# Validation of Illumina TruSeq® Amplicon Cancer Panel with Concordance Testing Using Ion AmpliSeq Cancer Panel and Other Methods

AACR 2013
Abstract #3229

Peng Fang, Zhenyu Yan, Weihua Liu , Agus Darwanto, Kim Pelak, Kim Anoe, Cindy Spittle, Sabita Sankar, Chad Galderisi, and Jin Li, MolecularMD Corp., Cambridge, MA and Portland, OR

## Introduction

The Illumina TruSeq® Amplicon Cancer Panel (TSACP), a highly multiplexed targeted resequencing assay for use on the Illumina MiSeq platform, is designed for detecting the hotspot mutations in 212 Regions of Interest (ROI) from 48 cancer-related genes (**Table 1**). We used well characterized cancer cell lines harboring clinically relevant variants as positive controls and HapMap samples NA12878 and NA19240 as wild type control samples. A total of 41 FFPE patient specimens representing a variety of cancer types were analyzed in a blinded fashion to evaluate the analytical sensitivity and specificity of the assay. With no gold standard available as a reference method to detect mutations with comparable sensitivity, concordance testing was performed using the Ion AmpliSeq Cancer Panel. Variants detected by both panels were considered as true positives. Variants that were either not covered by the Ion Panel or not concordantly detected were confirmed by a third method, either Sanger sequencing (variants with frequencies > 10%) or a MolecularMD custom Ion TargetSeq Assay (variants with frequencies <10%).

**Table 1. The 48 genes targeted by Illumina TSACP assay**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ABL1 | BRAF | EGFR | FGFR2 | GNAS | JAK3 | MLH1 | PDGFRA | RET | STK11 |
| AKT1 | CDH1 | ERBB2 | FGFR3 | HNF1A | KDR | MPL | PIK3CA | SMAD4 | TP53 |
| ALK | CDKN2A | ERBB4 | FLT3 | HRAS | KIT | NOTCH1 | PTEN | SMARCB1 | VHL |
| APC | CSF1R | FBXW7 | GNA11 | IDH1 | KRAS | NPM1 | PTPN11 | SMO | |
| ATM | CTNNB1 | FGFR1 | GNAQ | JAK2 | MET | NRAS | RB1 | SRC | |

## Materials and Methods

The DNA samples were quantified with Qubit 2.0 (Life Technologies), and the quality of the FFPE DNA was assessed with a qPCR based, Infinium HD FFPE QC Kit (Illumina). 150 ng genomic DNA or 250 ng qualified FFPE DNA was hybridized to the pre-mixed oligonucleotide probes that flank the 212 ROI. An extension reaction extends across the ROI, followed by ligation to unite probe pairs, creating new template strands. The templates generated from extension-ligation were PCR-amplified with the primers adding the index sequences for sample multiplexing as well as the common adapters required for the cluster generation. The indexed library products were purified with AMPure XP (Beckman-Coulter), normalized to equal concentration, pooled and loaded on MiSeq for automated cluster amplification and sequencing. Data were analyzed using MiSeq Reporter software and our proprietary analysis pipeline. In addition to reporting hotspot mutations, we also report "Critical Variants" such as non-synonymous coding mutations and splicing site mutations that fall within the ROI.

## Results

**ROI Coverage**: In the seven sequencing runs (16 samples per flow cell) performed during our validation study, the mean sequencing coverage of each run ranged from 1607 to 4029 (**Table 2**). Nine ROI, or 4.4% of total ROI, were consistently sequenced with less than 0.2x normalized coverage (**Table 2** and **Table 3**, highlighted green), whereas between 4 and 23 ROI had less than 500x read depth, or 2.2% to 12.6% of total ROI, depending on the mean coverage of each run (**Table 3**, highlighted orange).

