EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR MSK-IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets)

DECISION SUMMARY

A. DEN Number:

DEN170058

B. Purpose for Submission:

De novo request for evaluation of automatic class III designation for the MSK-IMPACT

C. Measurand:

Somatic single nucleotide variants, insertions, deletions, and microsatellite instability in genes in human genomic DNA obtained from formalin-fixed, paraffin-embedded tumor tissue.

Refer to Appendix 1a for complete list of hotspot mutations and Appendix 1b for complete list of genes included in this assay.

D. Type of Test:

Next generation sequencing tumor profiling test

E. Applicant:

Memorial Sloan Kettering (MSK)

F. Proprietary and Established Names:

MSK-IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets)

G. Regulatory Information:

1. Regulation section:

21 CFR 866.6080

2. Classification:

Class II

3. Product code:

PZM

4. **Panel:**

   Pathology

H. Indications for Use:

1. **Indications for Use:**
   The MSK-IMPACT assay is a qualitative in vitro diagnostic test that uses targeted next generation sequencing of formalin-fixed paraffin-embedded tumor tissue matched with normal specimens from patients with solid malignant neoplasms to detect tumor gene alterations in a broad multi gene panel. The test is intended to provide information on somatic mutations (point mutations and small insertions and deletions) and microsatellite instability for use by qualified health care professionals in accordance with professional guidelines, and is not conclusive or prescriptive for labeled use of any specific therapeutic product. MSK-IMPACT is a single-site assay performed at Memorial Sloan Kettering Cancer Center.

2. **Special conditions for use statement(s):**

   For prescription use.

   For in vitro diagnostic use.

3. **Special instrument requirements:**

   Illumina HiSeq™ 2500 Sequencer (qualified by MSK)

I. Device Description:

A description of required equipment, software, reagents, vendors, and storage conditions were provided, and are described in the product labeling (MSK-IMPACT manual). MSK assumes responsibility for the device.

1. **Sample Preparation**:
   The tumor volume and minimum tumor content needed to obtain sufficient DNA for testing to achieve the necessary quality performance are shown in the Table 1 below:

Table 1. Specimen Handling and Processing for Validated Specimen Types

Tissue Type	Volume	Minimum Tumor Proportion	Macrodissection requirements (Based on tumor proportion)	Limitations	Storage
FFPE sections	5-20 unstained sections, 10 microns thick	More than 10% of tumor cells; sections containing >20% viable tumor are preferred. For MSI testing, >25% tumor cells.	Yes, macrodissection to obtain non-neoplastic tissue for analysis	Archival paraffin-embedded material subjected to acid decalcification is unsuitable for analysis because acid decalcification severely damage nucleic acids.	Room temp

Genomic DNA is extracted from tissue specimens per protocol. DNA is quantified and concentrated if necessary. The amount of DNA required to perform the test is 100-250ng. DNA is run in singlicate. DNA shearing is conducted per protocol and a quality control check is performed. Average fragment size should be ~200bp. Sheared DNA is stored at -20°C if not proceeding directly to Library Preparation. The DNA can be stored at 37°C for 10-20 minutes, stored at 2–8°C for 24 hours, or at –20°C for longer periods.

2. **Library Preparation:**
Sequence libraries are prepared using KAPA Biosystems Library Preparation Reagents by first producing blunt-ended, 5'-phosphorylated fragments. To the 3' ends of the dsDNA library fragments, dAMP is added (A-tailing). Next, dsDNA adapters with 3'dTMP is ligated to the A-tailed library fragments. Library fragments with appropriate adapter sequences are amplified via ligation-mediated pre-capture PCR. A quality control check on the amplified DNA libraries is performed: Samples should be a smear; average fragment size with the peak at ~200bp; and concentration between 5-300ng/μL to ensure adequate hybridization for capture.

3. **Hybrid Capture NGS:**
Library capture is conducted using NimbleGen Capture reagents. Pooled sequencing libraries are hybridized to the vendor oligo pool. Capture beads are used to pull down the complex of capture oligos and genomic DNA fragments. Unbound fragments are washed away. The enriched fragment pool is amplified by ligation mediated-PCR. The success of the enrichment is measured as a quality control step: Samples should be a smear, average fragment size with the peak at ~300bp; the concentration of the amplified DNA library should be 5-45ng/μL; the LM-PCR yield should be ≥ 250ng. Reactions can be stored at 4°C until ready for purification, up to 72 hours.

4. **Sequencing and Data Analysis:**
Sequencing is conducted with the Illumina HiSeq2500 Sequencing Instruments and reagents and PhiX Control v3. The sequencing process uses multiple quality checks.

   a) **Data Management System (DMS):** Automated sample tracking and archival of run-associated metadata (barcode, run name, samples accession number, patient medical record number, source (class), specimen type, and panel version) is conducted with the following key functions: Tracking sample status through various stages of data analysis; tracking iterations of analysis applied to a given sample; recording versions of databases and algorithms used in analysis; archival of selected pipeline output files (FASTQ, BAM, VCF) and sequencing run statistics (e.g., cluster density, %clusters passing filter, unassigned read indices).

   b) **Demultiplexing and FASTQ generation:** The analysis pipeline uses software provided by Illumina. Two FASTQ files are generated per samples corresponding to full length forward and reverse reads. Demultiplexing quality control includes quality metrics for per-base sequence quality, sequence content, GC content and sequence length distribution, relative percentages of unmatched indices.

c) Indexing QC check: The potential for index contamination is managed by demultiplexing all sequencing reads for all possible barcodes. If the number of reads > 15,000 for any unused barcodes, then those reads are analyzed with the pipeline and the fingerprint SNPs are used to identify which of the barcodes used in the pool could be causing the appearance of extra reads.

d) Read alignment and BAM generation: Spurious adapter sequences are trimmed prior to read alignment. Reads are aligned in paired-end mode to the hg19 b37 version of the human genome. Aligned reads are written to a Sequence Alignment Map (SAM) file, which is then converted into Binary Alignment Map (BAM) format. PCR duplicates are removed. Each base within a read is assigned a base quality score by the sequencing software, which reflects the probability an error was made with the base call. To account for systemic biases that may not accurately reflect the actual error probabilities observed empirically, the analysis pipeline uses another tool to adjust the reported quality scores based on the selected covariates. Reassigned quality scores are subject to a threshold of 20, corresponding to a 1/100 chance of error.

e) Sample QC checks: The baits used for hybridization capture include custom intergenic and intronic probes targeting >1000 regions throughout the genome containing common single nucleotide polymorphisms (SNPs). The unique combination of SNPs specific to a given sample serves as a 'fingerprint' for the identity of the corresponding patient, and serves to identify potential sample mix-ups and contamination between samples and barcodes. QC checks involving the use of these 'fingerprint' SNPs are detailed below:

i. *Sample mix-up check:* The analysis pipeline computes the 'percent discordance' between a reference and query sample, defined as the percent of homozygous sites in the reference sample that are homozygous for the alternate allele in the query sample. The expected discordance between tumors and their respective matched normal should be low (<5%). Conversely, the expected discordance between samples from different patients should be high (~ 25%). Pairs of samples from the same patient with > 5% discordance ("unexpected mismatches") and from different patients with <5% discordance ("unexpected matches") are flagged.

ii. *Sample contamination checks*: Alternate alleles (percent heterozygous) at homozygous SNP sites (fingerprint SNPs) are assessed. A sample is flagged for review if the average minor allele frequency at these SNPs exceeds 2%.

iii. *Check for presence of tumor in normal:* Normal samples are expected to be free of known SNVs and insertions and deletions (indels) that are commonly (somatically) recurrent in tumor samples. As a first pass check, the pipeline genotypes normal samples at several known 'hotspot' locations derived from somatic mutation catalogs. If a known tumor-specific mutation (i.e. BRAF V600E) is detected with mutation frequency > 1% in a normal sample, the normal sample is flagged for review and possible exclusion from analysis. Tumor

samples with matched normal controls excluded due to possible tumor contamination will be considered as unmatched tumor samples for subsequent analyses.

f) Mutation calling – SNVs and Indels: The analysis pipeline identifies two classes of mutations: (1) single nucleotide variants (SNVs) and (2) indels. Paired sample mutation calling is performed on tumor samples and their respective matched normal controls. In instances where a matched normal sample is unavailable, or where the matched normal sample was sequenced with low coverage (< 50X), tumor samples will be considered as unmatched samples, and will be compared against a standard, in-batch pooled FFPE normal control for mutation calling. Filtering is performed to remove low quality sequence data, sources of sequencing artifacts, and germline results.

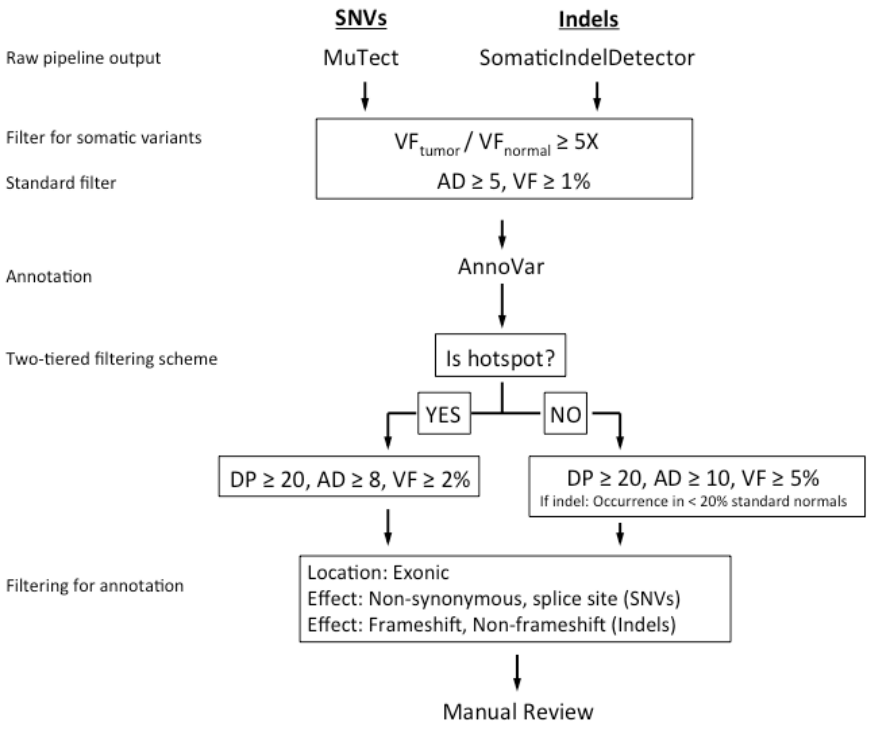
i. *Analysis of pooled FFPE positive and negative controls*: data from controls is used to confirm lack of contamination as well as analytical sensitivity.

ii. *Filters on sample coverage:* A sequence coverage ≥ 100X is required to achieve 95% power to detect mutations with underlying variant frequency of 10% or greater. To ensure that at least 98% of targeted exons meet this coverage, a per sample coverage requirement has been conservatively set at ≥ 200X. A lower coverage threshold for the matched normal is set at 50X.

iii. *Filtering for high confidence mutations:* Raw SNV and indel calls are subjected to a series of filtering steps to ensure only high-confidence calls are admitted to the final step of manual review. These parameters include (1) evidence of it being a somatic mutation (i.e., ratio between mutation frequencies in the tumor and normal samples to be ≥ 5.0); (2) whether the mutation is a known hotspot mutation (refer to Appendix 1a for details); (3) reference on in house 'standard normal' based on common artifacts; (4) technical characteristics that use coverage depth (DP), number of mutant reads (AD), mutation frequency (VF).

The filtering scheme and threshold are shown in Figure 1 below. The threshold values for the filtering criteria were established based on paired-sample mutation analysis on replicates of normal FFPE samples, and optimized to reject all false positive SNVs and almost all false positive indel calls from the reference dataset.

Figure 1. Summary of mutation filtering scheme



g) Mutation annotation: Predicted functional effect and clinical interpretation for each mutation is curated by automated software using information from several databases.

h) Microsatellite Instability (MSI) status calling: The somatic MSI status is inferred by interrogating all available genomic microsatellites covered by MSK-IMPACT within tumor samples against the matched normal DNA using the program MSIsensor (Nui B et al. 2014). Essentially, the sequencing results are analyzed via MSIsensor to assess the number and length of homo-polymers / microsatellites within the targeted regions of tumor-normal sample pair. This results in a continuous rather than categorical MSI score assignment for the tumor sample. Loci are considered unstable (somatic) if k-mer distributions are significantly different between the tumor and matched normal using a standard multiple testing correction of $\chi 2$ p-values. The percentage fraction of unstable sites is reported as the MSIsensor score. The assay uses a MSIsensor score threshold of 10 or greater to define MSI-H by MSIsensor.

5. Controls

a) Matched normal control: Genomic DNA is extracted from patient-matched normal tissue (when available) or peripheral blood, for use as a matched normal control. In the event a matched normal is unavailable, or where the matched normal sample was sequenced with low coverage (<50X), tumor samples will be compared against a standard, in-batch pooled FFPE normal control for mutation calling; mutations called under these circumstances may include rare germline mutations and cannot be guaranteed to be somatic.

b) Positive control: The positive control sample is a mixture of 3 tumor samples, each sample with a different confirmed SNV and at least one insertion or deletion, representing a range of mutation allele frequencies. Results are compared against a pooled FFPE negative control as an unmatched normal. Data generated from the mixed positive control sample are analyzed using the pipeline, and frequencies of the detected mutations are reviewed to determine if (1) the known mutations are among those called, and (2) the observed frequencies for the known mutations match their expected values within 5% of their values. The mixed FFPE positive control sample pools with expected variant frequency (VF) prior to pooling are shown in Table 2.

Table 2. Positive Controls and Expected Mutation Frequencies

Mixed Positive ID	Sample ID	VF	Known Mutation
M-1913-BF	M-1682-C3-T	17%	KRAS Q61H
	M-1791-8C-T	66%	EGFR L858R
	M-1754-DB-T	61%	KITexon9ins
M-1914-A2	M-1671-CE-T	25%	KITexon11del
	M-1693-5E-T	24%	PIK3CA H1047R
	M-1646-FC-T	41%	BRAF V600E
M-1915-CA	M-1612-28-3-T	32%	EGFR exon19 del
	M-1627-D9-T	52%	NRAS Q61H
	M-1625-1A-2A-T	28%	KRAS G12D

c) Negative control: The negative control sample is a mixture of FFPE normal samples verified in previous reruns to be free of tumor contamination and germline copy number mutations in target genes. Polymorphisms unique to each constituent normal sample in the pool have been identified in prior analyses and the expected frequencies for each polymorphism in the pooled negative control are confirmed. The observed mutation frequencies are compared against the expected mutation frequencies for the 862 common SNPs, and the degree of concordance is measured using Pearson's correlation. The correlation between expected and observed mutation frequencies is expected to be 0.9 or higher.

d) PCR reagent control [No Template Control (NTC)]: The NTC control should have a Qubit measurement of < 1.0ng/μL. Sequencing data from the NTC control sample will also be subjected to analysis using the pipeline, to verify that no known hotspot mutations are detected. Similar to the pooled FFPE negative control, if a hotspot mutation is detected, any samples containing that mutation in the pool will be reviewed to determine if a re-run is necessary.

6. <u>Result Reporting:</u>

- Oncopanel results are reported out under one of the two categories: "Cancer Mutations with Evidence of Clinical Significance" or "Cancer Mutations with

Potential Clinical Significance". The two categories are based on the supporting level of clinical evidence. Refer to the Clinical Performance Section for more information.

