cancer Panel



F. High concordance of allele frequencies based on GenapSys and Illumina sequencing of a hybrid capture library (R^2 = 0.98). The Horizon HD827 Oncospan reference DNA standard was used to generate a library using the IDT xGen Pan Cancer panel ver 1.5
G. Low frequency variant detection (down to 1%) using a multiplex Amplicon PCR based library. The x-axis shows expected allelic frequencies of 18 variants in HD827 (left) and 11 variants in the HD701 Quantitative Multiplex Reference Standard (right), quantified using ddPCR. The y-axis shows allelic frequencies based on GenapSys sequencing of a library generated using the AmpliSeq Cancer Hotspot panel v2 (R^2 ~ 0.99 for both samples).

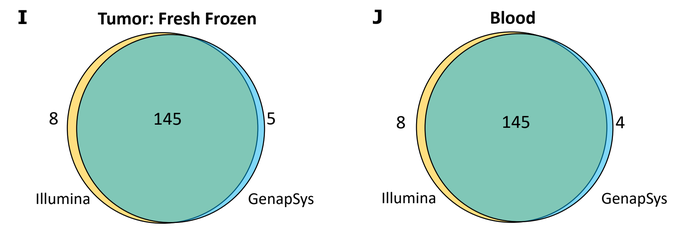
Performance on clinical samples: Tumor FFPE, Fresh Frozen, Blood

H. Whole Exome Sequencing of Tumor FFPE samples

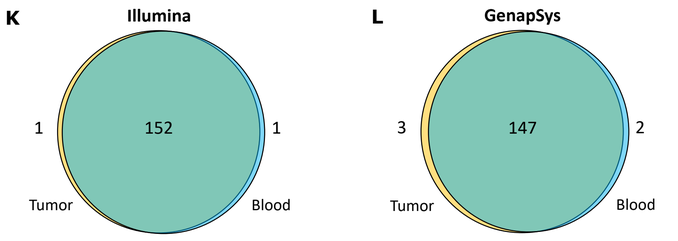


Statistic	FFPE
Average Depth (GenapSys)	201x
Fraction Covered at 10x (Gen)	98%
F1-score	95.5%
Precision	99.7%

I, J. Concordance of SNV calls: Pan Cancer Panel



K,L. Germline and Somatic Variant Calling



G. Comparison of SNV calls with the GenapSys and Illumina platforms, based on WES of a tumor FFPE sample. Genapsys and Illumina libraries were generated using the IDT Exome Research and the Agilent SureSelect exome panel, resp. **I, J.** Comparison of SNV calls with the GenapSys and Illumina platforms, following sequencing of IDT xGen Pan Cancer panel v1.5 libraries of matched fresh frozen tumor and blood samples. **K, L.** Comparison of the fresh frozen tumor and blood sample SNV calls (AF > 2%) shows potential germline and somatic variants. GenapSys data shows 147 potential germline variants (144 shared with Illumina), and 3 potential somatic SNVs (all 3 present in Illumina data, with only 1 SNV > 2% AF)

Summary

We demonstrate that the GenapSys Sequencing Platform is an accurate, scalable, and low cost solution for oncology research, using a wide range of sample types and NGS assays. Key highlights:

- High accuracy, scalable NGS platform with 1.2 - 2 Gb of data and low price per run
- Compatibility with Hybrid Capture and Multiplex PCR Amplicon-based cancer panel libraries
- Performs well with diverse sample types: Cell line DNA, Tumor FFPE, Fresh frozen, and Blood samples
- Accurate germline and somatic mutation calling, and low frequency allele detection, as low as 1%

GenapSys Sequencer System Specifications:

- 85% of bases >Q30
- Raw accuracy >99%
- 1.2 – 2Gb output
- 10 – 13M reads
- Average read length of 150bp
- Run Time ~ 24h
- Scalability: Low, medium & high throughput chips
- Low Cost per run

www.genapsys.com