thermoscientific

Development of Quality Control Materials for Characterization of Comprehensive Next-Generation Sequencing Panels Targeting Cancer Hotspots

Mark Christian[1], Samyuktha Dasari[1], Nancy Holihan[2], Geoffrey Bien[3], and Kara L Norman[1]

1. Thermo Fisher Scientific, Clinical Diagnostics Division, Fremont, CA, USA; 2. Health Network Laboratories, Allentown, PA, USA; 3. Thermo Fisher Scientific, Clinical Sequencing Division, Carlsbad, CA USA

ABSTRACT

Introduction: Targeted next-generation sequencing (NGS) panels can detect hundreds of mutations in key genes using amplification based and hybrid-capture based NGS technologies. Although NGS technology is a powerful tool, optimizing and characterizing test performance on hundreds of variants is extremely challenging, time consuming, and expensive. Samples must be sourced, variants identified and orthogonally confirmed, then quantified and diluted. This effort is then multiplied across dozens of samples, and then samples must be run over many runs and days to assess assay reproducibility, precision, sensitivity, etc. In this study, we developed a novel reference material, experimental design, and analysis pipeline that allows for highly streamlined NGS assay characterization, enabling thorough test characterization across 500+ variants within only 6 runs.

Methods: The AcroMetrix™ Hotspot Frequency Ladder was developed based on the highly characterized NIST Genome in a Bottle GM24385 DNA as the background genome. 555 SNV, Indel and MNVs were multiplexed into one sample, and a frequency ladder was created, targeting 50%, 25%, 15%, 10%, 5%, 2.5% and 0% allelic frequencies. Variant Allele Frequency (VAF) was confirmed using the Bio-Rad® Droplet Digital™ PCR (ddPCR™) system. To assess the ability of the frequency ladder to characterize assay performance, six independent replicate libraries were prepared for each of the seven levels. Libraries were run over six runs on the Ion Torrent™ Ion S5™ XL system using Oncomine Comprehensive Assay v3, allowing for 6 replicates of all variants at decreasing frequencies.

Results: Digital PCR VAF was observed to be 49.98%, 26.30%, 15.83%, 10.56%, 5.12% and 2.77% which compares well with target VAF. NGS data also showed good correlation and the observed frequencies were within ±20% of the expected target VAF for 131 variants covered by the assay. The sensitivity was >90% for 50%, 25% and 15% VAF controls. However, the sensitivity dropped to >70% for 5% and 10% VAF controls. Limit of detection was determined for all variants covered by the assay, using six replicates of seven frequency levels across six runs. Replicate data was also used to calculate assay reproducibility, precision, and specificity.

Conclusion: Carefully designed quality control materials, experimental design and analysis pipeline significantly accelerated the assessment of reproducibility, precision, sensitivity, and limit of detection of an NGS assay. Exhaustive analytical characterization of the NGS panel was completed across hundreds of variants within 3 days.

- Built on NIST Genome in a Bottle GM24385
- GM24385 DNA is the first NIST reference material for genome sequencing
- Platform agnostic
- Highly multiplexed control

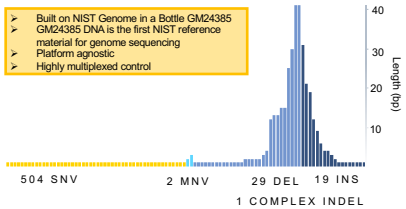

Figure 1. MegaMix™ Technology: Built on AcroMetrix™ MegaMix™ Technology + NIST Genome in a Bottle. 550+ common variants confirmed by sanger sequencing.

ABL1	CDH1	ERBB4	FLT3	HRAS	KIT	MSH6	PTEN	SMO
AKT1	CDKN2A	EZH2	FOXL2	IDH1	KRAS	NOTCH1	PTPN11	SRC
ALK	CSF1R	FBXW7	GNA11	IDH2	MAP2K1	NPM1	RM1	STK11
APC	CTNNB1	FGFR1	GNAQ	JAK2	MET	NRAS	RET	TP53
ATM	EGFR	FGFR2	GNAS	JAK3	MLH1	PDGFRA	SMAD4	VHL
BRAF	ERBB2	FGFR3	HNF1A	KDR	MPL	PIK3CA	SMARCB1	

Table 1. Core Targeted Genes: 555 mutations including 504 SNV, 2 MNV, 29 DEL, 19 INS and 1 complex indel all in one control sample. Data was generated for all hotspots, however analysis was only conducted on a subset of 42 relevant hotspots for genes shaded in blue above. See figure 7 for COSMIC IDs of each hotspot.

Frequency Ladder Format	VAF (%)
Panel Member 1	0% (Negative)
Panel Member 2	2.5%
Panel Member 3	5%
Panel Member 4	10%
Panel Member 5	15%
Panel Member 6	25%
Panel Member 7	50%

Table 2. Frequency Ladder Format: The frequency ladder features 555 hotspot variants in ascending %VAF from negative to 50%.

EXPERIMENTAL DESIGN FOR ANALYTICAL VALIDATION								
	Day 1	Operator 1 Instrument 1						
Run 1	Negative	50%	25%	15%	50%	25%	15%	
Run 2	Negative	10%	5%	2.50%	10%	5%	2.5%	
	Day 2	Operator 2 Instrument 1						
Run 1	Negative	50%	25%	15%	50%	25%	15%	
Run 2	Negative	10%	5%	2.50%	10%	5%	2.5%	
	Day 3	Operator 1 Instrument 2						
Run 1	Negative	50%	25%	15%	50%	25%	15%	
Run 2	Negative	10%	5%	2.50%	10%	5%	2.5%	
EXPERIMENTAL DESIGN FOR THIS STUDY								
	Day 1, 2, &3	Operator 1 Instrument 1						
Run 1	Negative	50%	25%	15%	50%	25%	15%	
Run 2	Negative	10%	5%	2.50%	10%	5%	2.5%	

Table 3. Testing Format: The experimental design for analytical validation is intended to assess multiple characteristics of NGS assay variability while minimizing time spent on additional testing events. For the purposes of this study, only runs indicated in the lower panel were run.

