GenepoweRx™

Genetic Science powered by K&H

OncoRx _Genomics Discovery:

Targeted Sequencing Solutions

White Paper | Version 1.0 | 16-Jan-2023

Prepared by: Dr.Kalyani Palasamudram

Reviewed by:

Dr. Kalyan Ram Uppaluri

Dr.Hima J. Challa

Uppaluri K&H Personalized Medicine Clinic Pvt. Ltd. Hyderabad

GenepoweRx™
Genetic Science powered by K&H

Title: OncoRx _Genomics Discovery: Targeted Sequencing Solutions

Objective: Demonstrate how and why OncoRx product/service is superior to other providers

Introduction Next-generation sequencing technologies have enabled application of targeted therapies in cancer by providing a high-throughput and cost-effective strategies to sequence thousands of cancer genomes.

There are seven major practicalities to consider before a cancer sample is sent for next generation sequencing analysis.

1. NGS Platform employed
2. NGS Genomic Target Amplification Chemistry
3. Depth of coverage
4. NGS Panel content
5. Sequencing types and error rates
6. Data Analysis and Report generation
7. Sequencing cost

1. NGS Platform_ Illumina's Novoseq 6000 NGS Analyzer

The concept behind NGS technology is the same in all analyzers. Instead of sequencing a single DNA fragment, NGS extends this process across millions of fragments in a massively parallel fashion. However, the chemistry behind this massively parallel sequencing is what matters. More than90% of the world's sequencing data are generated by Illumina sequencing by synthesis(SBS) chemistry. Illumina's **Novoseq 6000 NGS Analyzer thus delivers high accuracy, a high yield of error-free reads, and a high percentage of base calls above Q30** (PMID: 23718773, PMID: 18987734, PMID: 24167589).

2. NGS_ Genomic Target Amplification Chemistry

Amplicon-based chemistry vs. Hybridization-based capture technology.

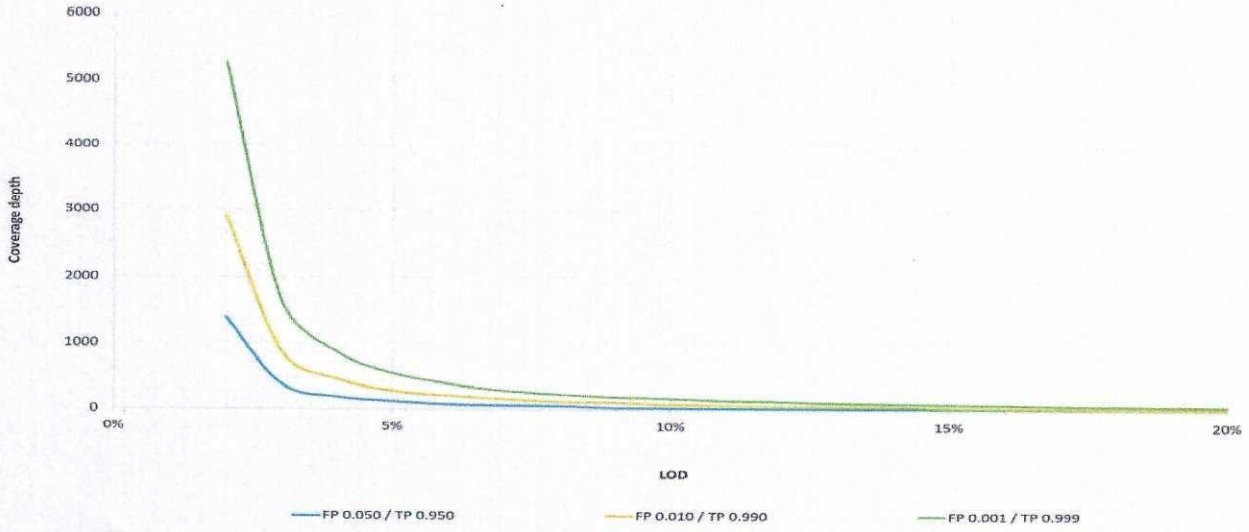
Amplicon-based assays provide high on-target specificity, but compromise on coverage of genomic regions (PMID: 20847746). Hybridization capture, a Probe based chemistry is **more precise, robust and has demonstrated to be a promising technology with better uniformity** (PMID: 26110913). Hybridization-

based capture strategies **can amplify larger regions of interest and provide accurate info on additional types of alteration, such as Copy Number Variations** (PMID: 24142049, PMID: 24189654, PMID: 24265154).

