Complementary Analysis of LKB1/STK11 Mutation and Protein Expression Status using Next-Generation Sequencing, Sanger Sequencing and Immunohistochemistry

AACR 2013
Abstract #3509

Kimberly Pelak, Jennifer Wright, Zhenyu Yan, Agus Darwanto, Weihua Liu, Peng Fang, Jin Li, Sabita Sankar and Chad Galderisi. MolecularMD, Portland, OR and Cambridge, MA

Introduction

LKB1 (liver kinase B1)/STK11 (serine-threonine kinase 11) is a tumor suppressor that encodes a serine/threonine kinase which negatively regulates the mTOR (mammalian target of rapamycin) signaling pathway. Somatic mutations in LKB1 occur most notably in lung adenocarcinoma, cervical cancers and melanoma. Inactivation of LKB1 is caused by point mutations, homozygous deletions or promoter methylation. LKB1 mutation status is a predictive marker for responsiveness to both MEK and PI3K inhibitors (ref. 1, 2). We have developed next-generation sequencing (NGS), Sanger sequencing and immunohistochemistry (IHC) assays to assess LKB1 mutation and protein expression status in a set of cell line samples and patient FFPE specimens. These assays provide complementary information about the range of mutations in the gene, as well as the impact that these may have on the protein.

Materials and Methods

Samples:

NGS: NA12878; 9 cancer cell lines – 5 with LKB1 mutations, 4 without LKB1 mutations (**Table 1**); dilutions of A549 and DU-145; 35 FFPE samples of unknown LKB1 mutation status that met our QC metrics

Sanger: 34 FFPE samples of unknown LKB1 mutation status (1 of 35 had insufficient DNA for Sanger)

IHC: A2058 melanoma cell line with LKB1 mutation; SKMel24 melanoma cell line without LKB1 mutation; 33 FFPE samples

NGS, alignment and variant calling: We used the Ion AmpliSeq Designer to design primers that targeted the coding region of LKB1, 10bp on the edge of each intron, and 700bp upstream of the gene. The IonAmpliSeq protocol was used to prepare and quantify the sample libraries, using just 20ng input DNA. Quantified libraries were amplified on the OneTouch system using the OneTouch template preparation v2 kit, and were sequenced on the PGM using the Ion PGM 200 bp sequencing kit and a 316 chip. Alignment was performed by Torrent Suite 3.4 and variants were called by the Torrent Suite 3.4 variant caller with modified parameters to allow optimal sensitivity and specificity. Only variants with a frequency of at least 2% are reported. Variants are called throughout the ROI, but our LKB1 Ion AmpliSeq targeted sequencing assay, and thus this poster, focus on the coding variants in LKB1, and variants that fall within 10bp of the exons (splice region). The identified variants are then filtered using MolecularMD's proprietary analysis and filtering pipeline.

Validation of NGS results: We sequenced the coding exons of LKB1 by Sanger sequencing. This allowed for confirmation of mutations with frequencies of greater than ~10% (true positives). It also allowed us to check whether any variants with a frequency of greater than ~10% were missed in the next-generation sequencing results (false negatives). Variants with frequencies less than 10% were validated by a restriction endonuclease digest method, as described in ref. 3. The resulting products were read on the Caliper Gx to quantify the portion of fragments that contained the mutant.

IHC: We optimized and validated an IHC assay for LKB1 using a concentrated preparation of anti-LKB1 antibody purchased from Cell Signaling Technologies. The assay was automated using the Ventana Benchmark. Initial assay validation was performed using formalin-fixed paraffin-embedded cell pellets of A2058 (LKB1 negative) and SKMel24 (LKB1 positive) cell lines (**Figure 1**). The assay was validated using 65 normal and FFPE tissues. Additionally, IHC was performed on the 33 FFPE tissues to compare IHC and NGS data. IHC expression scoring delineated below reflects both staining intensity and fraction of positive cells, taking into account both nuclear and cytoplasmic staining.

IHC scoring system:

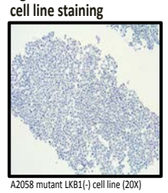
0 = no appreciable staining
1+ = faint/barely visible staining
2+ = weak to moderate staining
3+ = strong staining

A score of 0 and 1+ is a negative result; a score of 2+ is considered borderline or equivocal; a score of 3+ is a positive result.