**Table 2. Summary of ROI coverage**

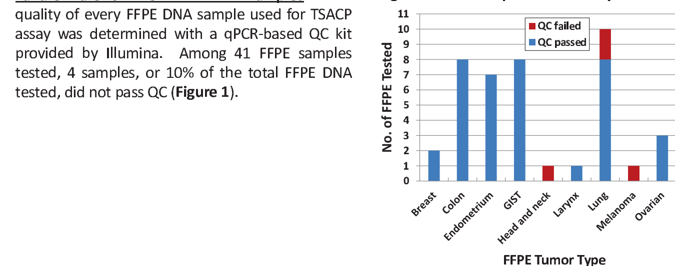
| Run# | Average coverage per sample | DNA loading on MiSeq (cluster density, k/mm²) | % ROI ≤500 depth | % ROI ≤0.2x coverage |
|---|---|---|---|---|
| Run#3 | 1607 | 443 | 12.6% | 4.4% |
| Run#4 | 1683 | 417 | 12.1% | |
| Run#1 | 1750 | 398 | 11.5% | |
| Run#6 | 2480 | 629 | 4.9% | |
| Run#5 | 2564 | 631 | 4.4% | |
| Run#2 | 4007 | 996 | 2.2% | |
| Run#7 | 4029 | 1095 | 2.2% | |

**Table 3. ROI covered at less than 500x read depth**

| Region of Interest | Gene | Normalized coverage | Average sequencing depth per ROI | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Run#3 | Run#4 | Run#1 | Run#6 | Run#5 | Run#2 | Run#7 |
| Average coverage per sample | | | 1607 | 1683 | 1750 | 2480 | 2564 | 4007 | 4029 |
| 9:21971149-21971282 | CDKN2A1 | 0.01 | 10 | 10 | 10 | 15 | 15 | 24 | 24 |
| 4:1808318-1808442 | FGFR3_8 | 0.10 | 154 | 161 | 168 | 238 | 246 | 384 | 386 |
| 20:36031762-36031895 | SRC | 0.11 | 184 | 193 | 201 | 284 | 294 | 459 | 462 |
| 19:3118827-3118963 | GNA11_5 | 0.12 | 189 | 198 | 206 | 292 | 302 | 471 | 474 |
| 3:10183795-10183926 | VHL1_2 | 0.14 | 228 | 239 | 249 | 352 | 364 | 569 | 573 |
| 7:128851591-128851708 | SMO5 | 0.15 | 244 | 256 | 266 | 377 | 390 | 609 | 612 |
| 4:1803564-1803685 | FGFR3_1_2 | 0.16 | 257 | 269 | 279 | 396 | 409 | 640 | 643 |
| 7:128846374-128846509 | SMO3 | 0.18 | 284 | 298 | 309 | 439 | 453 | 708 | 712 |
| 9:133738329-133738465 | ABL1_1 | 0.20 | 321 | 337 | 350 | 496 | 513 | 801 | 806 |
| 13:48955534-48955657 | RB1_4 | 0.23 | 369 | 386 | 401 | 569 | 588 | 919 | 924 |
| 9:139397767-139397901 | NOTCH1_1 | 0.24 | 387 | 405 | 422 | 597 | 617 | 965 | 971 |
| 19:3114847-3115111 | GNA11_4 | 0.25 | 401 | 420 | 436 | 618 | 639 | 999 | 1004 |
| 11:534279-534410 | HRAS6 | 0.25 | 409 | 428 | 446 | 631 | 653 | 1020 | 1026 |
| 4:1807859-1807999 | FGFR3_5_7 | 0.26 | 414 | 433 | 451 | 639 | 660 | 1032 | 1038 |
| 4:1806119-1806250 | FGFR3_3_4 | 0.26 | 422 | 442 | 460 | 651 | 673 | 1052 | 1058 |
| 4:55144546-55144676 | PDGFRA6 | 0.27 | 428 | 449 | 466 | 661 | 683 | 1068 | 1074 |
| 9:133747504-133747641 | ABL1_9 | 0.27 | 429 | 449 | 467 | 662 | 684 | 1069 | 1075 |
| 17:37880988-37881124 | ERBB2_4 | 0.27 | 432 | 452 | 470 | 666 | 688 | 1076 | 1082 |
| 10:43613834-43613971 | RET10 | 0.27 | 442 | 463 | 481 | 682 | 705 | 1101 | 1108 |
| 13:28610124-28610239 | FLT3_22 | 0.28 | 452 | 473 | 492 | 697 | 720 | 1126 | 1132 |
| 19:1220486-1220614 | STK11_4 | 0.28 | 457 | 478 | 497 | 705 | 729 | 1139 | 1145 |
| 7:128850336-128850472 | SMO4 | 0.30 | 476 | 498 | 518 | 734 | 759 | 1187 | 1193 |
| 1:115256527-115256653 | NRAS1_7 | 0.31 | 492 | 515 | 535 | 759 | 784 | 1226 | 1233 |
| 11:108218089-108218220 | ATM17 | 0.33 | 534 | 559 | 581 | 823 | 851 | 1330 | 1338 |