- Results are reported for point mutations and small insertions and deletions in protein-coding exons of the 468 gene panel. Refer to Appendix 1b for a list of genes.
- The MSK-IMPACT does not report mutations in 73 exons due to consistently low coverage in those exons. Refer to Appendix 1c for a list of excluded exons.
- Reporting takes in account the following quality metrics in the Table 3 below.

Table 3. Sample Level Quality Control Metrics

QC Metrics	Acceptance Criteria
Coverage	Average target coverage > 200X
Coverage Uniformity	≥ 98% target exons above 100X coverage
Base Quality	> 80% of bases with QS above > Q30
% Cluster passing	The percent cluster passing filter (Cluster PF) > 80%
% Reads passing filter	The percent reads passing filter (Reads PF) > 80%
Hotspot Mutation* calling threshold	Mutation Coverage (DP) ≥ 20, Number of Mutant Reads (AD) ≥ 8, Mutation Frequency (VF) ≥ 2%
Non-hotspot Mutation threshold**	DP ≥ 20, AD ≥ 10, VF ≥ 5%
Indels	Fewer than 20% of samples in an established 'standard normal'database
Positive Run Control	The difference between the observed and expected frequencies for the known mutations should be within 5%.
Negative Run Control	The correlation between expected and observed mutation frequencies should be 0.9 or higher
Sample-Mix up QC	Check over 1000 custom intergenic/intronic "fingerprint" SNPs. Flagged if pairs of samples from the same patient with > 5% discordance and from different patients with < 5% discordance
Major Contamination QC	% heterozygous sites at fingerprint SNPs < 55%; Average MAF at homozygous fingerprint SNPs < 2%
Criteria for calling test failure	If a sample presents with mean coverage across all exons < 50x and no mutations are detected due to the low overall coverage, the test is deemed "failed" for the sample.

*Defined as Hotspot SNVs in COSMICv68, mutation hotspots reported in TCGA, reported in Cheng *at. al.*(Nature Biotech, 2016) and indels in selected exons of established oncogenes.

**SNVs and Indels other than the ones defined as hotspot mutations above.

J. Standard/Guidance Document Referenced (if applicable):
Not applicable

K. Test Principle:
The MSK-IMPACT assay is a custom targeted sequencing platform, utilizing solution-phase exon capture and sequencing, to detect somatic alterations (point mutations, small insertions and deletions, and microsatellite instability) in tumor specimens. The MSK-IMPACT assay involves hybridization capture and deep sequencing of all protein-coding exons of 468 cancer-associated genes. The assay uses custom DNA probes corresponding to all exons and selected introns of oncogenes and tumor suppressor genes. Probes are synthesized by a secondary manufacturer and are biotinylated to enable sequence enrichment through capture by streptavidin-conjugated beads. Probes were designed to tile the entire length of each target sequence in an overlapping fashion, typically extending 20-50 base pairs beyond the boundaries of the target. In total, the probes target approximately 1.5Mb of the human genome.

Genomic DNA is extracted from tumor and patient-matched blood/normal tissue as a normal control when available. Sequence libraries are prepared through a series of enzymatic steps including shearing of double-stranded DNA, end repair, A-base addition, ligation of barcoded sequence adaptors, and low cycle PCR amplification. Multiple barcoded sequence libraries are pooled and captured using the custom-designed biotinylated probes. Captured DNA fragments are then sequenced on an Illumina HiSeq2500 as paired-end reads. Sequence reads are then aligned to the reference human genome. By comparing the identity of bases from the tumor DNA to the matched normal DNA and the reference human genome, somatic alterations are identified in the tumor.

L. Performance:

1. <u>Determination of pipeline thresholds:</u>

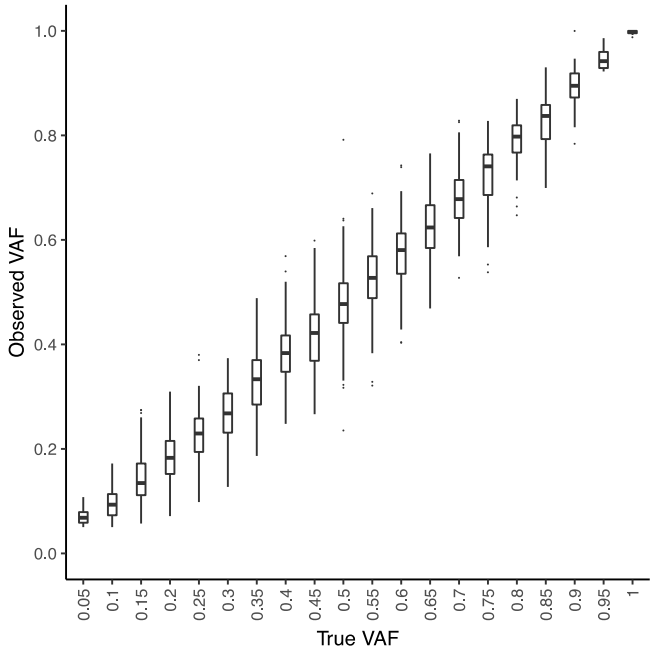
a) Requirements on exon coverage were established: A power analysis to compute the coverage or total number of reads needed to detect a mutation with true underlying mutation frequency 2% or greater, for varying levels of power (0.8 to 0.99), assuming a fixed alpha (Type I error rate) of 0.05 was conducted. Additionally, the 95% confidence interval ranges of observed mutation frequency as a function of coverage was also calculated. When the mutation is present at 10%, the 95% confidence interval with a coverage of 500X is expected to fall between 7.5% and 13%. When the overall coverage is 100X, the 95% CI for a mutation at 10% is estimated to fall between 5.0% and 17.6%.

To confirm these estimates, empirical data was obtained to measure the range of observed VF to expected VF using DNA from 10 normal FFPE samples from unrelated individuals which was mixed in equimolar parts so as to create a range of SNPs with expected frequencies as low as 5%. A total of 862 common SNPs were considered for this experiment.

A boxplot showing the observed mutation frequencies for the 862 common SNPs genotyped in the pooled normal sample binned by their true underlying mutation frequency is shown below. The results demonstrated that an observed VF range from 5.0% to 13.9% for a SNP with true underlying mutation frequency of 10% when the mean coverage of the sample was 480X. This range in values is roughly in line with what the theoretical statistical assessment for a coverage depth of 500X (7.5% to 13.0%). This data provided support for using a 5% as the lower limit for reporting mutations detected with true underlying frequency of 10%.

The boxplot in Figure 2 shows the correlation is 0.975, with a slope of 0.971 and intercept of -0.004. Consistent correlation is established as >0.9 as a QC metric for the whole pool analyzed.

Figure 2. Observed vs. Expected Variant Frequency



b) Requirements on sample coverage: Ten normal (diploid) FFPE samples were profiled in duplicate using the IMPACT assay (total = 20 replicates) to generate summary statistics across all targeted exons. The mean coverage across all targeted exons for the normal samples was 571X (SD = 373X). Summary statistics were also computed on coverage values per exon normalized by per-sample coverage. There were exons that presented with consistently low coverage values. None of the exons of the genes in the clinical validation are among those with consistently low coverage. It was determined the low coverage was due to sequence similarity with other loci, and high GC content. The exons were removed from the MSK-IMPACT assay. Of the remaining exons across all genes, 99.5% were sequenced to a depth of 100X or greater while 98.6% were sequenced to a depth of 250X or greater. This analysis of normal samples indicates that with a mean sample coverage of 571X, 98% of exons are sequenced with coverage greater than 306X, or with normalized coverage greater

than 0.54. (The ‘mean-normalized coverage’ is the coverage of the mutation divided by the mean coverage across all exons; it serves as a measure of how deeply the validation exon was sequenced relative to the overall coverage of the sample. A mean-normalized coverage below 1 indicates the exon coverage is below average; conversely if greater than 1, it indicates above average coverage.) The data are shown in Figures 3 and Figure 4

Figure 3. Distribution of mean coverage values for targeted exons. Dashed line indicates coverage at 100X.

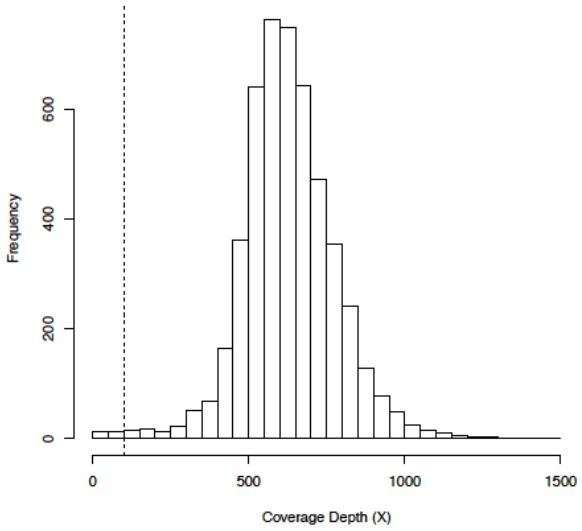
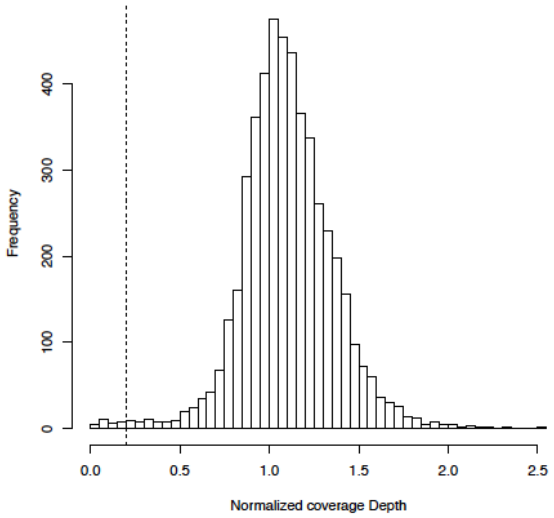


Figure 4. Distribution of mean coverage for targeted exons, normalized by per-sample coverage. Dashed line indicates 20% of mean sample coverage.



Based on the calculations, 98% of exons can be expected to be sequenced to coverage greater than 100X, when mean sample coverage is 185X (0.54* 185X = 100X). (A 100X minimum coverage threshold per exon is required based on the power

calculations, which showed 100X coverage was necessary to call mutations with true underlying mutation frequency 10% or greater, with 95% power at an alpha level of 0.05).

To be conservative, a threshold of 200X on mean sample coverage is used to determine if a sample is sequenced to sufficient depth for subsequent analysis. A sample is flagged as being at increased risk of false negatives if its mean coverage is below 200X.

To provide empirical data for these requirements, MSK utilizes the pool normal sample with known expected single nucleotide mutations (n = 2436) and the underlying mutation allele fractions (MAF). In silico downsampling analysis was conducted with a pool normal mix down to 45% where the sample coverage decreased from 452X to 203X. At this coverage level, 94% of the mutations with expected underlying VAF of 10% were called.

c) **Requirements on mutation coverage, allele depth and frequency for positive calls:** Permissive standard filters were used to intentionally generate false positives to identify suitable thresholds for parameters such as mutation coverage (DP), alternate allele depth (AD) and mutation frequency (VF) to optimize specificity. The following criteria allows optimal rejection of false positive SNVs (stratified by whether they are hotspots or not) and indel calls, while maintaining ability to detect true positive events with underlying frequency of 10% (5-17.6% observable). Potential strand-bias is also evaluated in the standard somatic mutation calling pipeline. An example of the number of false positive events detected pre and post filtering for coverage depth(DP), number of mutant reads (AD) and variant frequency (VF) is shown in Table 4.

Table 4. Sample error correction by DP/AD/VF filter

	Mutations –Cosmic database		**Mutations**	
Filter criteria	**DP ≥ 20X, AD ≥ 8,VF ≥ 2%**		**DP ≥ 20X, AD ≥ 10,VF ≥ 5%**	
	SNVs	Indels	SNVs	Indels
Pre-filter	1	24	342	40,793
Post-filter	0	0	0	8
Rejection Rate	1.00	1.00	1.00	0.999

2. **Pre-Analytical performance:**

Minimum DNA requirements were established by measuring assay performance based on different inputs from normal blood and FFPE tumor samples. DNA samples are normalized to yield 50 – 250 ng input and maximized to 55 ul prior to shearing. The normalization and DNA quantification are performed.

DNA extraction method was validated based on the invalid rates across multiple tumor types obtained from historical data. The data demonstrated that the DNA extraction has been optimized across tumor types to reasonably conclude that the analytical

performance presented is representative across FFPE tumor types. Table 5 shows the historical data for invalid rates from a retrospective chart review of >10,000 specimens tested with MSK-IMPACT. The range of invalid rates was 7.2% to 18.4%. The data shows that interference effects from different specimens are not significant across different tumor types supporting the performance of the pan-cancer specimen handling.

Table 5. Specimen Invalid Rates for 17 FFPE Tumor Types

			Pre-Run Invalids	Pre-Run Invalids	Post-Run Invalids	
Tumor Type	**Specimen Type**	**Number of Tests**	**Tumor Insufficient (Tumor % <20%)**	**DNA Insufficient (DNA yield <50ng)**	**Sequencing Failure (Coverage <50X)**	**Percent Invalids**
Non-Small Cell Lung Cancer	FFPE	1995	53	208	75	16.8
Breast Carcinoma	FFPE	1588	41	126	97	16.6
Colorectal Cancer	FFPE	1105	29	39	31	9.0
Prostate Cancer	FFPE	879	28	63	71	18.4
Glioma	FFPE	601	1	33	16	8.3
Pancreatic Cancer	FFPE	584	15	38	29	14.0
Soft Tissue Sarcoma	FFPE	479	3	21	13	7.7
Bladder Cancer	FFPE	480	12	20	25	9.8
Melanoma	FFPE	411	7	22	17	11.2
Renal Cell Carcinoma	FFPE	403	12	15	16	10.7
Hepatobiliary Cancer	FFPE	398	11	17	15	10.8
Esophagogastric Carcinoma	FFPE	374	5	12	16	8.8
Germ Cell Tumor	FFPE	332	9	13	30	8.1
Thyroid Cancer	FFPE	258	2	12	13	10.5
Ovarian Cancer	FFPE	244	4	8	8	8.2
Endometrial Cancer	FFPE	235	2	8	7	7.2
Head and Neck Carcinoma	FFPE	208	8	8	6	10.5
Cancer of Unknown Primary	FFPE	224	15	15	10	17.8

3. Analytical performance:

The hybridization-capture-based targeted re-sequencing assay is designed to detect point mutations [also referred to as single nucleotide variants (SNVs)] as well as small insertions/deletions (indels) < 30bp in length in the coding exons of 468 genes (Appendix 1b). A total of 6,357 exons are sequenced, 73 exons were excluded during assay development due to low sequence coverage and high GC content (Appendix 1c). A paired-sample analysis pipeline (tumor vs. matched normal) is used to identify somatic mutations in the targeted exons. MSK took a representative approach to validation of the SNVs and indels targeted in this panel, which is appropriate for variants of this type.[1]

a) Precision Studies: The objective of the precision studies was to assess between-run and within-run precision. Extracted DNA was run once per day for 3 days using different barcodes for inter-day assessment (n=3). For one run, a sample was run in triplicates for intra-day assessment, resulting in a total of 3+1+1=5 replicates. For each replicate tested, all observed mutations were reported and assessed for precision. Details of the study are described below.

i. Precision Panel: The precision of the MSK-IMPACT assay was assessed using 10 samples (9 FFPE specimens and one commercial cell line) to represent different tumor types, different mutation types, and the range of mutant allele frequencies. The panel included challenging specimens. The specimen panel was selected based on known mutations corresponding to "Cancer Mutations with Evidence of Clinical Significance" as well as the associated target tissue. The representative list of specimens is shown in Table 6.