3. Depth of coverage

Coverage refers to the average number of sequenced DNA bases that align to each base of the reference DNA. This coverage varies from one sequencing provider to the other. **NGS sequencing depth directly affects the reproducibility of variant detection** and varies from 1000x to 5000x with decrease in error rates. A depth of coverage of 1,650 or higher is generally suggested. However, **higher the sequencing coverage, greater would be the depth of sequencing and higher the number of aligned sequence reads and thus more accurate detection of base calls;** regardless of whether the base call is the same as the reference base or is mutated (PMID: 31552176).

Higher sequencing coverage also aids to address tumor heterogeneity issues and to precisely characterize subclonal variants and limited cancer cell clusters in the FFPE samples. The sequencing depth threshold is determined based on the intended Level of Detection (LOD), the tolerance for false positive(FP) or false negative results (FN), and the error rate of sequencing (PMID: 28341590, PMID: 29576615). 2500x to 3000x is considered for sequencing at GenepoweRx to ensure accurate identification of rare variants associated with cancer.



Courtesy: Petrackova A et.al. 2019.Frontiers in Oncology.

4. NGS Panel content

WGS Vs. Targeted Sequencing Panels

With targeted sequencing, a subset of genes or regions of the genome are isolated and sequenced. Targeted sequencing allows researchers to focus time, expenses, and data analysis on specific areas of interest and enables sequencing at much higher coverage levels. For example, a typical WGS study

achieves coverage levels of 30–50× per genome, while a targeted resequencing project can easily cover the target region at 500–1000× or higher.

Agilent SureSelect hybrid capture technology panel is employed at GenepoweRx for targeted sequencing of cancer samples. Hybrid capture target enrichment employs hybridization probes to capture and isolate target sequences from an NGS library. Hybrid capture is sensitive and detects single nucleotide variants, translocations, structural variants, insertions and deletions, and copy number variations. Hybrid capture is the appropriate choice for FFPE tissue samples including needle biopsies, or when sample is otherwise scarce.

Panel content is based on the Glasgow Precision Oncology Laboratory (GPOL) broad cancer genomic community testing. GPOL is a team of scientists with internationally recognized expertise in the technology, biology, and clinical utility of cancer genomics (PMID: 32025007, PMID: 25528188). GPOL has detailed curation of genomic data to define the landscape of clinically and biologically significant genomic events in cancer. This includes not only a published literature review, but also the International Cancer Genome Consortium (ICGC) and the Pan-Cancer Analysis of Whole Genomes (PCAWG) study (PMID: 32025007, PMID: 23539594, PMID: 24390350).

The cancer panel at GenepoweRx spans 174 high confidence onco genes related to response, resistance to therapies & 179 genes with proven and emerging clinical utilities (for drugs in clinical trials). The panel size is 3.96Mb and spans all FDA approved targeted therapy biomarkers for sequencing (Level 1, 2, 3) and resistance markers as well. It also covers all the 14 core genes of Homologous Recombination Repair pathway and all Microsatellite Instability markers for analysis.

5. Sequencing types and error rate

Single end sequencing Vs. Cutting-edge Paired end sequencing and duplex MBC sequencing.

In single-end reading, the sequencer reads a fragment from only one end to the other, generating the sequence of base pairs. Single-read sequencing can be a good choice for certain methods such as small RNA-Seq or chromatin immunoprecipitation sequencing (ChIP-Seq). However, for clinical utility, paired end sequencing remains the best choice.

GenepoweRx adopts Paired-end sequencing which involves **sequencing both ends of the DNA** fragments in a library and aligning the forward and reverse reads **as read pairs**. In addition to producing twice the number of reads for the same time and effort in library preparation, sequences aligned as read pairs allows highly precise, high-quality alignment across DNA regions containing repetitive sequences, and produce long contigs for *de novo* sequencing by filling gaps in the consensus sequence. This enables the ability to detect insertion-deletion (indel) variants, which is not possible with single-read data (PMID: 24167589).