**Failure Rate of the FFPE DNA Samples**: The quality of every FFPE DNA sample used for TSACP assay was determined with a qPCR-based QC kit provided by Illumina. Among 41 FFPE samples tested, 4 samples, or 10% of the total FFPE DNA tested, did not pass QC (**Figure 1**).

**Figure 1. FFPE sample QC summary**




**Mutations Identified**: Among 38 samples (37 FFPE and 1 mixed cell line DNA), with 5% LOD cut-off, 124 unique critical mutations (mutations that would affect the protein sequence) were identified. These included 107 single base substitutions (SBS, **Figure 2**), 14 unique single- or multi-base deletions (**Table 4**), and 3 single- or two- base insertions (**Table 5**). The mutation spectrum in each tumor type is shown in **Figure 3**.

**Figure 2. The number of SBS mutations identified** **Figure 3. Mutation spectrum detected in FFPE**

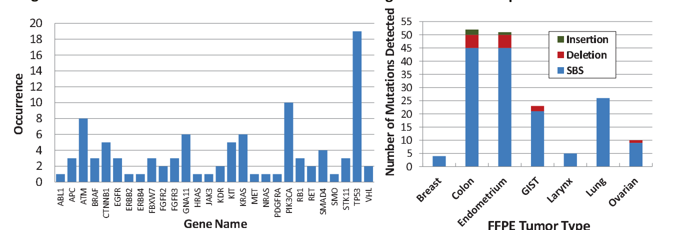



**Table 4. Mutations caused by deletions**

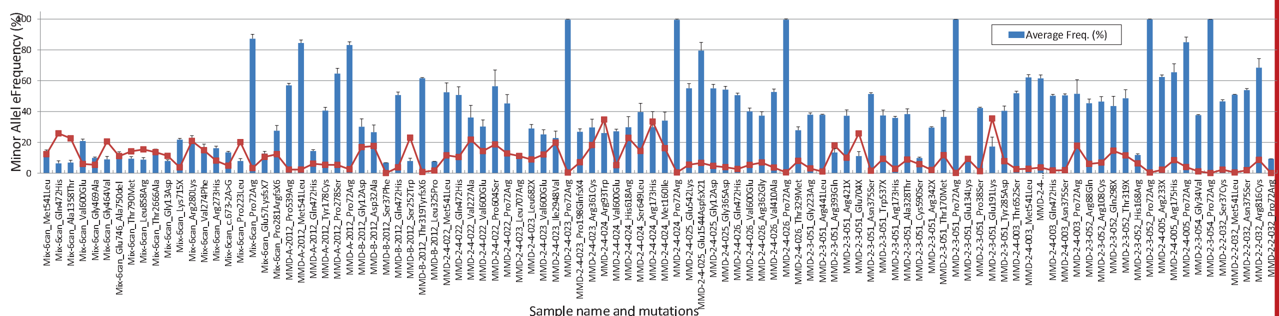
| Sample Name | Tumor Type | Chromosome coordinate | Gene name | Mutated CDS | Mutated Amino Acids | Freq (%) | Read Depth |
|---|---|---|---|---|---|---|---|
| MMD-2-4-002 | GIST | 4:55152092-55152103 | PDGFRA | c.2524_2535delGACATCATGCAT | p.Ile843_Asp846del | 25.9 | 8125 |
| MMD-2-4-003 | GIST | 4:55593589-55593594 | KIT | c.1655_1660delTGTATG | p.Met552_Glu554delinsLys | 60.0 | 1513 |
| MMD-2-4-004 | GIST | 4:55593640-55593660 | KIT | c.1706_1726delTTTACATAGACCCAACACAAC | p.Tyr570_Leu576del | 37.0 | 5682 |
| MMD-2-4-023 | colon | 4:55595555 | KIT | c.2045delT | p.Leu682X | 29.4 | 2811 |
| MMD-2-4-025 | colon | 5:112175923 | APC | c.4632delA | p.Glu1544AspfsX21 | 86.4 | 708 |
| Mix-6can | Cell line | 7:55242465-55242479 | EGFR | c.2235_2249delGGAATTAAGAGAAGC | p.Glu746_Ala750del | 10.5 | 3060 |
| MMD183 | colon | 9:133748412 | ABL1 | c.1073delA | p.Asn358ThrfsX14 | 9.0 | 4503 |