Table 6: Summary of the Specimens and Allele Frequencies in the Precision Studies

Tissue type	Mutation type	Gene/ exon	cDNA change	Amino acid change	Mutation frequency
Glioblastoma	INS	EGFR exon20	C2290_2310dup TACGTGATGGCCAGC GTGGAC	p.Y764_D770dup	~ 5%
Cutaneous Melanoma	DNV	BRAF exon15	c.1798_1799delinsAA	V600K	~6.5%
Uterine Endometrial Cancer	SNV	KRAS exon2	C35G>C	G12A	~7%
Lung Adenocarcinoma	INS	ERBB2 exon 20	2310_2311ins GCATACGTGATG	E770_A771insAYVM	~15%
Lung Adenocarcinoma	SNV	EGFR exon 21	2573T>G	L858R	~20%

[1] For complex structural variations, such as genomic rearrangements (fusions) and copy number variations (CNVs), the expectation is that the representative approach should be demonstrated at the gene level.

Tissue type	Mutation type	Gene/ exon	cDNA change	Amino acid change	Mutation frequency
CRC	SNV	KRAS exon 2	C34G>T	G12C	~30%
Lung Adenocarcinoma	DEL (15bp)	EGFR exon 19	2236_2250delGAATTAAGAGAAGCA	E746_A750del	~30%
CRC	SNV	BRAF exon 15	c.1799T>A	V600E	~40%
GIST	DEL (6bp)	Kit exon 11	1667_1672delAGTGGA	Q556_K558del	~50%
FFPE Cell Line	DEL, SNV	Hotspot mutations in BRAF, EGFR, FLT3, GNA11, IDH1, KRAS, NRAS and PIK3CA genes			~2%-15%

ii. *Precision- Panel-Wide Reproducibility*: The precision analysis was performed for the known mutations (as listed in Table 6), and also performed for all additional mutations identified in each specimen in any of the test replicates. A total of 69 mutations in the clinical specimens and 13 mutations in the cell line were detected for a total of 82 mutations. In addition to SNV/MNVs, there were 9 deletions and 8 insertions.

The results showed that all mutations have 100% concordance in all replicates except for 4 mutations in the clinical specimens and 3 mutations in the commercial sample. In the clinical specimen discordance was observed for an SNV (pQ64K) and a frameshift mutation (pL54fs) in AR exon1, an insertion (pA445_P446insP) ARID1B exon1; and a frameshift mutation (pT319Kfs*24) in PTEN exon 8. The discordance on AR and ARID1B mutations were due to poor mapping quality in the highly repetitive regions.

The 3 mutations from the commercial control sample that were discordant were 2 SNVs and one deletion (IDH1 exon4 R132H; BRAF exon15 V600M; EGFR exon19 E746_A750del). These 3 mutations were believed to be discordant because they have low frequencies near 2%.

The coefficient of variation (%CV) for the mutation allele frequency was also calculated for all 5 replicates. Thirty-four (45) of the 69 mutations in the clinical specimens had %CV ≤10%, 17/69 were between 10 and 20% and 7/69 were >21%. All results are summarized in Table 7. Each specimen is separated by a dark gray line. Known mutation within each specimen are in **bold**. Discordant cases are denoted in light grey. All runs passed the quality metrics criteria.

Table 7. Panel-wide precision summary for all 5 replicates Abbreviations: NC (normalized coverage); MAF (Mutant allele frequency)

Gene Exon	Mutation (cDNA/Protein Changes)	NC range	MAF range	MAF mean	MAF median	MAF (SD)	MAF (%CV)	Positive /Total Calls	Positive Call Rate (two-sided 95% CI)
EGFR exon19	**c.2236_2250delGAATTAAGAGAAGCA 746_750del**	0.84-1	0.311-0.342	0.323	0.316	0.013	4.0%	5/5	100.0% (47.8%, 100.0%)
PTEN exon2	c.T83G I28S	0.62-0.73	0.502-0.569	0.543	0.544	0.027	5.0%	5/5	100.0% (47.8%, 100.0%)
TET2 exon3	c.C311G S104C	1.04-1.32	0.085-0.103	0.098	0.102	0.008	8.2%	5/5	100.0% (47.8%, 100.0%)
TP53 exon7	c.C742T R248W	0.97-1.22	0.648-0.664	0.66	0.663	0.007	1.1%	5/5	100.0% (47.8%, 100.0%)
BRAF exon15	**c.T1799A V600E**	1.26-1.44	0.415-0.454	0.431	0.425	0.015	3.5%	5/5	100.0% (47.8%, 100.0%)
BRCA2 exon14	c.A7388G N2463S	0.84-0.96	0.19-0.23	0.209	0.21	0.015	7.2%	5/5	100.0% (47.8%, 100.0%)
BRD4 exon19	c.G3922A A1308T	0.44-0.56	0.5-0.636	0.553	0.54	0.054	9.8%	5/5	100.0% (47.8%, 100.0%)
FBXW7 exon9	c.G1268T G423V	0.91-1.05	0.369-0.418	0.395	0.391	0.02	5.1%	5/5	100.0% (47.8%, 100.0%)
GRIN2A exon7	c.C1514A A505E	0.92-1.1	0.194-0.211	0.202	0.203	0.006	3.0%	5/5	100.0% (47.8%, 100.0%)
PTPRD exon12	c.G10A V4I	0.5-0.63	0.281-0.361	0.336	0.35	0.034	10.1%	5/5	100.0% (47.8%, 100.0%)
RUNX1 exon9	c.806-1G>A NA	1.01-1.23	0.185-0.21	0.202	0.207	0.01	5.0%	5/5	100.0% (47.8%, 100.0%)
SPEN exon12	c.C10445T P3482L	0.94-1.03	0.189-0.235	0.208	0.2	0.018	8.7%	5/5	100.0% (47.8%, 100.0%)
SYK exon13	c.C1768T R590W	1.13-1.22	0.233-0.292	0.273	0.279	0.023	8.4%	5/5	100.0% (47.8%, 100.0%)
TP53 exon6	c.G610T E204X	0.9-1.01	0.525-0.56	0.547	0.551	0.013	2.4%	5/5	100.0% (47.8%, 100.0%)
APC exon16	c.G3856T E1286X	0.8-1.05	0.326-0.39	0.351	0.349	0.026	7.4%	5/5	100.0% (47.8%, 100.0%)

APC exon7	c.C646T R216X	0.87-1.06	0.148-0.185	0.162	0.16	0.015	9.3%	5/5	100.0% (47.8%, 100.0%)
CREBBP exon29	c.G4837A V1613M	1-1.19	0.159-0.196	0.178	0.18	0.017	9.6%	5/5	100.0% (47.8%, 100.0%)
KRAS exon2	**c.G34T G12C**	1.13-1.31	0.289-0.352	0.314	0.305	0.024	7.6%	5/5	100.0% (47.8%, 100.0%)
NOTCH1 exon34	c.7541dupC P2514fs	1.28-1.5	0.144-0.211	0.184	0.189	0.025	13.6%	5/5	100.0% (47.8%, 100.0%)
SMAD4 exon11	c.C1333T R445X	0.76-0.95	0.206-0.238	0.223	0.229	0.014	6.3%	5/5	100.0% (47.8%, 100.0%)
ALOX12B exon11	c.G1406A R469Q	1.03-1.31	0.333-0.377	0.355	0.356	0.016	4.5%	5/5	100.0% (47.8%, 100.0%)
ARID1B exon1	c.1333_1334insCGC A445_P446insP	0.2-0.2	0.2-0.2	0.2	0.2	NA	NA	1/5	20.0% (0.5%, 71.6%)
CDK8 exon10	c.C1014A D338E	0.59-0.7	0.256-0.336	0.303	0.315	0.032	10.6%	5/5	100.0% (47.8%, 100.0%)
DNMT1 exon36	c.T4380G H1460Q	1.18-1.51	0.51-0.558	0.534	0.53	0.017	3.2%	5/5	100.0% (47.8%, 100.0%)
ERBB2 exon2	c.G140A R47H	1.16-1.59	0.596-0.712	0.656	0.666	0.045	6.9%	5/5	100.0% (47.8%, 100.0%)
ERBB2 exon20	**c.2310_2311insGCATACGTGATG E770_A771insAYVM**	1.02-1.38	0.142-0.199	0.173	0.171	0.023	13.3%	5/5	100.0% (47.8%, 100.0%)
ERCC2 exon21	c.C1904T A635V	1.19-1.47	0.363-0.466	0.409	0.423	0.045	11.0%	5/5	100.0% (47.8%, 100.0%)
IRS1 exon1	c.C3639A S1213R	0.42-0.49	0.384-0.494	0.449	0.455	0.04	8.9%	5/5	100.0% (47.8%, 100.0%)
MED12 exon37	c.5258_5282delCTCCTACCCTGCTAGAGCCTGAGAA A1753fs	1.08-1.36	0.141-0.187	0.164	0.17	0.019	11.6%	5/5	100.0% (47.8%, 100.0%)
MED12 exon43	c.6339_6340insCAGCAACACCAG Q2113_Q2114insQQHQ	0.96-1.43	0.37-0.422	0.4	0.399	0.021	5.3%	5/5	100.0% (47.8%, 100.0%)
NF1 exon51	c.C7595T A2532V	0.92-1.04	0.627-0.68	0.664	0.676	0.022	3.3%	5/5	100.0% (47.8%, 100.0%)

NTRK1 exon1	c.G53A G18E	0.28-0.55	0.6-0.668	0.631	0.63	0.027	4.3%	5/5	100.0% (47.8%, 100.0%)
PDGFRB exon7	c.G946A V316M	0.73-1.14	0.615-0.681	0.646	0.642	0.026	4.0%	5/5	100.0% (47.8%, 100.0%)
PIK3CB exon15	c.A2150G N717S	0.67-0.85	0.273-0.317	0.299	0.308	0.018	6.0%	5/5	100.0% (47.8%, 100.0%)
PTPRS exon32	c.C4822T R1608W	0.79-1.06	0.526-0.562	0.543	0.542	0.013	2.4%	5/5	100.0% (47.8%, 100.0%)
RB1 exon2	c.138-2A>G splicing mutation	0.51-0.75	0.231-0.345	0.291	0.284	0.047	16.2%	5/5	100.0% (47.8%, 100.0%)
TET1 exon4	c.G3476A R1159Q	0.86-1.34	0.499-0.606	0.533	0.522	0.044	8.3%	5/5	100.0% (47.8%, 100.0%)
TP53 exon5	c.G524A R175H	0.75-1.11	0.247-0.344	0.314	0.337	0.04	12.7%	5/5	100.0% (47.8%, 100.0%)
EGFR exon21	**c.T2573G L858R**	1.4-1.44	0.172-0.225	0.199	0.203	0.02	10.1%	5/5	100.0% (47.8%, 100.0%)
HNF1A exon4	c.C934T L312F	0.35-0.54	0.033-0.077	0.057	0.059	0.016	28.1%	5/5	100.0% (47.8%, 100.0%)
MLL3 exon42	c.G9671A R3224H	1.27-1.4	0.089-0.118	0.104	0.105	0.011	10.6%	5/5	100.0% (47.8%, 100.0%)
NTRK3 exon14	c.1401delC P467fs	0.49-0.54	0.062-0.086	0.074	0.077	0.01	13.5%	5/5	100.0% (47.8%, 100.0%)
TP53 exon10	c.A1051T K351X	0.74-0.84	0.075-0.116	0.103	0.108	0.016	15.5%	5/5	100.0% (47.8%, 100.0%)
AR exon1	c.161_171delTGC TGCTGCTG L54fs	0.34-0.39	0.079-0.097	0.088	0.087	0.009	10.2%	3/5	60.0% (14.7%, 94.7.0%)
AR exon1	c.C190A Q64K	0.25-0.29	0.134-0.135	0.134	0.134	0.001	0.7%	2/5	40.0% (5.3%, 85.3%)
KIT exon11	c.1667_1672delA GTGGA 556_558del	1.65-1.86	0.554-0.595	0.569	0.566	0.016	2.8%	5/5	100.0% (47.8%, 100.0%)
KIT exon17	c.T2467G Y823D	1.28-1.49	0.619-0.658	0.646	0.655	0.016	2.5%	5/5	100.0% (47.8%, 100.0%)
RPS6KB2 exon10	c.G840T K280N	0.93-1.19	0.435-0.473	0.462	0.468	0.015	3.2%	5/5	100.0% (47.8%, 100.0%)

CARD11 exon25	c.3382T>A p.V1128I	1.34-1.58	0.276-0.293	0.284	0.278	0.009	3.2%	5/5	100.0% (47.8%, 100.0%)
EGFR exon20	**c.2290_2310dupTACGTGATGGCCAGCGTGGAC p.Y764_D770dup**	14.36-15.46	0.05-0.06	0.055	0.055	0.004	7.3%	5/5	100.0% (47.8%, 100.0%)
EGFR exon7	c.874G>T p.V292L	21.51-21.82	0.934-0.939	0.937	0.939	0.002	0.2%	5/5	100.0% (47.8%, 100.0%)
NOTCH3 exon22	c.3646G>A p.A1216T	1.35-1.52	0.247-0.318	0.281	0.281	0.026	9.3%	5/5	100.0% (47.8%, 100.0%)
PTEN exon5	c.395G>C p.G132A	0.6-0.72	0.605-0.667	0.635	0.631	0.029	4.6%	5/5	100.0% (47.8%, 100.0%)
RUNX1 exon8	c.899C>T p.T300M	0.81-0.92	0.244-0.274	0.26	0.266	0.015	5.8%	5/5	100.0% (47.8%, 100.0%)
STAG2 exon17	c.1544_1547delATAG p.D515Gfs*6	0.19-0.27	0.677-0.842	0.753	0.741	0.067	8.9%	5/5	100.0% (47.8%, 100.0%)
TERT Promoter	g.1295228C>T non-coding	0.55-0.67	0.388-0.467	0.421	0.417	0.033	7.8%	5/5	100.0% (47.8%, 100.0%)
AKT3 exon2	c.134T>G p.V45G	1.14-1.36	0.05-0.078	0.066	0.067	0.012	18.2%	5/5	100.0% (47.8%, 100.0%)
BRAF exon15	**c.1798_1799delinsAA p.V600K**	1.04-1.32	0.065-0.095	0.072	0.067	0.013	18.1%	5/5	100.0% (47.8%, 100.0%)
KIT exon11	c.1735_1737delGAT p.D579del	1.08-1.22	0.051-0.056	0.053	0.054	0.002	3.8%	5/5	100.0% (47.8%, 100.0%)
CTCF exon3	c.610dupA p.T204Nfs*26	0.68-0.86	0.041-0.072	0.057	0.061	0.014	24.6%	5/5	100.0% (47.8%, 100.0%)
EGFR exon20	c.2317_2319dupCAC p.H773dup	1.15-1.19	0.067-0.093	0.078	0.079	0.011	14.1%	5/5	100.0% (47.8%, 100.0%)
KDM5C exon23	c.3755G>A p.R1252H	0.88-1.17	0.064-0.13	0.088	0.084	0.026	29.5%	5/5	100.0% (47.8%, 100.0%)
KRAS exon2	c.35G>C p.G12A	0.78-0.94	0.044-0.106	0.076	0.074	0.023	30.3%	5/5	100.0% (47.8%, 100.0%)
PIK3R1 exon13	c.1672_1683delGAAATTGACAAA p.E558_K561del	0.43-0.52	0.067-0.116	0.085	0.081	0.019	22.4%	5/5	100.0% (47.8%, 100.0%)