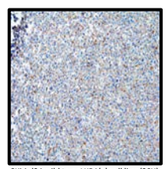
Table 1. LKB1 mutations in cell lines

Cell line Name	Cancer type	LKB1 mutations by cDNA change	Mutation by Protein Change	Known mutation status?
A549	Lung	c.109C>T	p.Q37*	Y
DU-145	Prostate	c.532_536delAAGCC	p.K178fs*86	Y
NCIH1395	Lung	c.165delG	p.E57fs*7	Y
G361	Skin	c.842delC	p.P281fs*6	Y
NCIH1975	Lung	None	-	Y
RKO	Colon	None	-	Y
MDAMB231	Breast	None	-	Y
A2058	Skin	c.145T>G	p.Y49D	N
SKMel24	Skin	None	-	N

Figure 1. LKB1 IHC control cell line staining



A2058 mutant LKB1(-) cell line (20X)



SKMel24 wild type LKB1(+) cell line (20X)

Results

Cell line sequencing: We used our LKB1 NGS targeted sequencing panel to sequence 9 cell lines (**Table 1**). Four of these were known to have LKB1 mutations, and all four of these mutations were identified in the NGS results, including two 1bp indels in homopolymer regions. Another 3 cell lines were known to have wild type LKB1, and no mutations were identified in these samples in the NGS results. Two cell lines had unknown LKB1 mutational status; one of these had an LKB1 variant, and the other had wild type LKB1.

Cell line dilutions and analytical sensitivity: An advantage of NGS is that it can detect low level mutations with about 5% sensitivity. To assess the limit of detection (LOD) of the LKB1 targeted sequencing panel, DNA from the cancer cell lines A549 and DU-145 was diluted into wild type NA12878. A549 has a known single-base change in LKB1 (c.109C>T), and DU-145 has a known 5bp deletion (c.532_536delAAGCC) in LKB1. We sequenced samples that were comprised of 50%, 20%, 10%, 5%, 2.5% and 1% cancer cell line DNA. Both the 5bp deletion in DU-145 and the SBS in A5459 showed the expected drop in frequency as the amount of DNA decreased, and the mutation frequencies that were called by Torrent Suite 3.4 approximately agreed with the expected mutation frequencies (**Table 2**). The LOD was in the range of 2.5% - 5%. Dilutions of DU-145 and A549 were done manually for each sequencing run, so slight deviations from the expected frequency are not surprising. The "IGV freq" column lists the percent of raw reads that support a variant, and was manually calculated by inspecting the reads in the Integrative Genomic Viewer (IGV) (ref. 4).

Table 2. Cell line dilutions: DU-145 and A549

	DU-145 19:1220438-1220442_c.531_535delCAAGC 5bp deletion			A549 19:1207021_c.109C>T C→T change		
Dilution	Mutation frequency (%)	IGV freq	Read Depth	Mutation frequency (%)	IGV freq	Read Depth
100% cell line	97.8	99.2	6050	99.8	99.7	2380
50% cell line	62.3	64.6	7905	42.8	42.2	2082
20% cell line	22.0	25.1	1812	16.0	15.7	2325
10% cell line	14.3	13.8	1367	8.1	8.0	5013
5% cell line	6.9	7.6	3590	4.0	4.0	4121
2.5% cell line	2.6	2.6	886	-	1.8	3229
1% cell line	-	1.0	2781	-	0.9	8903

Coverage: Another advantage to NGS is that it is possible to discover novel variants that occur at much lower frequencies than was possible by any previously existing method. As the coverage increases, the ability to identify low frequency variants also increases. In this assay, we targeted at least 500x read depth across our ROI, and we generally saw read depths >1000x. This gave us sufficient coverage to comfortably call mutations that occurred with frequencies as low as about 5%. The coverage across the LKB1 ROI was very consistent. Regions that showed sufficient coverage (>500x) were highly repeatable, as were the small number of regions that did not generally achieve sufficient coverage. **Figure 2** shows the average coverage across LKB1 in 29 FFPE samples. The vast majority (96%) of the coding region for LKB1 was efficiently targeted and captured by our assay with coverage >500x. There is a small section of Exon 3 that was not sufficiently covered (29bp of the coding region, plus the first 10bp of the intron). There is one variant in COSMIC in this region that would affect the protein (p.Q152* (c.454C>T); lung cancer) (ref. 5). There are a total of 175 coding substitutions/small indels in LKB1 in COSMIC, and our assay would identify 99.4% of these variants.