| MMD183 | colon | 10:89711914-89711916 | PTEN | c.532_534delTAT | p.Tyr178del | 23.8 | 1032 |
| MMD151 | endometrium | 10:89717758 | PTEN | c.783delG | p.Asn262ThrfsX4 | 97.8 | 2016 |
| MMD-2-3-063 | ovarian | 10:89720771 | PTEN | c.922delC | p.Arg308ValfsX9 | 42.0 | 7811 |
| MMD-2-3-052 | endometrium | 10:89720804-89720807 | PTEN | c.955_958delACTT | p.Thr319X | 52.5 | 4647 |
| MMD-2-3-053 | endometrium | 10:89720804-89720807 | PTEN | c.955_958delACTT | p.Thr319X | 46.6 | 4522 |
| MMD-2-4-023 | colon | 18:48581289 | SMAD4 | c.593delC | p.Pro198GlnfsX4 | 26.0 | 4686 |
| Mix-6can | Cell line | 19:1207081 | STK11 | c.169delG | p.Glu57LysfsX7 | 11.2 | 206 |
| Mix-6can | Cell line | 19:1221319 | STK11 | c.842delC | p.Pro281ArgfsX6 | 30.2 | 2032 |

**Table 5. Mutations caused by insertions**

| Sample Name | Tumor Type | Chromosome coordinate | Gene name | Mutated CDS | Mutated Amino Acids | Freq (%) | Read Depth |
|---|---|---|---|---|---|---|---|
| MMD-2-4-020 | colon | 5:112175480-112175481 | APC | c.4189dupG | p.Glu1397GlyfsX12 | 31.6 | 1372 |
| MMD183 | colon | 5:112175685-112175686 | APC | c.4393_4394dupAG | p.Ser1465ArgfsX9 | 26.2 | 6167 |
| MMD-B-2013 | endometrium | 10:89720803-89720804 | PTEN | c.954dupT | p.Thr319TyrfsX6 | 61.3 | 5547 |

**Analytical Reproducibility**: The analytical reproducibility of the TSACP assay was evaluated with a sample created by mixing DNA from 6 cancer cell lines (sample ID = Mix-6can) and DNA extracted from 13 FFPE tissue samples. With a 5% allele frequency cut-off, 100 of 102 expected mutations were reproducibly detected within or between runs (**Figure 4**). The two missed mutations were still correctly called at less than 5% allele frequency in those two samples, suggesting a variation of minor allele frequency determined by TSACP assay. The CV for the minor allele frequency of these mutations ranges from 0-38%, with the majority (88%) below 20%.