PIK3R1 exon9	c.1023dupA p.E342Rfs*4	0.41-0.58	0.056-0.102	0.083	0.086	0.017	20.5%	5/5	100.0% (47.8%, 100.0%)
PIK3R1 exon9	c.1024G>T p.E342*	0.42-0.59	0.064-0.108	0.093	0.095	0.017	18.3%	5/5	100.0% (47.8%, 100.0%)
PTEN exon6	c.493-1G>A p.X165_splice	0.53-0.64	0.173-0.208	0.192	0.187	0.015	7.8%	5/5	100.0% (47.8%, 100.0%)
PTEN exon8	c.956_959delCTT T p.T319Kfs*24	0.28-0.48	0.006-0.079	0.049	0.052	0.029	59.2%	3/5	60.0% (14.7%, 94.7.0%)
SOX17 exon1	c.287C>G p.A96G	1.16-1.51	0.061-0.074	0.069	0.069	0.005	7.2%	5/5	100.0% (47.8%, 100.0%)
BRAF exon15	**c.1798G>A V600M**	0.97-1.06	0.016-0.041	0.027	0.027	0.01	37.0%	3/5	60.0% (14.7%, 94.7.0%)
BRAF exon15	**c.1799T>A V600E**	0.97-1.06	0.051-0.08	0.064	0.067	0.012	18.8%	5/5	100.0% (47.8%, 100.0%)
EGFR exon18	**c.2155G>A G719S**	1.23-1.33	0.125-0.179	0.158	0.164	0.022	13.9%	5/5	100.0% (47.8%, 100.0%)
EGFR exon19	**c.2235_2249delG GAATTAAGAG AAGC E746_A750del**	1.01-1.19	0.009-0.043	0.023	0.019	0.013	56.5%	2/5	40.0% (5.3%, 85.3%)
FLT3 exon20	**c.2503G>T D835Y**	0.97-1.02	0.037-0.059	0.045	0.043	0.008	17.8%	5/5	100.0% (47.8%, 100.0%)
GNA11 exon5	**c.626A>T Q209L**	1.41-1.48	0.036-0.054	0.046	0.044	0.008	17.4%	5/5	100.0% (47.8%, 100.0%)
IDH1 exon4	**c.395G>A R132H**	0.5-0.53	0.038-0.049	0.035	0.044	0.020	57.1%	4/5	80.0% (28.4%, 99.5%)
KRAS exon2	**c.34G>A G12S**	0.9-1.03	0.026-0.057	0.041	0.039	0.011	26.8%	5/5	100.0% (47.8%, 100.0%)
KRAS exon2	**c.38G>A G13D**	0.91-1.06	0.217-0.249	0.231	0.229	0.012	5.2%	5/5	100.0% (47.8%, 100.0%)
KRAS exon4	**c.436G>A A146T**	0.82-0.88	0.031-0.055	0.042	0.044	0.009	21.4%	5/5	100.0% (47.8%, 100.0%)
NRAS exon3	**c.183A>T Q61H**	1.01-1.14	0.039-0.065	0.051	0.051	0.01	19.6%	5/5	100.0% (47.8%, 100.0%)
PIK3CA exon10	**c.1624G>A E542K**	0.67-0.87	0.038-0.047	0.042	0.042	0.004	9.5%	5/5	100.0% (47.8%, 100.0%)
PIK3CA exon21	**c.3140A>G H1047R**	0.62-0.72	0.222-0.331	0.276	0.258	0.05	18.1%	5/5	100.0% (47.8%, 100.0%)

iii. Per Specimen Precision: Results of the precision studies were combined and precision across all reportable genes was determined for each specimen. The positive call rate based on the total number of mutations along with the 2-sides 95% confidence interval were calculated. Results are summarized in Table 8.

Table 8. Precision per specimen across all reportable mutations (N – 5 replicates)

Specimen	Total No unique mutations detected across all 5 replicates*	*Positive call rate per mutation	Positive call rate* (two-sided 95% CI)	Negative call rate (two-sided 95% CI)
M15-22924	5	5/5 for all	25/25 100.0% (86.3%, 100.0%)	-
M15-3038	3	5/5 for all	15/15 100.0% (78.2%, 100.0%)	-
M16-19000	10	5/5 for 9 4/5 for 1	49/50 98.0% (89.4%, 99.9%)	-
M1688-5C	18	5/5 for 17 1/5 for 1	86/90 95.6% (89.0%, 98.8%)	4/5 80.0% (28.4%, 99.5%)
M-1698-A9	5	5/5 for all	25/25 100.0% (86.3%, 100.0%)	-
M-1654-CA	6	5/5 for all	30/30 100.0% (88.4%, 100.0%)	-
M-1612-28	4	5/5 for all	20/20 100.0% (83.2%, 100.0%)	-
M1648-D5	10	5/5 for all	50/50 100.0% (92.9%, 100.0%)	-
M-1707-12	5	5/5 for 3 3/5 for 1; 2/5 for 1	20/25 80.0% (59.3%, 93.2%)	3/5 60.0% (14.7%, 94.7%)
Commercial sample	13	5/5 for 10; 4/5 for 1; 3/5 for 1; 2/5 for 1	59/65 90.8% (81.0%, 96.5%)	3/5 60.0% (14.7%, 94.7%)

*Positive call rate is calculated based on variants with majority call detected as positive
#Negative call rate is calculated based on variants detected at least once, but with majority call as negative. For all other locations, the negative call rates are 100%.

The precision study was also evaluated for the intra-assay repeatability (within-run). All results were concordant except for ARID1B exon 2 insertion from clinical specimen M-1688, and BRAF V600M point mutation in the commercial control sample as described previously. Additionally, performance with respect to quality metrics (i.e., total depth of coverage and mutant allele coverage) in all replicates was also summarized and shown to meet the pre-specified acceptance criteria (data not shown).

iv. *Precision - Well-characterized reference material:* The precision of MSK-IMPACT was assessed through repeated measurements of a well characterized reference standard (HapMap cell line NA20810). To determine sequencing error rates for the reference sample, DNA extracted from the HapMap cell line was included in each run tested in the accuracy study. The study investigated whether the SNPs in the targeted exons were detected at their expected frequencies. Reference genotypes for 11,767 SNPs in the targeted exons using a whole genome sequencing BAM file for NA20810, were obtained from the 1000 Genomes database. A total of 11,443 SNPs (97.2%) were homozygous for the major allele (relative to the hg19 reference genome), 212 SNPs (1.8%) were heterozygous and 112 SNPs (0.95%) were homozygous for the minor allele. The strong bias towards alleles matching the reference genome was expected, given that these SNPs occur in coding exons and there is likely strong selective pressure against deviations from the reference sequence. NA20810 was profiled with the assay multiple times across different runs, for a total of 23 replicates. Zygosity results were 100% concordant and high levels of concordance – specifically, the difference between the expected and mean observed mutation frequencies was very small (absolute difference = 0.09%±0.45%). The data provide additional supplemental evidence of the reproducibility of the assay.

v. *Precision for Microsatellite Instability (MSI):* Precision of the MSI calling by MSIsensor was demonstrated with a total of 12 specimens: 6 MSI-H specimens (at three MSI-score levels, 3 replicates per sample) and 6 MSS specimens. Each DNA extracted sample was tested with 3 inter- and 3 intra-run replicates. Multiple barcodes were included. All samples had 100% agreement between calls**.** The total number of unstable loci relative to the total number of sites surveyed along with the mean, median and standard deviation (SD) and coefficient of variance (%CV) was also presented for each specimen and score. The results supported the precision of the MSIsensor scores greater than 0.5 Results are shown in Table 9.

Table 9. Precision of the MSIsensor Score Using 12 Specimens

N	Total Sites_ range	Unstable Loci_range	Mean	Median	SD	%CV	Positive Call Rate (two-sided 95% CI)
5	1227-1458	518-650	43.00	43.00	1.22	2.8%	100%(47.8%, 100.0%)
5	1158-1477	483-646	43.00	43.00	0.71	1.7%	100%(47.8%, 100.0%)
5	1187-1429	500-613	42.00	42.00	0.71	1.7%	100%(47.8%, 100.0%)
5	1287-1400	303-359	24.80	25.00	0.84	3.4%	100%(47.8%, 100.0%)
5	1251-1303	240-318	23.40	24.00	2.51	10.7%	100%(47.8%, 100.0%)
5	1154-1379	153-175	12.60	12.00	0.89	7.1%	100%(47.8%, 100.0%)

N	Total Sites_ range	Unstable Loci_range	Mean	Median	SD	%CV	Positive Call Rate (two-sided 95% CI)
5	1321-1545	46-58	3.60	4.00	0.55	15.3%	100%(47.8%, 100.0%)
5	1535-1604	44-64	3.40	3.00	0.55	16.2%	100%(47.8%, 100.0%)
5	1411-1612	28-38	2.20	2.00	0.45	20.5%	100%(47.8%, 100.0%)
5	1438-1528	6-9	0.48	0.50	0.08	16.7%	100%(47.8%, 100.0%)
5	1315-1487	0-2	0.02	0.00	0.04	223.6%	100%(47.8%, 100.0%)
5	1312-1532	0-1	0.01	0.00	0.03	223.6%	100%(47.8%, 100.0%)

b) Analytical Sensitivity – Limit of Detection (LoD): The LoD of the IMPACT assay is defined as the mutant allele fraction at which 95% of replicates across all replicates for a variant type are reliably detected. Studies were conducted to demonstrate a putative LoD for each variant type. In the first part, a dilution series was conducted to identify the lowest reliable mutant fraction. In part 2, the putative LoD was confirmed with multiple replicates.

i. Part 1: Dilution Series: The mean normalized coverage for all exons was determined for 10 normal FFPE specimens and the LoD was assessed with samples containing mutations in 5 validation exons (defined as representative exons harboring cancer mutations with evidence of clinical significance assessed in the accuracy study) with the lowest and highest coverage.

- The 5 validation exons with lowest coverage correspond to 3 exons harboring SNVs, (ERBB2 exon 20 (V777L), PDGFRA exon 18 (D842V), PIK3CA exon 10 (E545K), and 2 exons harboring indels (EGFR exon 19 and KIT exon 9).
- The 5 validation exons with highest coverage correspond to 3 exons harboring SNVs (BRAF exon 15 (V600E), KRAS exon 2 (G12D) and PIK3CA exon 2 (R88Q) and 2 exons harboring indels (KIT exon 11 and EGFR exon 20).

Five to eight serial dilutions were prepared using patient samples positive for the mutations listed above, where tumor samples were either diluted with their respective matched FFPE normal sample (when available) or a previously sequenced, unmatched normal FFPE sample. One replicate at each dilution was tested and the ability to detect the mutation of interest was measured. All results were called at the lowest dilution except for PIK3CA which was called wild-type at the lowest dilution. Results are shown in Tables 10A-J.

Table 10A. Limit of Detection –Part 1

SNV BRAF Exon 15 (Sample M-1648-D5-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.1799T>A	V600E	1018	410	0.4	Called
1:2			1044	319	0.31	Called
1:4			888	173	0.19	Called
1:8			999	91	0.09	Called
1:16			783	26	0.03	Called
1:32			845	20	0.02	Called

Table 10B

SNV KRAS Exon 2 (sample M-1807-ED-T)						
Dilution	**cDNA change**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.35G>A	G12D	907	405	0.45	Called
1:2			820	298	0.36	Called
1:4			400	97	0.24	Called
1:8			660	121	0.18	Called
1:16			665	59	0.09	Called
1:32			632	41	0.06	Called

Table 10C

SNV PIK3CA Exon 2 (Sample M-1729-E1-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.263G>A	R88Q	2029	629	0.31	Called
1:2			1008	211	0.21	Called
1:4			1140	145	0.13	Called
1:8			997	62	0.06	Called
1:16						WT

Table 10D

Kit exon 11 Deletion (Sample M-1621-AC-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.1667_1681delAGTGGAAGGTTGTTG	556_561del	2503	922	0.37	Called
1:2			1986	688	0.35	Called
1:4			1513	430	0.28	Called
1:8			1049	250	0.24	Called
1:16			792	138	0.17	Called
1:32			761	66	0.09	Called
1:64			618	37	0.06	Called
1:125			736	18	0.02	Called

Table 10E

EGFR exon 20 Insertion (sample M-1674-10-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.2308_2309insACT	D770_N771insY	1484	400	0.27	Called
1:2			777	166	0.21	Called
1:4			566	105	0.19	Called
1:8			595	55	0.09	Called
1:16			581	33	0.06	Called
1:32			608	21	0.03	Called

Table 10F

SNV ERBB2 exon 20 (sample M-1801-98-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.2525A>T	D842V	1471	408	0.28	Called
1:2			1482	240	0.16	Called
1:4			864	73	0.08	Called
1:8			903	38	0.04	Called
1:16			873	24	0.03	Called

Table 10G

SNV PDGFRα Exon 18 (sample M-1670-A6-T						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.1633G>A	E545K	448	236	0.53	Called
1:2			636	142	0.22	Called
1:4			962	95	0.1	Called
1:8			647	45	0.07	Called
1:16			707	16	0.02	Called

Table 10H

SNV PK3CA exon 10 (sample M-1434-A5-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.1633G>A	E545K	448	236	0.53	Called
1:2			636	142	0.22	Called
1:4			962	95	0.1	Called
1:8			647	45	0.07	Called
1:16			707	16	0.02	Called

Table 10I

EGFR exon 19 deletion (sample M-1809-C4-T)						
Dilution	**cDNAchange**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.2236_2250delGAATTAAGAGAAGCA	746_750del	1278	790	0.62	Called
1:2			1137	484	0.43	Called
1:4			792	207	0.26	Called
1:8			666	94	0.14	Called
1:16			622	49	0.08	Called
1:32			499	17	0.03	Called

Table 10J

Kit Exon 9 insertion (sample M-1754-DB-T)						
Dilution	**cDNA change**	**AA Change**	**DP**	**AD**	**VF**	**Result**
Neat	c.1502_1503insTGCCTA	S501_A502insAY	517	314	0.61	Called
1:2			512	187	0.37	Called
1:4			641	89	0.14	Called
1:8			486	27	0.06	Called
1:16			447	17	0.04	Called
1:32			521	14	0.03	Called

ii. *Part 2: Confirmation of the LoD.* A total of 5 replicates were tested for each of the 3 deletions, 4 insertions and 6 SNVs at 5% minor allele frequency. All variants have 100% positive call rates except for one replicate for a deletion on PTEN exon 6. This replicate also failed the mutation read depth and was below the estimated LoD of 5%. The results are shown in Table 11.