Blinded FFPE sequencing: We used our LKB1 targeted sequencing panel to sequence 35 FFPE samples of unknown LKB1 mutational status. We identified a total of 18 variants in these samples. Eight of these variants were unique, and ranged in frequency from 6% to 52% (**Table 3**). The variant 19:1207162_c.250A>T (rs137853076) has been reported as a germ line variant that can lead to Peutz-Jeghers syndrome. The other 10 variants are a shared germ line variant in the splice region (in an intron, 7bp away from exon 7) that is present in 10 samples. All of these variants were manually inspected in IGV and were determined to be likely true positive variants. None of the variants are listed in COSMIC. The FFPE samples in **Table 4** did not have any coding or splice variants (SBS or small indel) with a frequency of greater than 5% in the LKB1 NGS results.

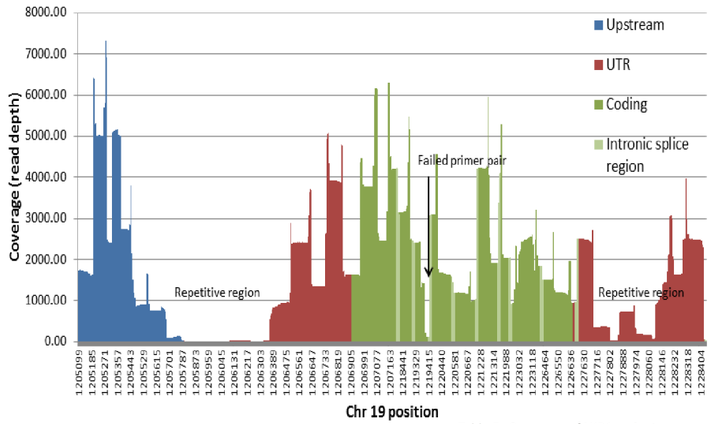
Table 3. IHC scores for FFPE samples with NGS variants

Sample	Tissue	Variant	Exon	Mutation frequency (%)	Amino acid change	Type of Mutation	Read Depth	Confirmation of NGS result	IHC Score
MMD-2-2-027	Lung	19:1222012_c.920+7G>C	7	44.8	-	Splice region	2770	Sanger	3+
MMD-2-2-028	Lung	19:1207162_c.250A>T	1	47.8	p.Lys84X	Nonsense	2978	Sanger	0
		19:1222012_c.920+7G>C	7	99.9	-	Splice region	1549	Sanger	
MMD-2-2-031	Lung	19:1221990_c.905A>C	7	51.7	p.Gln302Pro	Missense	2410	Sanger	0/1+
		19:1222012_c.920+7G>C	7	75.1	-	Splice region	2457	Sanger	
MMD-2-2-032	Lung	19:1222012_c.920+7G>C	7	58.8	-	Splice region	3063	Sanger	3+
MMD-2-2-033	Lung	19:1222012_c.920+7G>C	7	48	-	Splice region	2660	Sanger	3+
MMD-2-3-051	Uterus	19:1207024_c.112C>T	1	39.5	p.Pro38Ser	Missense	8504	Sanger	3+
MMD-2-3-052	Endometrium	19:1226561_c.1217C>T	9	6.1	p.Ala406Val	Missense	880	RE/Caliper	2+
MMD-2-3-063	Ovary	19:1220597_c.615G>A	5	21.3	p.205A	Synonymous	1230	Sanger	3+
MMD-2-3-064	Ovary	19:1222012_c.920+7G>C	7	34.2	-	Splice region	984	Sanger	0
MMD-2-3-070	Cervix	19:1207169_c.257G>C	1	10.6	p.Arg86Pro	Missense	2466	Sanger	0
		19:1222012_c.920+7G>C	7	25.6	-	Splice region	2954	Sanger	
MMD-2-3-071	Cervix	19:1223042_c.979G>A	8	23.7	p.Asp327Asn	Missense	2270	Sanger	3+
		19:1222012_c.920+7G>C	7	50.6	-	Splice region	2334	Sanger	
MMD-2-4-021	Colon	19:1226495_c.1151G>A	9	52.5	p.Arg384Gln	Missense	1475	Sanger	3+
		19:1222012_c.920+7G>C	7	47.8	-	Splice region	2503	Sanger	
MMD-2-5-007	Normal Pancreas	19:1222012_c.920+7G>C	7	52.9	-	Splice region	942	Sanger	3+

