Assuring quality for somatic tumor mutation profiling with the Seraseq™ Solid Tumor Mutation Mix-I (AF20)

Your sample has an unknown mutation profile: how do you know your systems are running at their optimum?

Next-generation sequencing (NGS) has been called "arguably one of the most significant technological advances in the biological sciences in the last 30 years."[1] While Sanger sequencing was the 'gold standard' for cancer mutation analysis, it had been hindered by low throughput and relatively low sensitivity to minor variants, in addition to relatively high cost and long turnaround time. With the advent of massively parallel sequencing, these concerns have been addressed, with much higher throughput, much better sensitivity, at lower cost and decreased turnaround time.

This major development can directly impact the practice of clinical oncology by enabling the examination of a large number of genes in a single assay. Yet there are currently no widely-accepted standard materials for multi-analyte next-generation sequencing-based clinical assays. The ability to compare the accuracy of different assays to process controls is hampered by the lack of broad availability of such standards. In addition, variability introduced by slight changes in NGS library construction, sequencing methodology, and computational analyses can influence the assay result and its resulting interpretation.

In addition to variability introduced by data generation and data analysis, the inherent heterogeneity of tumor tissue and differences in FFPE sample handling adds yet additional challenging dimensions. A recent publication examining over 500 FFPE solid tumor samples found the majority of somatic mutations in the 5% to 30% allele frequency range.[2] In addition, the common method of histopathology to estimate tumor-cell content has a low correlation to its allelic frequency (r^2=0.27, *p=0.0029*); these results indicate the histopathology measurement overestimates the analytically determined minor allele percentage.[3] Yet another source of variation is the manner of formalin fixation and paraffin-embedding; while there is evidence that these processes only have a minimal effect on enrichment uniformity and data variation[4], anecdotal evidence suggests that input FFPE processing variation has wide impact on sequencing quality.

Currently, laboratories that develop their own NGS-based Laboratory Developed Tests (LDT) for somatic mutation sequencing and analysis depend on their own methods for obtaining and characterizing the reference material they use for assuring the precision of their assay. They undertake the work of growing cell lines, characterizing them, purifying DNA, and making mixes of that DNA for their own use. In 2013, the Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Committee published draft guidelines as an educational resource for clinical laboratory geneticists to 'help them provide quality clinical laboratory

genetic service'.[5] Specific sections of these guidelines refer to topics such as Quality Management (§E.1.4, applied to every sequencing run), Clinical Sensitivity (§E.3.4, recommending laboratories 'track and share success rates with other laboratories') and Assay Precision (§E3.6, determining repeatability or within-run variability and determining reproducibility or between-run variability). All of these recommendations need high-quality reference material, manufactured under an appropriate quality system, rather than an improvised 'homebrew' approach.

The current environment for routine characterization of somatic mutations in clinical samples with NGS is reminiscent of the early phase of Sanger sequencing-based and real-time PCR-based assays in molecular diagnostics. In the 1980's, HIV, HCV, and HBV were serious global public health threats with an alarming rise in rate of infection. At that time individual laboratories made their own controls and characterized them in an *ad-hoc* fashion. SeraCare (through its predecessor company BBI) offered the first seroconversion panels for HIV, HCV, and HBV. For more than two decades since, SeraCare has provided specialized reagents and components for quality control in a wide variety of infectious disease tests for clinical laboratories and in-vitro diagnostic manufacturing partners (ACCURUN® is a well-known SeraCare product line serving this purpose). Due to its provenance, SeraCare has a longstanding commitment to and deep experience with process control and cGMP guidelines, in addition to ISO 9001 and ISO 13485 quality management system standards.

The Seraseq Solid Tumor Mutation Mix-I (AF20) takes a novel approach to assessing consistency and precision in somatic mutation sequencing with NGS. Synthetic DNA stretches of 22 actionable driver genes have been designed with two unique features: the first is the engineered mutation, and the second is an artificial 6-base "Internal Quality Marker" inserted within a 25 base-pair distance of that mutation.

The engineered mutations represent SNVs, SNVs as a portion of a homopolymer tract, insertions (either short or long), or deletions (either short or long). The synthetic constructs are quantitated with digital PCR and after normalization, added to the background genomic DNA to a 20% allele frequency ratio. The list of mutations is in Table 1, with their gene name, Catalog Of Somatic Mutations In Cancer (COSMIC, http://cancer.sanger.ac.uk[6]) identifier with mutation type, and coding amino acid change.