**Figure 4. Analytical reproducibility determined with DNA from 13 FFPE specimens and 1 cell line**




**Confirmation of Critical Mutations Identified**: For the 139 critical mutations identified in FFPE samples, 104 (75%) were cross-confirmed by the Ion Torrent AmpliSeq Cancer Panel Assay (**Figure 5**), while 32 were not covered by the ROI of the Ion assay. Of the 35 mutations either not covered by, or not concordant with the Ion assay, 3 mutations were not confirmed due to insufficient DNA available; 23 were analyzed by Sanger sequencing (all 23 were true positives), and 9 were analyzed by a MolecularMD custom Ion Torrent TargetSeq assay (5 of the 9 were true positives, **Table 6**). These same variant positions are also shown for the wild-type control NA12878. Of the 4 mutations called by the TSACP assay that were found to be false positives, 3 were located in the GNA11 gene which had less than 200x read depth (**Table 6**, highlighted).

**Figure 5. Mutation confirmation summary**

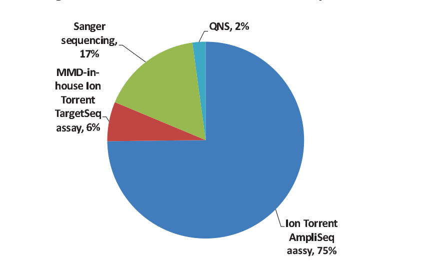



**Table 6. MolecularMD Custom Ion TargetSeq analysis for critical mutations below Sanger detection limit**

| Sample name | Chromosome coordinate | Gene name (HGNC) | Mutated CDS | Mutated Amino Acids | TSACP Allele Freq (%) | TSACP Read Depth | MMD-Ion TargetSeq Allele Freq (%) | MMD-Ion TargetSeq Read Depth | Result |
|---|---|---|---|---|---|---|---|---|---|
| MMD-2-2-025 | 3:10191533 | VHL | c.526A>T | p.Arg176Trp(176R>W) | 12.4 | 2714 | 13.4 | 47222 | TP |
| MMD-2-3-051 | 4:55144636 | PDGFRA | c.2110G>T | p.Glu704X(704E>*) | 9.1 | 561 | 12.8 | 7934 | TP |
| MMD-2-4-003 | 10:43613892 | RET | c.2356C>T | p.His786Tyr(786H>Y) | 5.4 | 594 | 0.0 | 45443 | FP |
| MMD-2-3-051 | 13:49027202 | RB1 | c.1769G>C | p.Cys590Ser(590C>S) | 10.2 | 763 | 10.1 | 28068 | TP |
| MMD-2-3-051 | 18:48575206 | SMAD4 | c.400G>A | p.Glu134Lys(134E>K) | 10.5 | 2134 | 9.2 | 21036 | TP |
| MMD-2-3-052 | 19:3114952 | GNA11 | c.487G>A | p.Asp163Asn(163D>N) | 8.3 | 168 | 0.0 | 5681 | FP |
| MMD-2-5-021 | 19:3114962 | GNA11 | c.497G>A | p.Arg166His(166R>H) | 5.8 | 120 | 0.0 | 1432 | FP |
| MMD-2-4-003 | 19:3114967 | GNA11 | c.502G>A | p.Ala168Thr(168A>T) | 5.6 | 195 | 0.0 | 2644 | FP |
| MMD-2-3-051 | 19:3115036 | GNA11 | c.571G>A | p.Glu191Lys(191E>K) | 13.0 | 1073 | 16.6 | 21984 | TP |
| NA12878 (wild-type control) | 3:10191533 | VHL | c.526A>T | p.Arg176Trp(176R>W) | | | 0.0 | 79589 | |
| | 4:55144636 | PDGFRA | c.2110G>T | p.Glu704X(704E>*) | | | 0.0 | 23737 | |
| | 10:43613892 | RET | c.2356C>T | p.His786Tyr(786H>Y) | | | 0.0 | 22460 | |
| | 13:49027202 | RB1 | c.1769G>C | p.Cys590Ser(590C>S) | | | 0.0 | 28524 | |
| | 18:48575206 | SMAD4 | c.400G>A | p.Glu134Lys(134E>K) | | | 0.0 | 13848 | |
| | 19:3114952 | GNA11 | c.487G>A | p.Asp163Asn(163D>N) | | | 0.0 | 2548 | |
| | 19:3114962 | GNA11 | c.497G>A | p.Arg166His(166R>H) | | | 0.0 | 6428 | |
| | 19:3114967 | GNA11 | c.502G>A | p.Ala168Thr(168A>T) | | | 0.0 | 2548 | |
| | 19:3115036 | GNA11 | c.571G>A | p.Glu191Lys(191E>K) | | | 0.0 | 16327 | |

