NGS Informatics

Efficient Biorepository Development using Oncomine Precision Assay and Genexus Integrated NGS Platform

Stephen Wunsch, Vinay Mittal, Amir Marcovitz, Jeoffrey Schageman, Tony Cooper, April Bigley, Nader Ezzedine, Diarra Hassell, Daniela Garcia, David Chi, Scott Myrand, Andrew Hatch, Seth Sadis, Kelli Bramlett

Thermo Fisher Scientific, Carlsbad, CA; South San Francisco, CA; Austin, TX; Ann Arbor, MI

Introduction

The development of tumor sample biorepositories is critically important to enable the development and analytical validation of next-generation sequencing (NGS) and other molecular assays. However, sequencing FFPE and plasma samples for the presence of relevant variants of interest has traditionally been a slow, labor-intensive, and typically expensive endeavor. To this end, we paired the targeted AmpliseqHD™ Oncomine Precision Assay (OPA) with the Genexus integrated NGS platform and successfully sequenced over 20,000 FFPE and plasma biospecimens, combining low input of DNA and RNA with rapid turn-around time and limited hands-on time.

Materials and methods

Sample Preparation

- KingFisher Flex Purification System was used for extraction and quantification of DNA/RNA concentrations performed via Quant-iT on FloroSkan

Workflow Method

- Automated Genexus workflow: Library prep, sequencing, and data analysis all performed by Genexus instrument

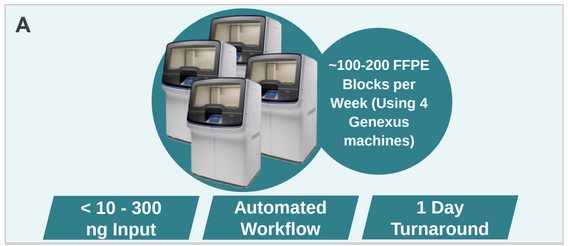


Figure 1(A). Sequencing setup for external vendors. (B) OPA Panel gene content table

Hardware/Software

- Torrent Variant Caller for annotation of VCF. RStudio, Plotly, ggplot2, and React used to generate plots/User Interface (UI)

FFPE Sample Sequencing

Genexus FFPE Sequencing Distributions

Utilizing 4 Genexus instruments, vendors were able to externally sequence 100-200 FFPE samples per week (16 samples/run), with minimal hand-on time and system failures (Figure 2). Runs completed without system failure at >94% rate. Figure 3 demonstrates the diversity of samples sequenced: Lung, Breast, and Colon being among the top.

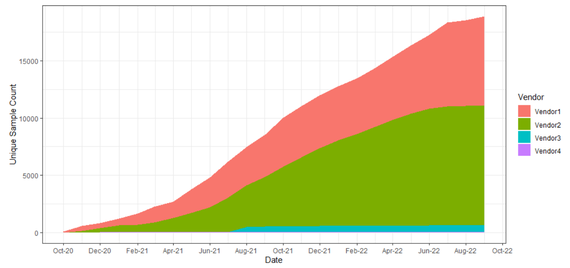


Figure 2. Area plot of cumulative FFPE unique sample counts sequenced per month by each Vendor

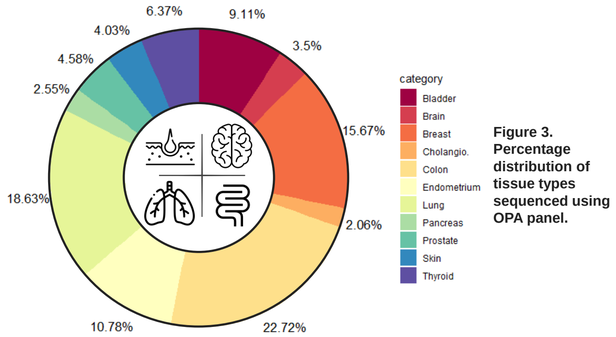


Figure 3. Percentage distribution of tissue types sequenced using OPA panel.

Sequencing Results

Table 1 indicates the number of unique FFPE/Plasma samples sequenced on Genexus with the OPA panel, and average DNA/RNA sequencing metrics.