Table 11. Limit of Detection– Part 2

Type	Mutation	GeneExon	Range DP	Range AD	Range MAF	Range NormDP	Positive Call Rate
DEL	In_Frame_Del c.1735_1737delGAT p.D579del	KIT exon11	509-693	26-38	0.051-0.056	1.08-1.22	100.0%
DEL	Frame_Shift_Del c.956_959delCTTT p.T319Kfs*24	PTEN exon8	197-242	7-19	0.036-0.079	0.31-0.48	80.0%
DEL	In_Frame_Del c.1672_1683delGAAATTGACAAA p.E558_K561del	PIK3R1 exon13	216-313	18-36	0.067-0.116	0.43-0.52	100.0%
INS	In_Frame_Ins c.2317_2319dupCAC p.H773dup	EGFR exon20	587-749	46-65	0.067-0.093	1.15-1.19	100.0%
INS	Frame_Shift_Ins c.1023dupA p.E342Rfs*4	PIK3R1 exon9	236-345	15-32	0.056-0.102	0.41-0.58	100.0%
INS	Frame_Shift_Ins c.610dupA p.T204Nfs*26	CTCF exon3	344-540	14-36	0.041-0.072	0.68-0.86	100.0%
INS	In_Frame_Ins c.2290_2310dupTACGTGATGGCCAGCGTGGAC p.Y764_D770dup	EGFR exon20	8601-9836	441-572	0.05-0.06	14.36-15.46	100.0%
SNV	Missense_Mutation c.134T>G p.V45G	AKT3 exon2	535-813	28-63	0.05-0.078	1.14-1.36	100.0%
SNV	Missense_Mutation c.1798_1799delinsAA p.V600K	BRAF exon15	489-747	33-71	0.065-0.095	1.04-1.32	100.0%
SNV	Missense_Mutation c.287C>G p.A96G	SOX17 exon1	672-805	45-59	0.061-0.074	1.16-1.51	100.0%
SNV	Missense_Mutation c.35G>C p.G12A	KRAS exon2	445-571	20-55	0.044-0.106	0.78-0.94	100.0%
SNV	Missense_Mutation c.3755G>A p.R1252H	KDM5C exon23	475-733	40-68	0.064-0.13	0.88-1.17	100.0%
SNV	Nonsense_Mutation c.1024G>T p.E342*	PIK3R1 exon9	242-355	18-37	0.064-0.108	0.42-0.59	100.0%

iii. *Microsatellite instability (MSI):* The minimum tumor proportion required to support the MSIsensor score robustness was assessed using CRC specimens. Five (5) replicates were run using multiple barcodes and runs. The data showed that qualitatively, the assay and score are reproducible to 8% tumor proportion, though a decreasing trend in the quantitative score was observed. Therefore, the minimum tumor proportion required for the assay was established as 25% with an average coverage of 200X. Separately, regardless of the tumor proportion, data showed that the score is robust across the MSIsensor score range (refer to Table 9 above and Table 12).

Table 12. Replicate Testing of the MSI Sensor Score at 8% Tumor Purity

Tumor Purity	Coverage	# Total site	# Unstable loci	MSIsensor Score (%)
Diluted to 8%	517	1420	182	13
Diluted to 8%	562	1389	175	13
Diluted to 8%	555	1352	185	14
Diluted to 8%	502	1361	135	10
Diluted to 8%	378	1273	152	12

iv. *DNA-Input:* The validated DNA concentration is the amount at which the average read depth over the exon regions was maintained at the criteria established (e.g., ≥20 reads per base), and have 100% positive mutation call rate. The optimized and recommended DNA concentration for the assay is 250ng. The DNA input range 50-250ng. was assessed for accuracy and sequencing failures as a function of the input DNA concentration. The results show that assay performance in terms of sequencing failures is a function of genomic DNA input values as shown in Table 13.

Table 13. Sequencing Failures Relative to DNA Input

DNA Input	Success	Sequencing Failure
250ng	97%	3%
201-249ng	87%	13%
151-200ng	87%	13%
101-150ng	81%	19%
50-100ng	78%	22%

c) Linearity/assay reportable range:
Not applicable

d) Traceability (controls, calibrators, or methods):
The MSK-IMPACT is not traceable to any known standard. Controls and quality metrics are described in the device description section.

e) Stability:

Reagent stability is based on manufacturer expiration dating, and supported by MSK verification. Stability of the reagents is monitored through the use of consistent controls.

f) Expected values:

The laboratory follows protocols for the use of controls consistent with CLIA regulation. The MSK-IMPACT does not use calibrators; however, the verification of mutant allele frequency is maintained by analysis of a pooled control with expected allele frequencies.

g) Analytical specificity:

High analytical specificity is maintained by paired tumor/matched normal sequencing, and was established during assay optimization.

Interference:

The MSK-IMPACT assay pre-analytic steps are designed to minimize interference. The invalid rates in the historical testing from >10,000 samples support that any interference from any challenging tissues is minimized.

h) Assay cut-off*:*

The MSK-IMPACT does not report mutations below 2% for known hotspot mutations and 5% for non-hotspot mutations.

i) Comparison studies:

i. Method comparison:

The MSK-IMPACT assay is designed to detect SNVs and small indels in 6284 exons from 468 genes. The accuracy of the MSK-IMPACT was assessed by comparison of the MSK-IMPACT result to the original results obtained with the validated orthogonal methods. Testing was conducted per protocol. A total of 267 unique mutations in 433 FFPE tumor specimens representing 48 exons in 20 genes were tested and are listed in Table 14 below.

Table 14. Mutations Represented in the Accuracy Summary Per Gene

Gene (n=20)	#Samples (n=433)	Exon (n=48)	Type	Mutations Assessed
AKT	10	exon3	SNV	E17K
ALK	3	exon23	SNV	F1174V/L;S1205F
	4	exon25	SNV	R1275Q;R1260T
BRAF	11	exon11	SNV	G466V/R;S467L;G469*
	19	exon15	SNV	D594G;V600*;K601I
EGFR	10	exon18	SNV	G719A/S; G724S
	12	exon19	**DEL**	745_750del; 746_748del; 746_750del; 747_753del; K754fs

Gene (n=20)	#Samples (n=433)	Exon (n=48)	Type	Mutations Assessed
	10	exon20	SNV	T790M
	16	exon20	**INS**	M766_A767insASV; V769_D770insDNP; D770_N771ins*;P772_H773ins*; H773_V774insY/H
	9	exon21	SNV	L858R
ERBB2	7	exon19	SNV	L755S;I767M;D769Y
	16	exon20	**INS**	E770_A771insAYVM; A771_Y772insYVMA;G776_G778ins*
	3	exon20	SNV	V777L;G776V
	7	exon8	SNV	S310F/Y; S305C
FGFR2	1	exon12	SNV	L528H
	1	exon7	SNV	S252W
	1	exon9	SNV	Y375C
FGFR3	2	exon18	SNV	P797L
	1	exon7	SNV	A261V;A265V
	5	exon9	SNV	F384L
	1	exon9	**INS**	G370_S371insH
GNA11	7	exon5	SNV	Q209L
GNAQ	5	exon5	SNV	Q209P/L
GNAS	5	exon8	SNV	R201C/H
HRAS	3	exon2	SNV	G10A; G13D/V
	5	exon3	SNV	A59V; Q61R/L/K
IDH1	8	exon4	SNV	R132G/C/H
IDH2	5	exon4	SNV	R172*;R140Q
	1	exon4	**DEL**	T146Lfs*15
KIT	9	exon11	**INS; DEL**	K550fs; 552_557del; 556_558del; 556_561del; 558_565del; 559_566del; P573_T574insTQLPS
	9	exon11	SNV	V555L; W557G;V559D; D572G;L576P.
	6	exon13	SNV	V654A; K642E
	5	exon17	SNV	D816H; D820E; N822K
	10	exon9	**INS**	S501_A502insAY; A502_Y503dup
KRAS	16	exon2	SNV	G12*; G13D
	13	exon3	SNV	Q61*
	10	exon4	SNV	K117N;G138E;A146*
MET	13	exon14	SNV	D1010*; Exon14 skipping
	19	exon14	**DEL**	Exon14 skipping; Other splicing defects
NRAS	4	exon2	SNV	G13*
	12	exon 3	SNV	Q61*
PDGFRA	12	exon18	SNV	D842V/I
	1	exon12	SNV	V561D
PIK3CA	4	exon10	SNV	E545A/K; E542K
	2	exon21	SNV	H1047R/Y
	1	exon21	**INS**	X1069delinsFL
	8	exon2	SNV/MNV	F83L;R88Q;R93Q;K111E/N
	2	exon2	**DEL**	E110del; 112_113del
	9	exon5	SNV	V344M;N345I/K
	9	exon8	SNV	E418K;C420R;P449R;E453K/Q
	1	exon8	**DEL**	E453_D454del
TP53	9	exon4	SNV/MNV	W53X;W91X;Q100X;G105V/C; S106R; F113C
	6	exon4	DEL	L35fs;P67fs;A84fs;109_109del;G108fs;

Gene (n=20)	#Samples (n=433)	Exon (n=48)	Type	Mutations Assessed
				R110fs
	3	exon4	INS	V73fs;L114fs;C124fs
	6	exon5	SNV	K132Q;W146X; Y163C; R175H; R158H
	3	exon5	INS	P153fs; M160_A161insRA; Q167_M170dup
	9	exon5	DEL	K132fs;A138fs;P152fs; R156fs; V157_R158del; K164fs; H178fs;D184fs
	2	exon6	SNV	R213L/X
	8	exon6	**DEL**	G187fs; L188fs; P191_Q192del; R196_L201del; D207fs; R209fs; F212fs
	10	exon7	SNV/MNV	Y234C; Y236C; M237I; R248G/Q; R249S; T256P
	3	exon7	**INS**	S241dup; R249fs; T253dup
	6	exon7	**DEL**	S241fs; M243X; G244fs; M246X; I255del; L257fs
	4	exon8	SNV/MNV	V272K; C275X; R282W; T284K
	4	exon8	**INS**	C275fs; N288fs; G302fs
	5	exon8	**DEL**	N263_N268del; N263fs; R267fs; P278fs; P301fs
	6	exon10	SNV	R337L; R342X; R337C
	1	exon10	**INS**	L344fs

Of the 433 specimens, 418 met the criteria of ≥200X coverage, 15 samples (3.5%) failed to achieve average coverage above 200X. The known mutation associated with each sample was successfully detected in 432 out of 433 cases (99.8% with two-sided 95% CI of (98.7%, 100.0%)). One discordant case was observed in sample M-1994-BC-T, which was used for the validation of insertions in EGFR exon 20. The known mutation for this sample was a 12bp duplication which began in the intron 5' of EGFR exon 20, potentially creating an alternative splice site acceptor for the exon. This duplication event was detected by the indel calling pipeline but was incorrectly filtered out because of the calling algorithm. (The filtering algorithm was modified to improve the detection accuracy for such mutations.)

The MSK-IMPACT accuracy study included 159 unique SNV/MNVs from 20 genes (45 exons), 49 unique deletions from 6 genes (11 exons), and 39 unique insertions from 6 genes (10 exons). Performance was stratified by mutation type and gene for percent positive agreement (PPA) with 95% confidence interval (CI). Results are shown in Table 15A-C.[2]

[2] Performance may be overestimated because specimens were selected based on the availability of results by the orthogonal methods (i.e., the specimen set may lack challenging specimens).

Table 15A.Percent Positive Agreement for SNV/MNVs by Gene

Gene	Number of exons	Number of unique mutations	Number of samples	PPA (95% CI)
AKT1	1	1	10	**100.0% (69.2%, 100.0%)**
ALK	2	5	7	**100.0% (59.0%, 100.0%)**
BRAF	2	13	30	**100.0% (88.4%, 100.0%)**
EGFR	3	6	30	**100.0% (88.4%, 100.0%)**
ERBB2	3	12	17	**100.0% (80.5%, 100.0%)**
FGFR2	3	3	3	**100.0% (29.2%, 100.0%)**
FGFR3	3	3	8	**100.0% (63.1%, 100.0%)**
GNA11	1	1	7	**100.0% (59.0%, 100.0%)**
GNAQ	1	2	5	**100.0% (47.8%, 100.0%)**
GNAS	1	2	5	**100.0% (47.8%, 100.0%)**
HRAS	2	7	8	**100.0% (63.1%, 100.0%)**
IDH1	1	3	8	**100.0% (63.1%, 100.0%)**
IDH2	1	4	6	**100.0% (54.1%, 100.0%)**
KIT	3	13	20	**100.0% (83.2%, 100.0%)**
KRAS	3	15	39	**100.0% (91.0%, 100.0%)**
MET	1	9	13	**100.0% (75.3%, 100.0%)**
NRAS	2	6	16	**100.0% (79.4%, 100.0%)**
PDGFRA	2	3	13	**100.0% (75.3%, 100.0%)**
PIK3CA	4	19	32	**100.0% (89.1%, 100.0%)**
TP53	6	32	37	**100.0% (90.5%, 100.0%)**

Table 15B. Percent Positive Agreement for insertions by gene

Gene	Number of exons	Number of unique mutations	Number of samples	PPA (95% CI)
EGFR	1	12	16	**93.8% (69.8%, 100.0%)**
ERBB2	1	8	16	**100.0% (79.4%, 100.0%)**
FGFR3	1	1	1	**100.0% (2.5%, 100.0%)**
KIT	1	3	10	**100.0% (69.2%, 100.0%)**
PIK3CA	1	1	1	**100.0% (2.5%, 100.0%)**
TP53	5	14	14	**100.0% (76.8%, 100.0%**

Table 15C. Percent Positive Agreement for deletions by gene

Gene	Number of exons	No. unique mutations	Number of samples	PPA (95% CI)
EGFR	1	6	12	**100.0% (73.5%, 100.0%)**
IDH2	1	1	1	**100.0% (2.5%, 100.0%)**
KIT	1	7	9	**100.0% (66.4%, 100.0%)**
MET	1	18	19	**100.0% (82.4%, 100.0%)**
PIK3CA	2	3	3	**100.0% (29.2%, 100.0%)**
TP53	5	14	14	**100.0% (76.8%, 100.0%)**

ii. Supplemental Method Comparison Study for Wildtype Calls:
A supplemental study was conducted to assess accuracy for 33 "hotspots" within 10 genes. A total of 95 specimens were tested and the accuracy of

MSK-IMPACT results at all 33 positions was compared to results obtained with a single orthogonal method. Within the 95 specimens, there were 109 mutations across samples and 3026 wild-type calls. Variant-level concordance (PPA and NPA) was 100% for all results with two-sided 95% confidence intervals of (96.7%, 100.0%) for mutations (PPA) and (99.9%, 100.0%) for wild-type locations (NPA).

iii. *Method Comparison of the MSK-IMPACT MSIsensor:*
The somatic MSI status is inferred by interrogating all available genomic microsatellites covered by MSK-IMPACT within tumor samples against the matched normal DNA using the MSIsensor program as described in the Device Description section above. An MSIsensor score assigned to each tumor sample is used to distinguish MSS from MSI-H by MSIsensor.

The cutoff was first established using a training specimen dataset consisting of 138 colorectal cancer (CRC) and 40 endometrial carcinoma (EC) specimens with matched normal and having MSI status results from a validated MSI-PCR or MMR IHC test. MSIsensor scores ranged from 0 to 47.7 for CRC and 0 to 43.7 for EC. Based on concordance to either mismatch repair immunohistochemistry (MMR IHC) for MLH1, MSH2, MSH6 and PMS2 expression, or a commercially available PCR assay that detects 5 mononucleotide microsatellite loci including MR-21, BAT-25, MONO-27, NR-24 and BAT-26, a MSIsensor cut-off of 10 was established to delineate microsatellite stable (MMS) from high microsatellite instability (MSI-H).

A separate data set was obtained to validate this cut-off. A retrospective-prospective chart review of 135 CRC patients was conducted to identify cases that had both MSK-IMPACT MSI results and results by a validated IHC panel (MLH1, MSH2, MSH6 and PMS2). A total of 66 specimens had both sets of results. Of these, there were two discordant cases. The estimated positive predictive value (PPV) was 92.3% (12/13) with two-sided 95% confidence interval of 64.0%-99.8% and the estimated negative predictive value (NPV) was 98.1%. (52/53) with two-sided 95% confidence interval of 90.0%, 100.0%. The results are shown in Table 16 below.