The second unique feature of the Seraseq Solid Tumor Mutation Mix-I (AF20) is the insertion of a 6bp 5'-ACATCG-3' sequence as a quality marker. In all cases, this Internal Quality Marker (IQM) is present within a distance of 25 bases from the associated mutation, allowing both features to be analyzed in context on the same sequence read. As a result, the IQM and mutation of interest are present at approximately the same frequencies in final sequence data. As such, the IQM can be used, with its nearby insertion, as a benchmarking aid to evaluate assay performance and pipeline analysis procedures. For example, the inserted sequence is not only easily visible using tools such as the Integrated Genome Viewer (IGV, http://www.broadinstitute.org/igv/ Figure 2), but also can be a shortcut to evaluate a given 'pileup' of reads for a quick estimate of minor allele frequency.

	Gene	COSMIC ID of Mutation	Position (hg19)	CDS	Mutation Type	Amino Acid Change	Target VAF	Ion AmpliSeq™ Cancer Hotpot v2	Illumina TruSeq™ Cancer Panel
1	EGFR	COSM6224	55259515	c.2573T>G	SNV (Homopolymer)	p.L858R	20%	●	○
2	FGFR3	COSM715	1803568	c.746C>G	SNV (Homopolymer)	p.S249C	20%	●	⊙
3	GNAQ	COSM28758	80409488	c.626A>C	SNV (Homopolymer)	p.Q209P	20%	●	⊙
4	AKT1	COSM33765	105246551	c.49G>A	SNV (Homopolymer)	p.E17K	20%	●	●
5	ATM	COSM21924	108117846	c.1058_1059delGT	Small Deletion	p.C353fs*5	20%	●	○
6	SMAD4	COSM14105	48603093	c.1394_1395insT	Small Insertion	p.A466fs*28	20%	●	●
7	NPM1	COSM17559	170837547	c.863_864insTCTG	Large Insertion	p.W288fs*12	20%	●	●
8	EGFR	COSM6225	55242465	c.2236_2250del15	Large Deletion	p.E746_A750delELREA	20%	●	●
9	BRAF	COSM476	140453136	c.1799T>A	SNV	p.V600E	20%	●	●
10	KRAS	COSM521	25398284	c.35G>A	SNV	p.G12D	20%	●	●
11	PIK3CA	COSM775	178952085	c.3140A>G	SNV	p.H1047R	20%	●	●
12	PIK3CA	COSM763	178936091	c.1633G>A	SNV	p.E545K	20%	●	●
13	NRAS	COSM584	115256529	c.182A>G	SNV	p.Q61R	20%	●	●
14	TP53	COSM10648	7578406	c.524G>A	SNV	p.R175H	20%	●	●
15	CTNNB1	COSM5664	41266124	c.121A>G	SNV	p.T41A	20%	●	●
16	IDH1	COSM28747	209113113	c.394C>T	SNV	p.R132C	20%	●	●
17	EGFR	COSM6240	55249071	c.2369C>T	SNV	p.T790M	20%	●	●
18	MPL	COSM18918	43815009	c.1544G>T	SNV	p.W515L	20%	●†	●
19	APC	COSM13127	112175639	c.4348C>T	SNV	p.R1450*	20%	●	●
20	FLT3	COSM783	28592642	c.2503G>T	SNV	p.D835Y	20%	●	●
21	PDGFRA	COSM736	55152093	c.2525A>T	SNV	p.D842V	20%	●	●
22	RET	COSM965	43617416	c.2753T>C	SNV	p.M918T	20%	●	⊙
23	GNAS	COSM27887	57484420	c.601C>T	SNV	p.R201C	20%	●	○
24	TP53	COSM10662	7577538	c.743G>A	SNV	p.R248Q	20%	●	⊙
25	KIT	COSM1314	55599321	c.2447A>T	SNV	p.D816V	20%	●	○
26	JAK2	COSM12600	5073770	c.1849G>T	SNV	p.V617F	20%	●	⊙

Legend:

● Mutation observed to appear at ~20% allelic frequency

⊙ Mutation observed at ~5%-15% allelic frequency due to assay primer consideration

○ Mutation not observed due to assay primer consideration

† Some minor variability observed with strand bias and inadequate coverage

Table 1: *List of mutations included in the Seraseq Solid Tumor Mutation Mix-I (AF20). The presence of the mutation in a particular assay depends upon the enrichment strategy and sequencing platform used. The 26 mutations listed above have been observed to appear at ~20% allelic frequency using the Ion AmpliSeq™ Cancer Hotspot Panel v2 on the Ion Torrent PGM™ Sequencing system, while 17 of these mutations have been observed to appear at ~20% allelic frequency using the Illumina TruSeq® Amplicon - Cancer Panel on the MiSeq® System.*