Type	Unique Sample Count	DNA Mapped Reads	RNA Mapped Reads	DNA MRL*	RNA MRL	MAPD	Uniform. Base Cov.
FFPE	18,140	976,973	136,231	85	66	0.4668	91.87%
Plasma	3,570	10,938,958	271,147	102**	NA	0.2225	98.82%

Table 1. Sequencing metric averages for FFPE and Plasma samples using OPA.
*** MRL: Mean Read Length**
****Plasma MRL calculated together for DNA and RNA**

Variant Analysis Results

DNA Variant Analysis

Out of the 18,140 unique FFPE samples analyzed, 14,362 (79.2%) contained at least one positive SNV/INDEL mutation (Figure 4). 2,821 (15.6%) of samples contained positive CNV variations, with an average MAPD of 0.28.

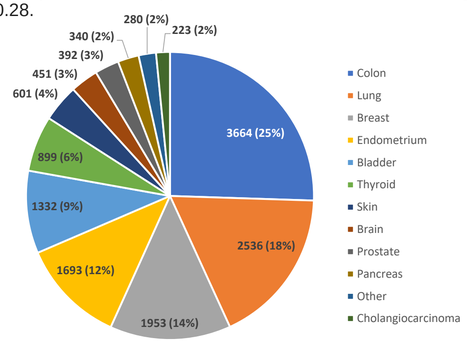
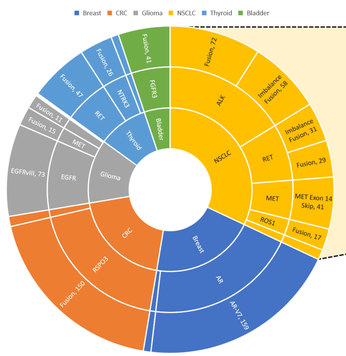


Figure 4. Pie chart displaying counts and percentages for unique FFPE samples containing positive SNV/INDEL mutations found per tissue type

RNA Variant Analysis

Utilizing dynamic fusion calling, targeted/non-targeted fusions, fusion detection by proprietary exon tiling imbalance, and RNA exon variations were observed in sequencing results obtained from FFPE samples. NCSLC had the highest frequency of positive fusions with ALK as the top isoform driver gene.



Driver Gene	Partner Gene	Count (Freq. %)	Avg. Read Count	Avg. Mol. Count
ALK	EML4	63 (2.060)	799	54
MET	MET	41 (1.341)	10,658	798
RET	KIF5B	18 (0.589)	2,478	302
ROS1	CD74	8 (0.262)	3,986	476
NRG1	CD74	7 (0.229)	5,653	383
RET	CCDC6	6 (0.196)	1,908	176
ROS1	SDC4	4 (0.131)	4,045	597
ROS1	EZR	3 (0.010)	1779	212

Figure 5. Sunburst plot of positive RNA variations found across diverse FFPE tissue types

Table 2. Top 8 driver/partner gene combinations observed in FFPE NSCLC samples with positive RNA variations. Average Read and Molecular Counts displayed in last two columns.

Biorepository

An automated pipeline was used to seamlessly parse Genexus sequencing data and archive results to an internal database (DB). The DB can be queried by user interface (UI) to explore diverse sample metadata and compare sequencing results (Figure 6).

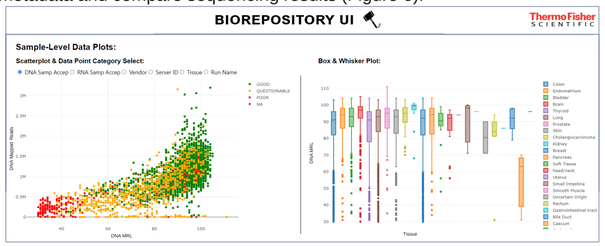


Figure 6. Developed UI/DB for accessing repository data and compare sequencing performance

Conclusions

Next day 'sample to sequencing results' turn-around time, allowed Biobank vendors to analyze samples at high-throughput volume, with reduced need for manual intervention (hands-on time). Over the course of 2 years >20,000 FFPE and Plasma samples were sequenced using the OPA panel on Genexus, delivering the following critical endpoints:

- Fast access to sequencing results with >94% sequencing run success rate with samples of varying quality
- The frequency of DNA and RNA alterations were comparable to reference studies
- Pooled sample metadata and sequencing results analyzed within Custom Database

Disclaimer

For Research Use Only. Not for use in diagnostic procedures.

Acknowledgements

We would like acknowledge the sequencing work performed externally by Biorepository vendors.

Science at a scan
Scan the QR code on the right with your mobile device to download this and many more scientific posters.

Learn more at thermofisher.com/FURL