Table 16. MSIsensor Results Compared to IHC MMR for CRC

CRC/EC Concordance with IHC		**MMR-D***	**MMR-P***	**Total**
MSI Sensor	**MSI-H ≥ 10**	12	1	13
	MSS < 10	1	52	53
Total		13	53	66
PPV = 92.3% (12/13) 95% CI 64.0%-99.8%				
NPV = 98.1%. (52/53) 95% CI 90.0%, 100.0%				

*MMR-D refers to deficient in mismatch repair proteins and MMR-P indicates not deficient

To evaluate the ability of the MSIsensor to determine MSI status in cancer types other than CRC or EC cancer types, 119 unique non-CRC and non-EC tumor-normal pair samples covering 25 tumor types were assessed for MSI by both MSIsensor and a validated MSI-PCR test. The results are shown in Table 17. Excluding the specimens without a MSI-PCR result from the total number of specimens analyzed, PPV is 46/49=93.9% (83.1%, 98.7%), and NPV is 58/60=96.7% (88.5%, 99.6%). When including all missing data in the analysis (i.e., consider all PCR unknown data as discordant results), the PPV=46/59=78.0% (65.3%, 87.7%), NPV= 58/60=96.7% (88.5%, 99.6%). (The MSIsensor MSI-H/MSS definition is based on genome wide analysis of over 1000 microsatellite markers and not based on the 5 or 7 MSI loci described in current clinical practice guidelines.)

Table 17. MSIsensor Results Compared to PCR 5 Loci MSI Panel for Other Cancer Types

Non CRC/EC concordance with MSI-PCR		**PCR Results**			
		MSI-H	**MSI-L/MSS**	**Unknown***	**Total**
MSIsensor	MSI-H (≥10)	46	3	10	59
	MSS (≥2 & <10)	2	58	0	60
	Total	48	61	10	119
Excluding missing specimens with 95%CI	PPV is 46/49=93.9% 95%CI (83.1%, 98.7%)				
	NPV is 58/60=96.7% (88.5%, 99.6%)				
Accounting for missing specimens with 95%CI	PPV=46/59=78.0% (65.3%, 87.7%)				
	NPV= 58/60=96.7% (88.5%, 99.6%)				

* In exploratory analysis, the 10 without PCR results were all MMR-D by IHC, consistent with the MSI-H by MSIsensor findings.

MSI Supplemental Information:
The mean, median and range of MSIsensor score was determined in a large cohort of 10,900 patients with 66 different types advanced solid tumor. The MSIsensor scores ranged from 0 to 48.5, mean 1.2, median 0.4. The prevalence of MSI-H by MSIsensor was also determined, and the findings are consistent with the MSI-H prevalence as described in the literature (data not shown).

3. **<u>Clinical Performance</u>:**

MSK-IMPACT assay is a molecular profiling platform using next generation sequencing to detect somatic alterations (point mutations and small insertions and deletions and microsatellite instability) in tumor specimens using a 468 gene panel. The genes in the panel were selected for their role in cancer pathogenesis and tumor

suppression, or for clinical or mechanistic information of relevance in the management of cancer patients. The assay reports mutations under two categories: "Cancer mutations with evidence of clinical significance" and "Cancer mutations with potential clinical significance" consistent with the intended use clinical settings. Mutations with evidence of clinical significance are represented in professional guidelines as established by consensus opinion of experts in the health care community.

Clinical Evidence Curation:
MSK-IMPACT uses a clinical evidence curation resource (OncoKB) to facilitate the clinical interpretation of detected mutations. OncoKB is a knowledge base that includes biologic, clinical and therapeutic information curated from multiple information resources including professional guidelines and recommendations, therapeutic labeling, disease specific expert and advocacy group recommendations, and medical literature. OncoKB information is publicly available through an interactive web site. Classification criteria were developed by MSK to communicate the level of clinical evidence available for individual mutations in the test report. The mutations are reported under two categories (i.e., cancer mutations panel with evidence of clinical significance and cancer mutations panel with potential clinical significance) based on the pre-specified classification criteria. OncoKB undergoes periodic updates through the review of new information by a panel of experts.

4. Clinical cut-off:

Not applicable.

5. Expected values:

The prevalence of somatic mutations was explored through a large-scale, prospective clinical sequencing initiative using a comprehensive assay, MSK-IMPACT, through which tumor and matched normal sequence data from a cohort of more than 10,000 patients with advanced cancer and available pathological and clinical annotations was compiled. The prevalence of mutations and cancer type via the link to the publicly accessible data on cohort of tested patients and available pathological and clinical annotations was published by Zehir, A. et al., "Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients." 2017. 23(6):703-713. This information is also available at the following website. (http://www.cbioportal.org/study?id=msk_impact_2017#summary)

N. Instrument Name
Illumina HiSeq 2500 (qualified by MSK)

O. System Descriptions:

1. Modes of Operation:
The Illumina HiSeq2500 is a high throughput sequencing system using Sequencing-By-Synthesis chemistry.

2. Software:
FDA has reviewed applicant's Hazard Analysis and software development processes for this line of product types:

Yes _X_ or No ___

3. Level of Concern:
Moderate

4. Specimen Handling
Refer to Device Description section above.

5 Calibration and Quality Controls:
Refer to Device Description section above.

P. Other Supportive Instrument Performance Characteristics Data Not Covered In The "Performance Characteristics" Section Above:

To support the continuous implementation of process improvements to the existing 468 gene panel, protocols with specific procedures and acceptance criteria for modifications that could be anticipated at the time of submission were provided, reviewed by FDA, and cleared as part of this marketing authorization. Future modifications by MSK for the specified types of changes below that are made in accordance with the applicable validation strategy and the pre-specified success criteria would not require a new 510(k) submission. Significant changes such as adding new genes or variant types to the panel would require a new submission with appropriate validation.

Type of change	**Validation Strategy**	**Pre-specified success criteria**
New pre-analytical protocol, kits or reagents	Sequence at least 10 specimens with known mutations. Measure sequence coverage distribution, and call somatic mutations in all samples.	For cases sequenced to >200x, ensure that 95% of exons are covered to 100x or more. Concordance for known mutations should be >95%.
New library preparation protocol, kits, or reagents	Sequence at least 40 DNA specimens (tumor / normal pairs) or three pools previously sequenced by MSK-IMPACT. Measure sequence coverage distribution, and call	For cases sequenced to >200x, ensure that 95% of exons are covered to 100x or more. Concordance for calling somatic mutations with variant allele fraction >10% should be >98%.

Type of change		Validation Strategy	Pre-specified success criteria
		somatic mutations in all samples.	
Changes to probes for already analytically validated genes		Re-capture existing sequence libraries from at least 3 runs (at least 40 samples) with new probes, sequence, and analyze.	For cases sequenced to >200x, ensure that 95% of exons in analytically validated genes are covered to 100x or more. Concordance for calling somatic mutations with variant allele fraction >10% should be >98%.
New sequencing instrument or reagents using similar chemistry and technology, and the sequence depth and read length are not changed from previous platform.		Re-sequence existing captured libraries from at least 3 runs, and call somatic mutations in all samples.	Sequence coverage distribution and GC bias across targeted regions should be within 5% of prior sequencing runs. Concordance for calling somatic mutations with variant allele fraction >10% should be >98%.
Bioinformatics pipeline	Update to underlying annotation database or transcript isoforms	Reanalyze FASTQ files (raw sequencing reads) from at least 3 runs (at least 40 samples). Compare variants calls between the clinical analysis results and the current modified results	Confirm the changes do not change the variant call results. Confirm the annotations for the unaffected transcripts do not change. Confirm the annotations for the affected transcripts are modified as expected.
	Update to data management system and system database	Reanalyze FASTQ files (raw sequencing reads) from at least 3 runs (at least 40 samples) in production mode. Compare variants calls between the clinical analysis results and the current modified results	Ensure that all previously called mutations are recovered and the variants in the database of results are concordant with the variants in the pipeline output files

Type of change		Validation Strategy	Pre-specified success criteria
	Modification to an existing component of the analysis pipeline (e.g., tool or algorithm) where the underlying algorithm or main parameter settings (e.g. minimal coverage/VAF threshold for SNV/indel calling; MSIsensor score cut-off for MSI-H calling, etc.) are not changed.	Reanalyze FASTQ files (raw sequencing reads) from at least 3 runs (at least 40 samples). Compare variants calls between the clinical analysis results and the current modified results	Ensure that all previously called mutations are recovered and that newly detected mutations can be explained by pipeline modifications.

Q. Proposed Labeling:

The labeling is sufficient and it satisfies the requirements of 21 CFR Parts 801 and 809, as applicable, and the special controls for this device type.

R. Patient Perspectives

This submission did not include specific information on patient perspectives for this device.

S. Identified Risks to Health and Identified Mitigations:

Identified Risks to Health	Identified Mitigations
Incorrect performance of the test leading to false positives, false negatives	General controls and special control (b)(1)
Incorrect interpretation of test results	General controls and special controls (b)(1)(iii)(E) and (b)(2)

T. Benefit/Risk Determination

Summary of the Benefit(s)	The MSK-IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets) test provides comprehensive genomic profiling of tumor samples (point mutations, small insertions and deletions and microsatellite instability), in previously diagnosed cancer patients, for use by qualified health professionals in accordance with professional guidelines. There is probable clinical benefit of the device based on evidence from peer-reviewed clinical literature and analytical performance of the device in identifying genomic alterations.
Summary of the Risk(s)	Erroneous device results could adversely influence clinical interpretation and consultation for patients. The risk of an erroneous test result is mitigated by the analytical performance of this device. The accuracy of the test was demonstrated using clinical specimens covering a variety of clinically relevant variants across multiple tumor types and variant categories (i.e., point mutations, small insertions and deletions and microsatellite instability). The output of this device demonstrated a high degree of analytical concordance to comparator assays across multiple tumor types. Thus, the probable risk of this device is mitigated by the supportive analytical performance for the device, when clinical limitations and the established special controls, in combination with general controls, are considered.
Summary of Other Factors	Limitations statements in the test report and the established special controls, in combination with general controls, serve to mitigate the risks associated with the use of this device.
Conclusions Do the probable benefits outweigh the probable risks?	The probable clinical benefits of this device, which allows for detection of somatic mutations and MSI status in patients previously diagnosed with cancer, outweigh the probable risks that are mitigated by the special controls established for this device type, in combination with general controls.

U. Conclusion:

The information provided in this *de novo* submission is sufficient to classify this device into class II under regulation 21 CFR 866.6080. FDA believes that special controls, along with the applicable general controls, provide reasonable assurance of the safety and effectiveness of the device type. The device is classified under the following:

Product Code:	PZM
Device Type:	Next Generation Sequencing Based Tumor Profiling Test.
Class:	II (special controls)
Regulation:	21 CFR 866.6080

(a) *Identification.* A next generation sequencing (NGS) based tumor profiling test is a qualitative in vitro diagnostic test intended for NGS analysis of tissue specimens from malignant solid neoplasms to detect somatic mutations in a broad panel of targeted genes to aid in the management of previously diagnosed cancer patients by qualified health care professionals.

(b) *Classification.* Class II (special controls). A next generation sequencing based tumor profiling test must comply with the following special controls:

(1) Premarket notification submissions must include the following information:

(i) A detailed description of all somatic mutations that are intended to be detected by the test and that are adequately supported in accordance with paragraph (b)(1)(v) of this section and reported in the test results in accordance with paragraph (b)(2)(iv) of this section, including:

(A) A listing of mutations that are cancer mutations with evidence of clinical significance.

(B) As appropriate, a listing of mutations that are cancer mutations with potential clinical significance.

(ii) The indications for use must specify the following:

(A) The test is indicated for previously diagnosed cancer patients.

(B) The intended specimen type(s) and matrix (e.g., formalin-fixed, paraffin-embedded tumor tissue).

(C) The mutation types (e.g., single nucleotide variant, insertion, deletion, copy number variation or gene rearrangement) for which validation data has been provided.

(D) The name of the testing facility or facilities, as applicable.

(iii) A detailed device description including the following:

(A) A description of the test in terms of genomic coverage, as follows:

(*1*) Tabulated summary of all mutations reported, grouped according to gene and target region within each gene, along with the specific cDNA and amino acid positions for each mutation.

(*2*) A description of any within-gene targeted regions that cannot be reported and the data behind such conclusion.

(B) Specifications for specimen requirements including any specimen collection devices and preservatives, specimen volume, minimum tumor content, specimen handling, DNA extraction, and criteria for DNA quality and quantity metrics that are prerequisite to performing the assay.

(C) A detailed description of all test components, reagents, instrumentation, and software required. Detailed documentation of the device software including but not limited to, software applications and hardware-based devices that incorporate software.

(D) A detailed description of the methodology and protocols for each step of the test, including description of the quality metrics, thresholds, and filters at each step of the test that are implemented for final result reporting and a description of the metrics for run-failures, specimen-failures, invalids, as applicable.

(E) A list of links provided by the device to the user or accessed by the device for internal or external information (e.g., decision rules or databases) supporting clinical significance of test results for the panel or its elements in accordance with paragraphs (b)(1)(v) and (b)(2)(vi) of this section.

(F) A description of internal and external controls that are recommended or provided and control procedures. The description must identify those control elements that are incorporated into the testing procedure.

(iv) Information demonstrating analytical validity of the device according to analytical performance characteristics, evaluated either specifically for each gene/mutation or, when clinically and practically justified, using a representative approach based on other mutations of the same type, including:

(A) Data that adequately supports the intended specimen type (e.g., formalin-fixed, paraffin-embedded tumor tissue), specimen handling protocol, and nucleic acid purification for specific tumor types or for a pan-tumor claim.

(B) A summary of the empirical evidence obtained to demonstrate how the analytical quality metrics and thresholds were optimized.

(C) Device precision data using clinical samples to adequately evaluate intra-run, inter-run, and total variability. The samples must cover all mutation types tested (both positive and negative samples) and include samples near the limit of detection of the device. Precision must be assessed by agreement within replicates on the assay final result for each representative mutation, as applicable, and also supported by sequencing quality metrics for targeted regions across the panel.

(D) Description of the protocols and/or data adequately demonstrating the interchangeability of reagent lots and multiplexing barcodes.

(E) A description of the nucleic acid assay input concentration range and the

evidence to adequately support the range.

(F) A description of the data adequately supporting the limit of detection of the device

(G) A description of the data to adequately support device accuracy using clinical specimens representing the intended specimen type and range of tumor types, as applicable.

(*1*) Clinical specimens tested to support device accuracy must adequately represent the list of cancer mutations with evidence of clinical significance to be detected by the device.

(*2*) For mutations that are designated as cancer mutations with evidence of clinical significance and that are based on evidence established in the intended specimen type (e.g., tumor tissues) but for a different analyte type (e.g., protein, RNA) and/or a measurement (e.g., incorporating a score or copy number) and/or with an alternative technology (e.g., IHC, RT-qPCR, FISH), evidence of accuracy must include clinically adequate concordance between results for the mutation and the medically established biomarker test (e.g., evidence generated from an appropriately sized method comparison study using clinical specimens from the target population).

(*3*) For qualitative DNA mutations not described in paragraph (b)(1)(iv)(G)(*2*) of this section, accuracy studies must include both mutation-positive and wild-type results.

(H) Adequate device stability information.

(v) Information that adequately supports the clinical significance of the panel must include:

(A) Criteria established on what types and levels of evidence will clinically validate a mutation as a cancer mutation with evidence of clinical significance versus a cancer mutation with potential clinical significance.

(B) For representative mutations of those designated as cancer mutations with evidence of clinical significance, a description of the clinical evidence associated with such mutations, such as clinical evidence presented in professional guidelines, as appropriate, with method comparison performance data as described in paragraph (b)(1)(iv)(G) of this section.

(C) For all other mutations designated as cancer mutations with potential clinical significance, a description of the rationale for reporting.