**Sensitivity and Specificity of the TSACP assay**: The results shown above indicate that the sensitivity and specificity of the TSACP assay are 99% and 97%, respectively (**Table 7**). To further enhance the specificity of the assay, a stringent variant filter could be applied, requiring that the mutated nucleotide has to be covered by at least 25x read depth in order to be reported. This filter removes 3 FP mutations and 1 TP mutation. This stringent coverage filter reduces the FP rate from 3% to 1%, with 2 FN identified. Consequently, with the filter the sensitivity and specificity are 98% and 99% respectively (**Table 7**).

**Table 7. Sensitivity and specificity of the TSACP assay**

| | No Filter | Filter |
|---|---|---|
| Number of mutations reported without confirmatory test | 136 | 132 |
| True Positive (TP) | 132 | 131 |
| False Positive (FP) | 4 | 1 |
| Number of False Negative (FN) mutations missed | 1 | 2 |
| | | |
| **Sensitivity** (TP/(TP+FN)) | **99%** (132/(132+1)) | **98%** (131/(131+2)) |
| **Specificity** (TP/(TP+FP)) | **97%** (132/(132+4)) | **99%** (131/(131+1)) |

**False Negative Mutation Discovered**: One FN mutation, which was covered in the ROI but missed by TSACP, was identified when the results from TSACP assay were compared with those from the Ion Torrent AmpliSeq assay. The mutation was not detected because of the inefficient amplification of the amplicon due to a SNP at the 3' end of the primer region (**Figure 6**).

**Figure 6. False negative variant missed in TSACP assay**

## Conclusions

- Our validation study demonstrated that the Illumina TruSeq® Amplicon Cancer Panel combined with MolecularMD's proprietary analysis pipeline provides a highly specific and sensitive test suitable for screening patient FFPE tumor specimens for a spectrum of clinically relevant somatic mutations with analytical sensitivity of 5% minor allele frequency.
- With more than 95% ROI sequenced at least 0.2x normalized coverage, the analytical reproducibility was above 98%.
- In addition to reporting hotspot mutations, we also report other variants that would impact protein sequence. We report all mutations in our ROI, except for those characterized as synonymous or intronic.
- The diagnostic sensitivity and specificity for FFPE samples, under these validation conditions, are 98% and 99% respectively when a stringent coverage filter is applied during data analysis.
- In this study, we developed a unique cross-validation strategy using two different NGS platforms. We also established a custom Ion TargetSeq® assay for confirming variants at frequencies below 10%. In total, 98% of all the mutations identified in FFPE samples were cross-examined by other platforms, i.e., the Ion Torrent AmpliSeq Cancer Panel, the custom Ion Torrent TargetSeq, or Sanger sequencing.

**For Further Information** please contact info@molecularmd.com or visit www.molecularmd.com.