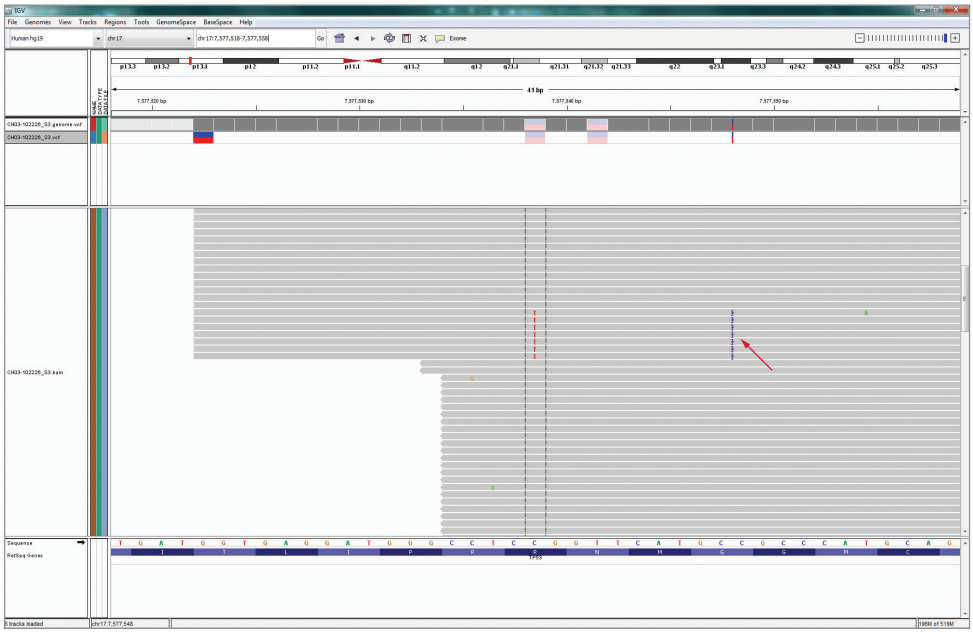

Figure 2: *Integrated Genome Viewer (IGV) screenshot of a TP53 mutation from the Seraseq Solid Tumor Mutation Mix. The six-base Internal Quality Marker (IQM) insert is highlighted.*

Observed variant frequencies may appear higher or lower depending on a number of assay-specific factors. Slight differences in G-C composition between assays can have a profound effect upon relative abundance, for example. Figure 3 illustrates results obtained with two independent cancer panel designs using the Ion AmpliSeq™ technology. Note that while one assay may have a lower observed frequency than the other, a different assay may have the opposite (the designs of these assays, while targeting the same mutation, are independent of each other, and thus perform differently with the same target sequence).

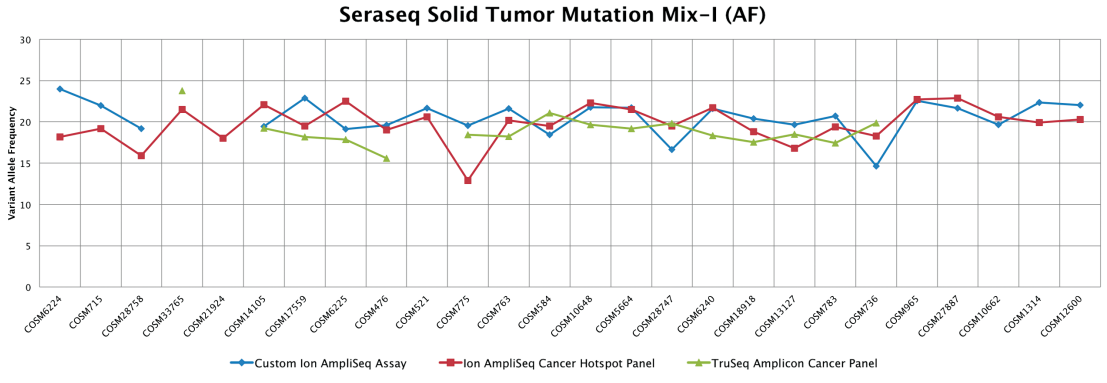

Figure 3: *Seraseq Solid Tumor Mutation Mix-I (AF20) run with two independent Ion AmpliSeq™ cancer panel designs and the TruSeq® Amplicon Cancer Panel run on the MiSeq® Sequencing System.*