(2) The 21 CFR 809.10 compliant labeling and any product information and test report generated, must include the following, as applicable:

(i) The intended use statement must specify the following:

(A) The test is indicated for previously diagnosed cancer patients.

(B) The intended specimen type(s) and matrix (e.g., formalin-fixed, paraffin-embedded tumor tissue).

(C) The mutation types (e.g., single nucleotide variant, insertion, deletion, copy number variation or gene rearrangement) for which validation data has been provided.

(D) The name of the testing facility or facilities, as applicable.

(ii) A description of the device and summary of the results of the performance studies performed in accordance with paragraphs (b)(1)(iii), (b)(1)(iv), and (b)(1)(v) of this section.

(iii) A description of applicable test limitations, including, for device specific mutations validated with method comparison data to a medically established test in the same intended specimen type, appropriate description of the level of evidence and/or the differences between next generation sequencing results and results from the medically established test (e.g., as described in professional guidelines).

(iv) A listing of all somatic mutations that are intended to be detected by the device and that are reported in the test results under the following two categories or equivalent designations, as appropriate: "cancer mutations panel with evidence of clinical significance" or "cancer mutations panel with potential clinical significance."

(v) For mutations reported under the category of "cancer mutations panel with potential clinical significance," a limiting statement that states "For the mutations listed in [cancer mutations panel with potential clinical significance or equivalent designation], the clinical significance has not been demonstrated [with adequate clinical evidence (e.g., by professional guidelines) in accordance with paragraph (b)(1)(v) of this section] or with this test."

(vi) For mutations under the category of "cancer mutations panel with evidence of clinical significance," or equivalent designation, link(s) for physicians to access internal or external information concerning decision rules or conclusions about the level of evidence for clinical significance that is associated with the marker in accordance with paragraph (b)(1)(v) of this section.

Appendix 1a:
List of hotspot mutations (i.e., commonly somatically mutated in cancers) for all genes in the MSK-IMPACT panel

Gene	Codons
ABL1	G250, Q252, Y253, E255, T315, F317, M351, F359, H396R
AKT1	E17,Q124,G171,E170
AKT2	V140
ALK	K1062,D1091,C1156,M1166,I1171,F1174,L1196,A1234,F1245,I1250,R1275,Y1278
APC	S1234,I1307,E1309,E1317,P1319,G1339,S1341,P1361,P1372,P1373,R1399,S1400, S1407,S1411,V1414,S1415,S1421,T1438,P1439,P1440,T1445,P1453,N1455,E1464, S1465,T1487,L1488,F1491,T1493,E1494,T1537,K1555,T1556,I1557,C1578
AR	T878,T8782,Q581
ARAF	S214
ARID1A	D1850,G2087
ARID2	R314,S297,R285,A1773
ASXL1	Y591,E635,G645,G646,E1102D
ASXL2	R591
ATM	D1853,R3008,R3376,E2164
ATRX	K1936,E625
BARD1	P24
BCL6	R594,R618
BCOR	N1425,N14591
BRAF	G464,G466,G469,Y472,N581,D594,F595,G596,L597,A598_T599,V600,V600_K601, K601,V60010,K6010,G4694,N5810,G4660
CARD11	R170
CBL	Y371,L380,C384,C404,R420Q
CDH1	T263
CDK4	R24
CDKN2A	S43,P48,A57,A68,D74,L78,P81,H83,D84,L97,D108,P114,H831,D1081,P1140
CEBPA	P23,H24,Q83,K304_Q305,E309_T310,Q312_K313,K313_V314,K313_V314,K313, E316_L317,E316_L317insQ
CHEK2	K373,K3732
CIC	R215
CREBBP	R1446,S1680,R14460
CRLF2	F232C
CSF1R	Y969C
CTCF	R377
CTNNB1	D32,S33,G34,I35,H36,S37,T40,T41,T42,A43,P44,S45,G48,K49,E53,K335,S376, S334,D324,T412,G349,S455,C619

DICER1	E1813
DIS3	R382,D488
DNMT1	E432
DNMT3A	G543,R635,S714,F731,R882,R8820
DOT1L	G1386
EGFR	R108,A289,G598,R677,E709,G719,K745_E749,K745_E746,E746_A750,E746_S752,E746_T751,E746_E749,E746_T751,L747_P753,L747_A750,L747_T751,L747_S752,L747_T751,L747_E749,L747,T751,S752_I759,D761,S768,V769_D770,D770_N771,H773_V774,R776,T790,L833,H835,T847,P848,T854,L858,L861,G863,L8587,A2898,R252,R222
EP300	D1399,D13990,C1164
EPHB1	R170
ERBB2	S310,L755,D769,A775_G776,G776,V777,V842,S3108,L7553,E930,R678
ERBB3	V1043,D297,M91
ERBB4	R711
ERCC2	D312
ESR1	Y537
ETV1	R187
ETV6	R369
EZH2	Y646,R690
FBXW7	G423,R465,R479,R505,S582,R689,R4652,R5054,R4792
FGFR2	S252,P253,C382,N549,N550,K659
FGFR3	R248,S249,G370,S371,Y373,G380,A391,K650,G697,S2492,Y3730
FGFR4	V550
FLT3	D835,I836,D8358
FOXL2	C134W
FUBP1	R430
GATA1	M1,S30,V74I
GATA2	G320,L321,L359,R362Q
GNA11	R183,Q209,R256
GNAQ	R183,Q209
GNAS	R201,Q227,R8448
GRIN2A	R1067
HIST1H3B	E74
HNF1A	W206,P291,G292
HRAS	G12,G13,Q61,E62,Q614,G136,G122
IDH1	G70,V71,R132,V178,R13239,P33
IDH2	R140,R172,V294,R1402,R1721
IL7R	K395
IRS2	G1057

JAK1	R873
JAK2	F537_K539,H538_K539,K539,I540_E543,R541_E543,N542_E543,E543_D544, V617, R683
JAK3	A572,A573,R657Q
KDR	S1100,E759
KEAP1	R470
KIT	D52,D419,Y503_F504,K509,M541,K550_K558,P551_V555,P551_E554,P551_M552, Y553_K558,E554_K558,Q556_V560,W557_K558,W557,W557_V559,W557_E561, W557_V559,K558_E562,K558,K558_V560,V559,V559_V560,V559_E561,V560,E56 Y570_L576,D572,L576,D579,K642,V654,T670,S715,D816,K818,D820,N822,Y823,V 25, D8160
KMT2C	V656
KRAS	G10_A11,G12,G13,V14,L19,Q22,T58,A59,Q61,K117,A146,G1242,G133,Q619, A146
LATS2	A3243,G3630
MAP2K1	Q56,K57,D67,P124,P1240,F53,E203
MAP2K4	R134
MAP3K1	S1330,S939
MAPK1	E322
MAX	R600
MED12	L36,Q43,G44,L1224,L12240
MEF2B	D83V
MET	T1010,Y1248,Y1253,M1268,K1360
MLL3	K2797
MPL	S505,W515,W515R
MSH6	F1088,T1219I
MTOR	S22152,F1888
MYC	T58
MYCN	P44
MYD88	S219,S243,L265P
NF1	L844
NFE2L2	D29,L30,G31,R34,E79,T80,G81,E82,E794,D294,R342
NOTCH1	L1574,L1575,V1578,L1585,L1586,F1592,L1593,L1594,R1598,R1599,L1600, L1601,L1678,L1679,Q2460,P2514,A1944
NOTCH2	E385,N463
NPM1	W288,W290
NRAS	G12,G13,A18,G60,Q61,Q6193,G128,G138
NTRK1	T264
PAK7	E144
PARP1	I562
PAX5	P80R

PDGFRA	V561,S566_E571,N659,D842,I843_D846,D1071N
PIK3C2G	S670
PIK3CA	R38,E81,R88,R93,G106,R108,K111,G118,V344,N345,C378,E418,C420,E453, P539,E542,E545,Q546,E547,S553,K567,H701,E726,C901,G1007,Y1021,T1025, M1043,N1044,D1045,A1046,H1047,G1049,T1052,A1066,N1068,E54534,H104715, E54217,Q5467,R887,N3453,C4209,G1187,E7265,E4535, K1113, R932, R382, R1080, E39
PIK3R1	G376,D560,N564,K567
POLE	P2864,V4111
PPP2R1A	P179,R182,R183,S256,W257,R258,R1832
PREX2	G233C
PTCH1	P1315
PTEN	K6,P38,L42,H61,Y68,Y76,Y88,H93,I101,C105,L112,H123,A126,G129,R130,C136, A151,Y155,R159,K164,G165,S170,R173,N184,E242,P246,P248,C250,K267,V290, L318,T319,T321,N323,F347,R1309,R1730,K128
PTPN11	G60,D61,E69,A72,T73,E76,S502,G503,Q510
PTPRD	S431,P666
RAC1	P295
RAF1	S2570
RET	E632_T636,E632_L633,C634,M918T
RHOA	E40,Y42
RICTOR	S1101
RIT1	M90
RUNX1	L56,R107,D198,R201,R204,R162,R205
SDHA	S4560,A466,R465
SF3B1	E622,R625,H662,K666,K700,K7002
SMAD4	A118,D351,R361,G386,R3619,D537,P356
SMARCA4	T910,G1232
SMARCB1	R377,A382,P383
SMO	W535L
SPOP	F133,F1338,W131,F102
SRSF2	P95,P95_R102,P107H
STAG2	R370
STK11	D194,P281,F354L
TET2	C25,C262,Q764,F868,R1261,H1380,V1718L
TNFAIP3	L324
TP53	E11,D49,P82,T102,G105,Y107,R110,L111,F113,K120,T125,Y126,Y126_K132,S127, P128,L130,N131,K132,M133,F134,C135,A138,K139,T140,C141,P142,V143,Q144, L145,V147,S149,P151,P152,P153,G154,T155,R156,V157,R158,A159,M160,A161,

	I162,Y163,K164,S166,H168,M169,T170,E171,V172,V173,R174,R175,C176,P177, P177_C182,H178,H179,E180,R181,C182,D184,D186,G187,P190,P191,Q192,H193, L194,I195,R196,V197,E198,G199,N200,R202,V203,Y205,D208,R209,T211,F212,R21 , ,S215,V216,V217,V218,Y220,E224,G226,S227,D228,C229,T230,I232,Y234,N235,Y2 6, M237,C238,N239,S240,S241,C242,M243,G244,G245,M246,N247,R248,R249,P250 I251,L252,T253,I254,I255,L257,E258,D259,G262,L265,G266,R267,F270,E271,V272, R273,V274,C275,A276,C277,P278,G279,R280,D281,R282,R283,T284,E285,E286,E28 7, N288,R290,K291,K292,E294,P300,P301,S303,K320,G334,R337,R27328,R24892, R17538,R2820,G2451,Y2202,H1938,H1797,R1583,C1763,P2783,Y1633,R2800, G2660,I1950,S2419,R2499,V1577,C2386,E2856,R3375,G2445,V1733,P1512,C2752, K1321,Y2050,V2720,C1359,D2818,E2718,V2168,M2378,Y2347,E2867,L1946, A1596,R2675,S1275,C2425,Y2364,C1414,F2704,A1613,V2743,S2153,R2132,H2142, R1101,N2390,T1550,P1520,P2500,G1050,L1300,Q136,F109
TP63	R379
TSC2	N1515
TSHR	M453,I486,L512,I568,D619,A623,L629,I630,T632,D633,D633E
U2AF1	S34,Q157,S347
VHL	V62,S65,S72,V74,F76,N78,S80,P81,L85,P86,L89,N90,S111,G114,H115,L118,D121, L128,V130,G144,F148,I151,L153,V155,L158,E160,C162,V166,R167,L169,L184
WT1	V303,R312,A314,R394,D396,R462
XPO1	E571,R749

Appendix 1b: List of genes/transcripts included on the MSK-IMPACT panel

Gene Name	Transcript ID
ABL1	NM_005157
ACVR1	NM_001111067
AGO2	NM_012154
AKT1	NM_001014431
AKT2	NM_001626
AKT3	NM_005465
ALK	NM_004304
ALOX12B	NM_001139
AMER1	NM_152424
ANKRD11	NM_013275
APC	NM_000038
AR	NM_000044
ARAF	NM_001654
ARID1A	NM_006015
ARID1B	NM_020732
ARID2	NM_152641
ARID5B	NM_032199
ASXL1	NM_015338
ASXL2	NM_018263
ATM	NM_000051
ATR	NM_001184
ATRX	NM_000489
AURKA	NM_003600
AURKB	NM_004217
AXIN1	NM_003502
AXIN2	NM_004655
AXL	NM_021913
B2M	NM_004048
BABAM1	NM_001033549
BAP1	NM_004656
BARD1	NM_000465
BBC3	NM_001127240
BCL10	NM_003921
BCL2	NM_000633
BCL2L1	NM_138578
BCL2L11	NM_138621
BCL6	NM_001706
BCOR	NM_001123385
BIRC3	NM_182962
BLM	NM_000057
BMPR1A	NM_004329
BRAF	NM_004333
BRCA1	NM_007294
BRCA2	NM_000059
BRD4	NM_058243
BRIP1	NM_032043
BTK	NM_000061
CALR	NM_004343
CARD11	NM_032415
CARM1	NM_199141
CASP8	NM_001080125
CBFB	NM_022845
CBL	NM_005188
CCND1	NM_053056
CCND2	NM_001759
CCND3	NM_001760
CCNE1	NM_001238
CD274	NM_014143
CD276	NM_001024736
CD79A	NM_001783
CD79B	NM_001039933
CDC42	NM_001791
CDC73	NM_024529

CDH1	NM_004360
CDK12	NM_016507
CDK4	NM_000075
CDK6	NM_001145306
CDK8	NM_001260
CDKN1A	NM_078467
CDKN1B	NM_004064
CDKN2Ap14ARF	NM_058195
CDKN2Ap16INK4A	NM_000077
CDKN2B	NM_004936
CDKN2C	NM_078626
CEBPA	NM_004364
CENPA	NM_001809
CHEK1	NM_001274
CHEK2	NM_007194
CIC	NM_015125
CREBBP	NM_004380
CRKL	NM_005207
CRLF2	NM_022148
CSDE1	NM_001242891
CSF1R	NM_005211
CSF3R	NM_000760
CTCF	NM_006565
CTLA4	NM_005214
CTNNB1	NM_001904
CUL3	NM_003590
CXCR4	NM_003467
CYLD	NM_001042355
CYSLTR2	NM_020377
DAXX	NM_001141970
DCUN1D1	NM_020640
DDR2	NM_006182
DICER1	NM_030621
DIS3	NM_014953
DNAJB1	NM_006145
DNMT1	NM_001379
DNMT3A	NM_022552
DNMT3B	NM_006892
DOT1L	NM_032482
DROSHA	NM_013235
DUSP4	NM_001394
E2F3	NM_001949
EED	NM_003797
EGFL7	NM_201446
EGFR	NM_005228
EIF1AX	NM_001412
EIF4A2	NM_001967
EIF4E	NM_001130678
ELF3	NM_004433
EP300	NM_001429
EPAS1	NM_001430
EPCAM	NM_002354
EPHA3	NM_005233
EPHA5	NM_004439
EPHA7	NM_004440
EPHB1	NM_004441
ERBB2	NM_004448
ERBB3	NM_001982
ERBB4	NM_005235
ERCC2	NM_000400
ERCC3	NM_000122
ERCC4	NM_005236
ERCC5	NM_000123
ERF	NM_006494
ERG	NM_182918
ERRFI1	NM_018948