Analysis of differential read-pair spacing also **allows removal of PCR duplicates**, a common artifact resulting from PCR amplification during library preparation. Paired-end sequencing **facilitates reliable detection of genomic rearrangements and repetitive sequence elements, a higher number of SNV calls following read-pair alignment and novel transcripts.**

GenepoweRx's chosen Agilent's sureselect platforms for cancer sample screening significantly improves the accuracy of low variable allele frequency detection using information from both strands (duplex MBC) which is critical in FFPE and liquid biopsy applications. This has been proven with duplex sequencing in Illumina's seq as well.

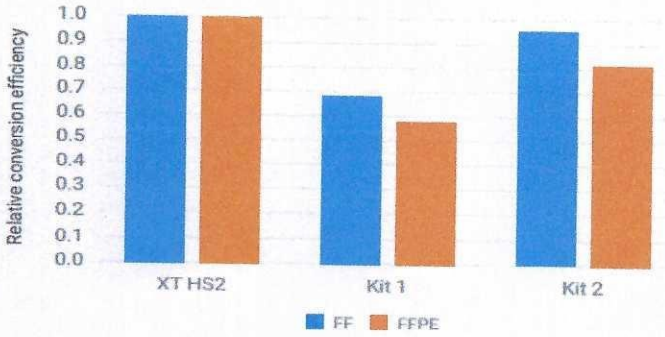
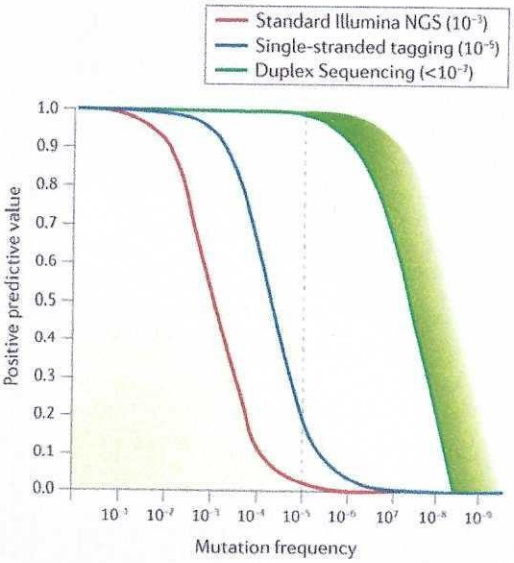


Figure **Improved conversion efficiency for both fresh frozen and FFPE samples.** 10 ng of DNA from lung tumor fresh frozen and normal kidney FFPE (ddCq=1.8, DIN=3.2) samples served as the input for library preparation using the SureSelect XT HS2 DNA reagent kit and two competing kits (Kit 1 and Kit 2). For Kit 1 and Kit 2, libraries were prepared using library preparation kit from vendor K and different adaptors from vendor ID. Sample quality was determined by the Agilent TapeStation (DIN) and Agilent NGS FFPE QC kit (ddCq value). Limited PCR cycles were performed to ensure linear amplification. Relative conversion efficiency was calculated using XT HS2 as reference.



Courtesy: https://www.agilent.com/cs/library/datasheets/public/XT-HS2-Datasheet-5994-1687EN-1-5.pdf ; Nat Rev Genet. 2018 May; 19(5): 269–285. PMID: 29576615

The error rates of standard Illumina Sequencing and single-stranded tag-based error correction result in critical losses in positive-predictive value at variant frequencies of ~ 1/100 and 1/1000 respectively. The extremely low error rate conferred by Duplex Sequencing enables confident identification of variants below 1/100,000

6. Data Analysis and Report generation

GenepoweRx uses standardardized and well validated pipelines with annotations from FDA approved database. GenepoweRx collaborated with Memorial Sloan Kettering Cancer Center (MSK), the world's oldest and largest private cancer centre to utilize MSK's clinical and research insights into gene mutations associated with solid tumors. MSKs clinical information along with GenepoweRx proprietary database information is utilized to provide accurate recommendations for Indian cancer populace.

7. Sequencing cost

With paired-end sequencing, hybrid-capture technology, high sequencing coverage & depth and Illumina's Novoseq 6000 NGS Analyzer, in-depth deep sequencing & analysis is carried out at GenepoweRx with best combination sequencing parameters and most competitive price.

References:

https://sapac.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

https://sapac.support.illumina.com/documentation.html

https://www.agilent.com/en/product/next-generation-sequencing/hybridization-based-next-generation-sequencing-ngs

Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, Temple-Smolkin RL, Voelkerding KV, Nikiforova MN. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagn. 2017 May;19(3):341-365. doi: 10.1016/j.jmoldx.2017.01.011. Epub 2017 Mar 21. PMID: 28341590; PMCID: PMC6941185.

Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. Nat Rev Genet. 2018 May;19(5):269-285. doi: 10.1038/nrg.2017.117. Epub 2018 Mar 26. PMID: 29576615; PMCID: PMC6485430.

Samorodnitsky E, Datta J, Jewell BM, Hagopian R, Miya J, Wing MR, Damodaran S, Lippus JM, Reeser JW, Bhatt D, Timmers CD, Roychowdhury S. Comparison of custom capture for targeted next-generation DNA sequencing. J Mol Diagn. 2015 Jan;17(1):64-75. doi: 10.1016/j.jmoldx.2014.09.009. PMID: 25528188; PMCID: PMC4279426.

ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature. 2020 Feb;578(7793):82-93. doi: 10.1038/s41586-020-1969-6. Epub 2020 Feb 5. PMID: 32025007; PMCID: PMC7025898.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013 Mar 29;339(6127):1546-58. doi: 10.1126/science.1235122. PMID: 23539594; PMCID: PMC3749880.

Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014 Jan 23;505(7484):495-501. doi: 10.1038/nature12912. Epub 2014 Jan 5. PMID: 24390350; PMCID: PMC4048962

Larotectinib for treating NTRK-fusion positive solid tumors. Technology appraisal guidance Published: 27 May 2020 www.nice.org.uk/guidance/ta630

Nakazato T, Ohta T, Bono H. Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive. PLoS One. 2013 Oct 22;8(10):e77910. doi: 10.1371/journal.pone.0077910. PMID: 24167589; PMCID: PMC3805581.

Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, et.al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nat Biotechnol. 2013 Nov;31(11):1023-31. doi: 10.1038/nbt.2696. Epub 2013 Oct 20. PMID: 24142049; PMCID: PMC5710001.

Pritchard CC, Salipante SJ, Koehler K, Smith C, Scroggins S, Wood B, Wu D, Lee MK, Dintzis S, Adey A, Liu Y, Eaton KD, Martins R, Stricker K, Margolin KA, Hoffman N, Churpek JE, Tait JF, King MC, Walsh T. Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. J Mol Diagn. 2014 Jan;16(1):56-67. doi: 10.1016/j.jmoldx.2013.08.004. Epub 2013 Nov 2. PMID: 24189654; PMCID: PMC3873496.

Wagle N, Berger MF, Davis MJ, Blumenstiel B, Defelice M, Pochanard P, Ducar M, Van Hummelen P, Macconaill LE, Hahn WC, Meyerson M, Gabriel SB, Garraway LA. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer Discov. 2012 Jan;2(1):82-93. doi: 10.1158/2159-8290.CD-11-0184. Epub 2011 Nov 7. PMID: 22585170; PMCID: PMC3353152.

Wagle N, Van Allen EM, Treacy DJ, Frederick DT, Cooper ZA, et.al. MAP kinase pathway alterations in BRAF-mutant melanoma patients with acquired resistance to combined RAF/MEK inhibition. Cancer Discov. 2014 Jan;4(1):61-8. doi: 10.1158/2159-8290.CD-13-0631. Epub 2013 Nov 21. PMID: 24265154; PMCID: PMC3947296.

Bentley DR, Balasubramanian S, Swerdlow HP, et.al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6;456(7218):53-9. doi: 10.1038/nature07517. PMID: 18987734; PMCID: PMC2581791.

Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010 Oct;11(10):685-96. doi: 10.1038/nrg2841. PMID: 20847746.

Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, Nusbaum C, Jaffe DB. Characterizing and measuring bias in sequence data. Genome Biol. 2013 May 29;14(5):R51. doi: 10.1186/gb-2013-14-5-r51. PMID: 23718773; PMCID: PMC4053816.

Bewicke-Copley F, Arjun Kumar E, Palladino G, Korfi K, Wang J. Applications and analysis of targeted genomic sequencing in cancer studies. Comput Struct Biotechnol J. 2019 Nov 7;17:1348-1359. doi: 10.1016/j.csbj.2019.10.004. PMID: 31762958; PMCID: PMC6861594.

Petrackova A, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, Papajik T, Kriegova E. Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. Front Oncol. 2019 Sep 4;9:851. doi: 10.3389/fonc.2019.00851. PMID: 31552176; PMCID: PMC6738196.