RESULTS

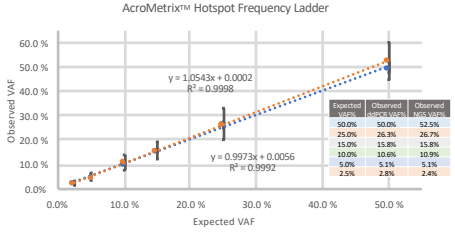

Figure 2. AcroMetrix™ Hotspot Frequency Ladder ddPCR Results: AcroMetrix™ MegaMix™ technology containing 550+ variants was used to prepare highly multiplexed samples at six different allelic frequencies. The blue line above is measurements observed on ddPCR, while the orange is measurements observed on NGS.

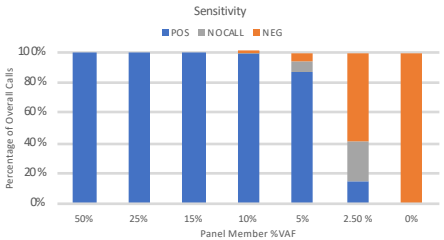

Figure 3. Frequency Ladder Sensitivity: The columns above represent variant calls for 42 hotspots from 6 replicates for each panel member (n=252 per column). Panel Members 7, 6, and 5 (50%, 25%, and 15% VAF) have 100% sensitivity. Panel Member 4 (10% VAF) sensitivity is 99.21%. Panel Member 3 (5% VAF) has some replicates drop out, giving a specificity of 87.70%. The majority of variants for panel member 2 (2.5% VAF) expectedly drop out as they are below the variant caller cutoff.

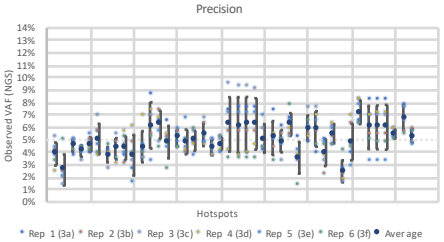

Figure 4. Precision Assessment for Panel Member 3 (5% VAF): The graph above displays data for all 6 replicates of each of the 42 COSMIC IDs assessed for Panel Member 3 at 5% VAF. Precision varied by hotspot. Average standard deviation across hotspots was 1.2% observed VAF.

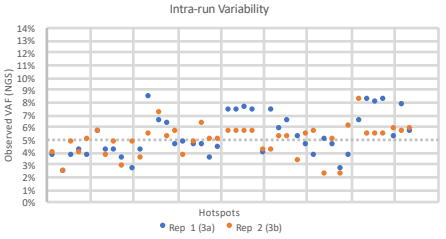

Figure 5. Intra-Run Variability for Panel Member 3 (5% VAF): Replicates from the same chip, (replicate a and b), were used to assess intra-run variability displayed above. Replicates displayed an average variation of 1.2%, 1.5%, and 1.6% VAF intra-run when evaluating replicate a and b, c and d, and e and f, respectively.

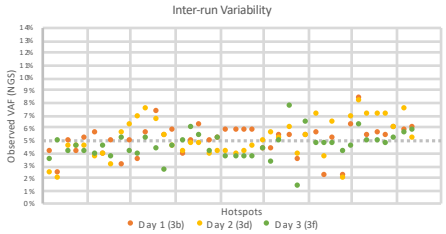

Figure 6. Inter-run Variability for Panel Member 3 (5% VAF): Replicates across days (samples b, d, and f) were used to assess inter-run variability. Replicates displayed an average variation of 1.2% VAF across days.

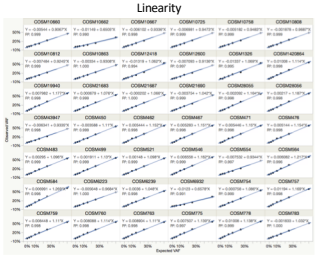

Figure 7. Linearity Assessment: A plot of VAF observed on NGS versus VAF expected for each hotspot assessed yielded equation values plotted above. All hotspots analyzed had R²values of ≥0.99.

	Analytical Validation Metrics
Reproducibility	X
Inter-Operator	x
Intra-run	x
Inter-run	x
Inter-instrument	x
Accuracy	X
Sensitivity	x
Specificity	x
Cross-sample contamination	X
On-instrument	x
Cross-library prep	x
Limit of Detection	X
Coverage	x
Frequency	x
DNA Input	x

Table 4. Analytical Validation Data Generated: The AcroMetrix™ Frequency Ladder can be used to generate data for the listed analytical validation metrics.

CONCLUSIONS

- A comprehensive control using a focused study format can generate sufficient data to perform analytical validation of a diverse NGS assay in 3 days.
- Assay performance and detection varied between targeted analytes and runs. All variants showed high linearity with expected %VAF across the full range of the NGS assay (0-50%).
- Analytical validation was completed on 42 variants in 6 runs over 3 days.

ACKNOWLEDGEMENTS

David Chi, Laura Ramsundar-Otero, Liang-Chun Liu, Lu Zheng, Imteaz Siddique, Kunal Banjara

TRADEMARKS

This control is currently in development and not available for use.