Reference Standards for the validation of myeloid sequencing assays

Ming Wang[1], Hannah Child[1], Jesse Stombaugh[1], Maureen A. Jakubowski[2], Concetta R. Peck[2], James R. Cook MD, PhD.[2], Yu-Wei Cheng PhD.[2], Zheng J. Tu PhD.[2] and Julie Wickenden[1].

[1] Horizon Discovery, [2] Cleveland Clinic, Ohio

Introduction

Myeloid neoplasms are diseases of the hematopoietic stem cells. They can be categorized into five primary types based on WHO classification, with Acute Myeloid Leukemia (AML) occurring at the highest incidence. With the advent of more robust sequencing capabilities, pathogenic mutations have now been identified in genes from five main groups: signaling pathway proteins (e.g. *CBL*), transcription factors (e.g. *RUNX1*), epigenetic regulators (e.g. *ASXL1*), tumor suppressors (e.g. *TP53*) and components of the spliceosome (e.g. *SF3B1*). A recent influx of new myeloid gene panels for sequencing of these genes requires rigorous validation to ensure sufficient clinical accuracy, as errors can often occur in library preparation, sequencing, analysis and interpretation of results. In order to support this effort, Horizon has developed a novel cell line-derived reference standard containing 22 variants across 19 genes involved in all of the main pathway groups, including *CBL, RUNX1, ASXL1, TP53, SF3B1, JAK2, FLT3, ABL, KRAS* and *DNMT3A*.

Methods

Engineered cancer lines were single-cell diluted and investigated for their genomic status by Sanger sequencing at key loci identified as relevant to myeloid cancer in the literature. Target variants were then further characterized by ddPCR, to enable the variant allele frequency to be reliably determined. The material was sent to seven test labs for assessment across a range of mainstream myeloid NGS gene panels *(listed in Table 1)* and compared to the gold standard ddPCR data.

Lab ID	Location	Myeloid NGS Test Assay
Lab 1	USA	Custom ArcherDx VariantPlex panel
Lab 2	USA	Archer Myeloid VariantPlex panel
Lab 3	USA	Custom 62 gene myeloid panel
Lab 4	USA	Custom 62 gene myeloid panel
Lab 5	USA	Custom 63 gene myeloid panel
Lab 6	USA	Proprietary custom myeloid panel
Lab 7	UK	Illumina TruSight panel

Table 1: List of the 7 NGS myeloid beta test centres

Conclusion

This is the first fully cell line-derived Myeloid Reference Standard containing 22 relevant mutations across 19 genes implicated in the diagnosis, management and treatment of myeloid cancers. Results show it to be a consistent, commutable and high quality DNA reference standard that closely mimics DNA from real clinical samples to support the genetic testing of myeloid cancer patients.

Key features such as the 300bp Internal Tandem Duplication *(ITD300)* engineered into the FLT3 gene, and SNP mutations in both IDH1 and IHD2 genes make this a highly clinically relevant reference standard for the validation and optimization of myeloid sequencing tests.

Results

Sequencing by both Sanger and ddPCR confirmed the presence of 22 variants across 19 genes relevant to myeloid cancer, at a blended ratio of 5-70% variant allele frequency. Multiple reproducible batches of gDNA were tested producing consistent sequencing results. The reference material displayed good coverage across some of the most commonly used clinical myeloid NGS gene panels, and performance testing yielded accurate variant calling of the confirmed mutations at the expected allele frequencies as originally verified by ddPCR.