ESR1	NM_001122740
ETV1	NM_001163147
ETV6	NM_001987
EZH1	NM_001991
EZH2	NM_004456
FAM175A	NM_139076
FAM46C	NM_017709
FAM58A	NM_152274
FANCA	NM_000135
FANCC	NM_000136
FAT1	NM_005245
FBXW7	NM_033632
FGF19	NM_005117
FGF3	NM_005247
FGF4	NM_002007
FGFR1	NM_001174067
FGFR2	NM_000141
FGFR3	NM_000142
FGFR4	NM_213647
FH	NM_000143
FLCN	NM_144997
FLT1	NM_002019
FLT3	NM_004119
FLT4	NM_182925
FOXA1	NM_004496
FOXL2	NM_023067
FOXO1	NM_002015
FOXP1	NM_001244814
FUBP1	NM_003902
FYN	NM_153047
GATA1	NM_002049
GATA2	NM_032638
GATA3	NM_002051
GLI1	NM_005269
GNA11	NM_002067
GNAQ	NM_002072
GNAS	NM_000516
GPS2	NM_004489
GREM1	NM_013372
GRIN2A	NM_001134407
GSK3B	NM_002093
H3F3A	NM_002107
H3F3B	NM_005324
H3F3C	NM_001013699
HGF	NM_000601
HIST1H1C	NM_005319
HIST1H2BD	NM_021063
HIST1H3A	NM_003529
HIST1H3B	NM_003537
HIST1H3C	NM_003531
HIST1H3D	NM_003530
HIST1H3E	NM_003532
HIST1H3F	NM_021018
HIST1H3G	NM_003534
HIST1H3H	NM_003536
HIST1H3I	NM_003533
HIST1H3J	NM_003535
HIST2H3C	NM_021059
HIST2H3D	NM_001123375
HIST3H3	NM_003493
HLA-A	NM_001242758
HLA-B	NM_005514
HNF1A	NM_000545
HOXB13	NM_006361
HRAS	NM_001130442
ICOSLG	NM_015259

ID3	NM_002167
IDH1	NM_005896
IDH2	NM_002168
IFNGR1	NM_000416
IGF1	NM_001111285
IGF1R	NM_000875
IGF2	NM_001127598
IKBKE	NM_014002
IKZF1	NM_006060
IL10	NM_000572
IL7R	NM_002185
INHA	NM_002191
INHBA	NM_002192
INPP4A	NM_001134224
INPP4B	NM_001101669
INPPL1	NM_001567
INSR	NM_000208
IRF4	NM_002460
IRS1	NM_005544
IRS2	NM_003749
JAK1	NM_002227
JAK2	NM_004972
JAK3	NM_000215
JUN	NM_002228
KDM5A	NM_001042603
KDM5C	NM_004187
KDM6A	NM_021140
KDR	NM_002253
KEAP1	NM_203500
KIT	NM_000222
KLF4	NM_004235
KMT2A	NM_001197104
KMT2B	NM_014727
KMT2C	NM_170606
KMT2D	NM_003482
KNSTRN	NM_033286
KRAS	NM_033360
LATS1	NM_004690
LATS2	NM_014572
LMO1	NM_002315
LYN	NM_002350
MALT1	NM_006785
MAP2K1	NM_002755
MAP2K2	NM_030662
MAP2K4	NM_003010
MAP3K1	NM_005921
MAP3K13	NM_004721
MAP3K14	NM_003954
MAPK1	NM_002745
MAPK3	NM_002746
MAPKAP1	NM_001006617
MAX	NM_002382
MCL1	NM_021960
MDC1	NM_014641
MDM2	NM_002392
MDM4	NM_002393
MED12	NM_005120
MEF2B	NM_001145785
MEN1	NM_000244
MET	NM_000245
MGA	NM_001164273
MITF	NM_198159
MLH1	NM_000249
MPL	NM_005373
MRE11A	NM_005591
MSH2	NM_000251

MSH3	NM_002439
MSH6	NM_000179
MSI1	NM_002442
MSI2	NM_138962
MST1	NM_020998
MST1R	NM_002447
MTOR	NM_004958
MUTYH	NM_001128425
MYC	NM_002467
MYCL1	NM_001033082
MYCN	NM_005378
MYD88	NM_002468
MYOD1	NM_002478
NBN	NM_002485
NCOA3	NM_181659
NCOR1	NM_006311
NEGR1	NM_173808
NF1	NM_001042492
NF2	NM_000268
NFE2L2	NM_006164
NFKBIA	NM_020529
NKX2-1	NM_001079668
NKX3-1	NM_006167
NOTCH1	NM_017617
NOTCH2	NM_024408
NOTCH3	NM_000435
NOTCH4	NM_004557
NPM1	NM_002520
NRAS	NM_002524
NSD1	NM_022455
NTHL1	NM_002528
NTRK1	NM_002529
NTRK2	NM_006180
NTRK3	NM_001012338
NUF2	NM_031423
NUP93	NM_014669
PAK1	NM_002576
PAK7	NM_177990
PALB2	NM_024675
PARK2	NM_004562
PARP1	NM_001618
PAX5	NM_016734
PBRM1	NM_018313
PDCD1	NM_005018
PDCD1LG2	NM_025239
PDGFRA	NM_006206
PDGFRB	NM_002609
PDPK1	NM_002613
PGR	NM_000926
PHOX2B	NM_003924
PIK3C2G	NM_004570
PIK3C3	NM_002647
PIK3CA	NM_006218
PIK3CB	NM_006219
PIK3CD	NM_005026
PIK3CG	NM_002649
PIK3R1	NM_181523
PIK3R2	NM_005027
PIK3R3	NM_003629
PIM1	NM_002648
PLCG2	NM_002661
PLK2	NM_006622
PMAIP1	NM_021127
PMS1	NM_000534
PMS2	NM_000535
PNRC1	NM_006813

POLD1	NM_002691
POLE	NM_006231
PPARG	NM_015869
PPM1D	NM_003620
PPP2R1A	NM_014225
PPP4R2	NM_174907
PPP6C	NM_002721
PRDM1	NM_001198
PRDM14	NM_024504
PREX2	NM_024870
PRKAR1A	NM_212471
PRKCI	NM_002740
PRKD1	NM_002742
PTCH1	NM_000264
PTEN	NM_000314
PTP4A1	NM_003463
PTPN11	NM_002834
PTPRD	NM_002839
PTPRS	NM_002850
PTPRT	NM_133170
RAB35	NM_006861
RAC1	NM_018890
RAC2	NM_002872
RAD21	NM_006265
RAD50	NM_005732
RAD51	NM_002875
RAD51B	NM_133509
RAD51C	NM_058216
RAD51D	NM_133629
RAD52	NM_134424
RAD54L	NM_001142548
RAF1	NM_002880
RARA	NM_000964
RASA1	NM_002890
RB1	NM_000321
RBM10	NM_001204468
RECQL	NM_032941
RECQL4	NM_004260
REL	NM_002908
RET	NM_020975
RFWD2	NM_022457
RHEB	NM_005614
RHOA	NM_001664
RICTOR	NM_152756
RIT1	NM_006912
RNF43	NM_017763
ROS1	NM_002944
RPS6KA4	NM_003942
RPS6KB2	NM_003952
RPTOR	NM_020761
RRAGC	NM_022157
RRAS	NM_006270
RRAS2	NM_012250
RTEL1	NM_032957
RUNX1	NM_001754
RXRA	NM_002957
RYBP	NM_012234
SDHA	NM_004168
SDHAF2	NM_017841
SDHB	NM_003000
SDHC	NM_003001
SDHD	NM_003002
SESN1	NM_014454
SESN2	NM_031459
SESN3	NM_144665
SETD2	NM_014159

SETD8	NM_020382
SF3B1	NM_012433
SH2B3	NM_005475
SH2D1A	NM_002351
SHOC2	NM_007373
SHQ1	NM_018130
SLX4	NM_032444
SMAD2	NM_001003652
SMAD3	NM_005902
SMAD4	NM_005359
SMARCA4	NM_003072
SMARCB1	NM_003073
SMARCD1	NM_003076
SMO	NM_005631
SMYD3	NM_001167740
SOCS1	NM_003745
SOS1	NM_005633
SOX17	NM_022454
SOX2	NM_003106
SOX9	NM_000346
SPEN	NM_015001
SPOP	NM_001007228
SPRED1	NM_152594
SRC	NM_198291
SRSF2	NM_003016
STAG2	NM_001042749
STAT3	NM_139276
STAT5A	NM_003152
STAT5B	NM_012448
STK11	NM_000455
STK19	NM_004197
STK40	NM_032017
SUFU	NM_016169
SUZ12	NM_015355
SYK	NM_003177
TAP1	NM_000593
TAP2	NM_018833
TBX3	NM_016569
TCEB1	NM_005648
TCF3	NM_001136139
TCF7L2	NM_001146274
TEK	NM_000459
TERT	NM_198253
TET1	NM_030625
TET2	NM_001127208
TGFBR1	NM_004612
TGFBR2	NM_001024847
TMEM127	NM_001193304
TMPRSS2	NM_001135099
TNFAIP3	NM_006290
TNFRSF14	NM_003820
TOP1	NM_003286
TP53	NM_000546
TP53BP1	NM_001141980
TP63	NM_003722
TRAF2	NM_021138
TRAF7	NM_032271
TSC1	NM_000368
TSC2	NM_000548
TSHR	NM_000369
U2AF1	NM_006758
UPF1	NM_002911
VEGFA	NM_001171623
VHL	NM_000551
VTCN1	NM_024626
WHSC1	NM_001042424

WHSC1L1	NM_023034
WT1	NM_024426
WWTR1	NM_001168280
XIAP	NM_001167
XPO1	NM_003400
XRCC2	NM_005431
YAP1	NM_001130145
YES1	NM_005433
ZFHX3	NM_006885

Appendix 1c: List of genes/exons excluded from reporting due to consistently low coverage.

Gene	Transcript ID	Chromosome Coordinates	Exon	cDNA	Amino Acid
AGO2	NM_012154	8:141645584-141645605	1	1_22	1_8
ANKRD11	NM_013275	16:89334886-89335071	13	7807_7992	2603_2664
CD276	NM_001024736	15:73995113-73995427	4	419_733	140_245
CD276	NM_001024736	15:73996517-73996813	6	1073_1369	358_457
CHEK2	NM_007194	22:29085123-29085203	14	1462_1542	488_514
FAM58A	NM_152274	X:152864420-152864529	1	1176_1	392_1
FLT3	NM_004119	13:28674605-28674647	1	1_43	1_15
H3F3A	NM_002107	1:226259052-226259180	4	283_411	95_137
HIST2H3C	NM_021059	1:149812319-149812729	1	411_1	137_1
HIST2H3D	NM_001123375	1:149784826-149785236	1	1_411	1_137
HLA-A	NM_001242758	6:29911899-29912174	4	620_895	207_299
INSR	NM_000208	19:7293803-7293902	1	1_100	1_34
KMT2C	NM_170606	7:151970790-151970952	7	850_1012	284_338
KMT2C	NM_170606	7:151962123-151962294	8	1013_1184	338_395
KMT2C	NM_170606	7:151935792-151935911	15	2533_2652	845_884
KMT2C	NM_170606	7:151932902-151933018	16	2653_2769	885_923
KMT2C	NM_170606	7:151927008-151927112	18	2872_2976	958_992
KMT2C	NM_170606	7:151921520-151921701	19	2977_3158	993_1053
KMT2C	NM_170606	7:151921100-151921264	20	3159_3323	1053_1108
KMT2C	NM_170606	7:151919658-151919767	21	3324_3433	1108_1145
KMT2C	NM_170606	7:151904385-151904513	24	3713_3841	1238_1281
MST1	NM_020998	3:49726031-49726124	1	1_94	1_32
MST1	NM_020998	3:49724380-	6	608_728	203_243

		49724500			
MST1	NM_020998	3:49724117-49724235	7	729_847	243_283
MST1	NM_020998	3:49723746-49723914	8	848_1016	283_339
MST1	NM_020998	3:49723495-49723625	9	1017_1147	339_383
MST1	NM_020998	3:49722695-49722815	13	1424_1544	475_515
MST1	NM_020998	3:49722445-49722522	14	1545_1622	515_541
MST1	NM_020998	3:49721983-49722089	16	1770_1876	590_626
MST1	NM_020998	3:49721747-49721886	17	1877_2016	626_672
MYCL1	NM_001033082	1:40367480-40367560	1	1_81	1_27
NOTCH2	NM_024408	1:120611948-120612020	1	1_73	1_25
NOTCH2	NM_024408	1:120572529-120572610	2	74_155	25_52
NOTCH2	NM_024408	1:120547952-120548211	3	156_415	52_139
NOTCH2	NM_024408	1:120539620-120539955	4	416_751	139_251
NOTCH3	NM_000435	19:15311599-15311716	1	1_118	1_40
PDPK1	NM_002613	16:2588114-2588137	1	1_24	1_8
PDPK1	NM_002613	16:2607704-2607964	2	25_285	9_95
PDPK1	NM_002613	16:2611481-2611523	3	286_328	96_110
PDPK1	NM_002613	16:2611772-2611909	4	329_466	110_156
PDPK1	NM_002613	16:2615554-2615698	5	467_611	156_204
PDPK1	NM_002613	16:2616357-2616454	6	612_709	204_237
PDPK1	NM_002613	16:2627426-2627501	7	710_785	237_262
PDPK1	NM_002613	16:2631296-2631364	8	786_854	262_285
PDPK1	NM_002613	16:2631608-2631704	9	855_951	285_317
PDPK1	NM_002613	16:2633413-2633586	10	952_1125	318_375
PIK3CA	NM_006218	3:178937737-178937840	13	1912_2015	638_672
PIK3R2	NM_005027	19:18272089-18272305	6	599_815	200_272
PMS2	NM_000535	7:6022455-6022622	12	2007_2174	669_725
PMS2	NM_000535	7:6018227-6018327	13	2175_2275	725_759
PMS2	NM_000535	7:6017219-6017388	14	2276_2445	759_815
PMS2	NM_000535	7:6013030-6013173	15	2446_2589	816_863
PPP4R2	NM_174907	3:73096337-73096507	3	117_287	39_96

PTEN	NM_000314	10:89725044-89725229	9	1027_1212	343_404
PTPRT	NM_133170	20:41818286-41818373	1	1_88	1_30
RECQL	NM_032941	12:21623128-21623280	16	1798_1950	600_650
RECQL4	NM_004260	8:145743085-145743168	1	1_84	1_28
SDHA	NM_004168	5:254508-254621	14	1795_1908	599_636
SDHC	NM_003001	1:161332119-161332223	6	406_405	136_135
SDHD	NM_003002	11:111965529-111965694	4	315_480	105_160
SETD8	NM_020382	12:123873980-123874101	2	11_132	4_44
SETD8	NM_020382	12:123892040-123892250	8	849_1059	283_353
STAT5A	NM_003152	17:40452148-40452299	8	682_833	228_278
STAT5A	NM_003152	17:40452733-40452888	9	834_989	278_330
STAT5B	NM_012448	17:40371330-40371481	7	682_833	228_278
STAT5B	NM_012448	17:40370741-40370896	8	834_989	278_330
STK19	NM_004197	6:31948781-31948826	8	1050_1095	350_365
SUZ12	NM_015355	17:30267305-30267351	2	275_321	92_107
SUZ12	NM_015355	17:30267441-30267505	3	322_386	108_129
SUZ12	NM_015355	17:30274636-30274704	4	387_455	129_152
SUZ12	NM_015355	17:30300165-30300250	6	506_591	169_197
SUZ12	NM_015355	17:30310018-30310123	9	918_1023	306_341
TGFBR1	NM_004612	9:101867488-101867584	1	1_97	1_33