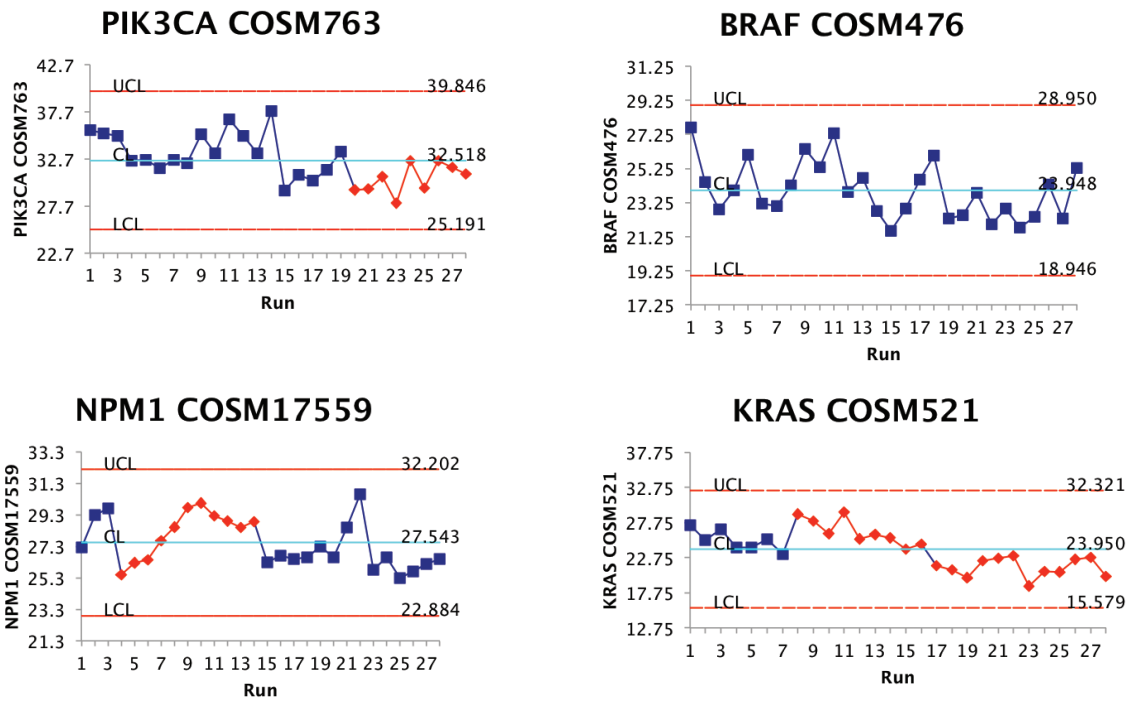

Figure 4: *Levey-Jennings plot of four mutations in the Seraseq Mutation Mix-I (AF20) over the course of 28 runs in 8 months. Data kindly provided by Dr. P. Mickey Williams, Molecular Characterization Laboratory, National Cancer Institute, Frederick MD USA.*

At one large clinical laboratory, the performance of a similar synthetic mix was measured on every run they performed over the course of 8 months. As seen in figure 4, the measurement of allele frequency from the same reference material over time for a given locus could vary as much as plus or minus 5%. One other observation to make from this dataset is the robust stability over time.

Seraseq™ Solid Tumor Mutation Mix-I (AF20)

Item Number: 1600-0098

1 vial, 25 µL per vial at 25 ng/µL concentration

For Research Use Only. Not for use in Diagnostic Procedures.

To learn more about Seraseq™ precision oncology products, visit http://seracare.com/oncology

References

1. Meldrum C, Doyle MA, Tothill RW. Next-generation sequencing for cancer diagnostics: a practical perspective. Clin Biochem Rev. 2011 Nov;32(4):177-95. PubMed PMID: 22147957.

2. Chen K, Meric-Bernstam F, Zhao H, Zhang Q, Ezzeddine N, Tang LY, Qi Y, Mao Y, Chen T, Chong Z, Zhou W, Zheng X, Johnson A, Aldape KD, Routbort MJ, Luthra R, Kopetz S, Davies MA, de Groot J, Moulder S, Vinod R, Farhangfar CJ, Shaw KM, Mendelsohn J, Mills GB, Eterovic AK. Clinical actionability enhanced through deep targeted sequencing of solid tumors. Clin Chem. 2015 Mar;61(3):544-53.doi:10.1373clinchem.2014.23110.

3. Querings S, Altmüller J, Ansén S, Zander T, Seidel D, Gabler F, Peifer M, Markert E, Stemshorn K, Timmermann B, Saal B, Klose S, Ernestus K, Scheffler M, Engel-Riedel W, Stoelben E, Brambilla E, Wolf J, Nürnberg P, Thomas RK. Benchmarking of mutation diagnostics in clinical lung cancer specimens. PLoS One. 2011 May 5;6(5):e19601. doi:10.1371/journal.pone.0019601.

4. Kerick M, Isau M, Timmermann B, Sültmann H, Herwig R, Krobitsch S, Schaefer G, Verdorfer I, Bartsch G, Klocker H, Lehrach H, Schweiger MR. Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. BMC Med Genomics. 2011 Sep 29;4:68. doi: 10.1186/1755-8794-4-68.

5. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, Friez MJ, Funke BH, Hegde MR, Lyon E; Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Commitee. ACMG clinical laboratory standards for next-generation sequencing. Genet Med. 2013 Sep; 15(9):733-47. doi: 10.1038/gim.2013.92.

6. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 2015 Jan; 43(Database issue):D805-11. doi:10.1093/nar/gku1075.

800.676.1881
508.244.6400
info@seracare.com
SeraCare.com