Verified Mutations

Gene	Amino Acid Change	Genomic Position	Reference	Alternative	Variant Type	COSMIC ID	Expected AF(%)	Actual AF (%) (ddPCR results)	Acceptance Criteria (%) Low	Acceptance Criteria (%) High	Pass/Fail
ABL1	T315I	133748283	C	T	SNP	COSM12560	5	4.81	4.00	6.00	PASS
ASXL1	G646fs*12	31022441	A	AG	INS	COSM1411076	40	39.80	36.00	44.00	PASS
ASXL1	W796C	31022903	G	T	SNP	COSM1681610	5	5.08	4.00	6.00	PASS
BCOR	Q1174fs*8	39923086	G	GT	INS	COSM3732385	70	68.80	63.00	77.00	PASS
CBL	S403F	119148988	C	T	SNP	COSM1676499	5	5.52	4.00	6.00	PASS
DNMT3A	R882C	25457243	G	A	SNP	COSM53042	5	4.73	4.00	6.00	PASS
EZH2	R418Q	148514471	C	T	SNP	COSM3259655	5	4.67	4.00	6.00	PASS
FLT3	D835Y	28592642	C	A	SNP	COSM783	5	5.10	4.00	6.00	PASS
FLT3	ITD300	28608047	A	300bp ins	INS	N/A	5	5.55	4.00	6.00	PASS
GATA1	Q119*	48650385	C	T	SNP	N/A	10	10.20	9.00	11.00	PASS
GATA2	G200fs*18	128204841	AC	A	DEL	COSM1418772	35	35.00	31.50	38.50	PASS
IDH1	R132C	209113113	G	A	SNP	COSM28747	5	5.12	4.00	6.00	PASS
IDH2	R172K	90631838	C	T	SNP	COSM33733	5	4.91	4.00	6.00	PASS
JAK2	F537-K539>L	5070021	TTCACAA	T	DEL	COSM24437	5	5.03	4.00	6.00	PASS
JAK2	V617F	5073770	G	T	SNP	COSM12600	5	5.23	4.00	6.00	PASS
KRAS	G13D	25398281	C	T	SNP	COSM532	40	38.70	36.00	44.00	PASS
NPM1	W288fs*12	170837543	C	CTCTG	INS	COSM158604	5	5.06	4.00	6.00	PASS
NRAS	Q61L	115256529	T	A	SNP	COSM583	10	9.92	9.00	11.00	PASS
RUNX1	M267I	36206711	C	T	SNP	COSM1681955	35	35.80	31.50	38.50	PASS
SF3B1	G740E	198266713	C	T	SNP	COSM133120	5	5.27	4.00	6.00	PASS
TET2	R1261H	106164914	G	A	SNP	COSM211643	5	4.90	4.00	6.00	PASS
TP53	S241F	7577559	G	A	SNP	COSM10812	5	5.06	4.00	6.00	PASS

Table 2: Annotation of the 22 variants across 19 genes, with associated COSMIC IDs and confirmation of variant AF by ddPCR

Figure 1: Final product image

Gene	Variant	Expected Allele Frequency (%)	Test Lab ID: Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
ABL1	T315I	5		x	x	x			
ASXL1	G646fs*12	40	x	x	x				x
ASXL1	W796C	5		x	x	x	x		x
BCOR	Q1174fs*8	70		x	x	x	x		x
CBL	S403F	5		x	x	x	x	x	x
DNMT3A	R882C	5	x	x	x	x	x		x
EZH2	R418Q	5		x	x	x	x		x
FLT3	D835Y	5			x	x	x	x	x
FLT3	ITD300	5			x		x		
GATA1	Q119*	10					x		
GATA2	G200fs18	35		x	x		x		x
IDH1	R132C	5	x	x	x	x	x	x	x
IDH2	R172K	5		x	x	x	x	x	x
JAK2	F537-K539>L	5	x	x	x	x			x
JAK2	V617F	5	x	x	x	x	x	x	x
KRAS	G13D	40		x	x	x	x	x	x
NPM1	W288fs*12	5		x	x				x
NRAS	Q61L	10		x	x	x		x	x
RUNX1	M267I	35		x	x	x	x		x
SF3B1	G740E	5		x	x	x	x		x
TET2	R1261H	5		x	x	x			x
TP53	S241F	5	x	x	x	x	x	x	x
Total number of variants called per lab			6	19	21	17	16	8	19

Table 3: NGS myeloid beta test results – across 7 external test centres

300bp FLT3 ITD

Due to high clinical relevance of Internal Tandem Duplications (ITDs) within the FLT3 gene, we engineered a 300bp FLT3 ITD *(termed ITD300)* into the cell lines that make this Reference Standard *(Fig 2A)*. Results show ITD300 was detectable by ddPCR *(Table 2)*, Sanger *(Fig 2B)* and NGS *(Table 3)*.

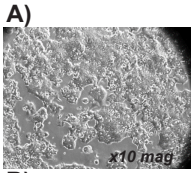

Figure 2: A) Image of engineered HCT116 cells containing the 300b FLT3 insertion
B) Sanger sequencing results of the engineered 300bp FLT3 insertion

In partnership with Test Lab 3:

Cleveland Clinic

horizondiscovery.com