Table 4. IHC scores for NGS wild type FFPE samples

Sample	Tissue	IHC Score
MMD-1-1-006	Normal Skin	-
MMD-2-1-016	Skin	3+
MMD-2-1-017	Skin	3+
MMD-2-1-018	Skin	3+
MMD-2-1-019	Skin	3+
MMD-2-1-020	Skin	3+
MMD-2-2-025	Lung	3+
MMD-2-2-026	Lung	1+
MMD-2-2-030	Lung	3+
MMD-2-2-034	Lung	2+
MMD-2-2-035	Lung	2+
MMD-2-3-053	Uterus	3+
MMD-2-3-054	Endometrium	3+
MMD-2-3-058	Endometrium	2+
MMD-2-3-061	Endometrium	3+
MMD-2-3-062	Endometrium	2+
MMD-2-3-073	Breast	0
MMD-2-3-074	Breast	-
MMD-2-4-020	Colon	3+
MMD-2-5-007	Pancreas	3+
MMD-2-5-008	Pancreas	1+/2+
MMD-2-5-009	Pancreas	3+

Figure 2. Average coverage across 29 FFPE libraries



Validation of NGS results:

Sanger sequencing: The 9 coding exons of LKB1 in 34 of the 35 FFPE samples were Sanger sequenced. The final sample did not have sufficient DNA for Sanger sequencing. Sanger sequence traces for each sample were manually inspected, with a focus on the coding regions and the sequence immediately adjacent to the coding regions. Variants that were identified by NGS were compared with the variants identified by Sanger sequencing. *All of the NGS variants with frequencies >10% were confirmed by the Sanger sequence data.* None of the NGS variants with frequencies greater than 10% were false positive variant calls. We could also use the Sanger sequence data to cross-validate our variant calls from the NGS data. We did not identify any additional variants in the Sanger sequence data that we had not previously identified in the NGS data. This means that we had 100% specificity to identify variants with frequencies of at least 10-20% (i.e., above the Sanger LOD).

RE digest/Caliper Gx: A mutation was identified by NGS that had a frequency below 10% (p.Ala406Val), and therefore could not be confirmed by Sanger sequencing. This variant was confirmed by a restriction endonuclease-based fragment analysis method.

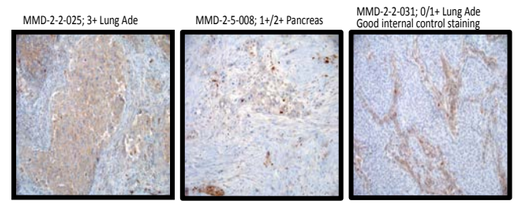
Table 5. Summary of LKB1 variants identified by NGS

Variant Type	FFPE Specimens	Cell Lines	Combined Total
Frameshift	0	3	3
Missense	6	1	7
Nonsense	1	1	2
Splice region	10	3	13
Synonymous	1	0	1
Total	**18**	**8**	**26**

Immunohistochemistry: Thirty-three FFPE tissue specimens analyzed by NGS were evaluated for LKB1 protein expression by IHC. **Figure 3** shows representative images of IHC scores ranging from 0 to 3+. Interestingly, there was a specimen without an NGS variant (WT mutation status) but with a negative IHC score.

NGS and IHC comparison: The FFPE sample with a nonsense variant in LKB1, as reported by the NGS results, had an IHC score of 0, as would be expected. The FFPE sample with a synonymous variant in LKB1 had an IHC score of 3+, as would be expected. The 6 FFPE samples with missense variants had IHC scores ranging from 0 to 3+, as would be expected, since some missense changes would have a larger effect on the protein than others. The samples without NGS variants generally had an IHC score of 2+ or 3+. The samples with no missense/nonsense NGS as well as splice region variants but an IHC score of 0 or 1+ are being investigated for possible copy loss.

Figure 3. Representative IHC Images (20X)



Conclusions

- NGS and IHC can provide complementary information about the LKB1 mutational status.
- The LKB1 targeted sequencing panel combined with MolecularMD's proprietary analysis pipeline provides a robust and accurate test to identify clinically relevant mutations (SBS, indels) with frequencies as low as ~5%. Our targeted panel provides a comprehensive analysis of LKB1 mutations throughout the 9 coding exons of LKB1, whereas the Ion AmpliSeq cancer panel only covers parts of 4 exons.
- We identified 18 mutations in LKB1 across 35 FFPE samples, all of which were confirmed by Sanger sequencing or a restriction endonuclease digest method (**Table 5**).
- We identified 8 mutations in cell lines (**Table 5**).
- Both the sensitivity and specificity of the NGS assay were 100% (no false positives or false negatives).
- The LKB1 targeted sequencing panel is well-suited for small clinical samples, since it only uses 20ng DNA input.
- We are currently investigating ways to use NGS to call copy number changes in LKB1.
- We have developed and validated an LKB1 IHC assay using the Ventana Benchmark to assess LKB1 status in FFPE specimens.
- Our data highlight the need for complementary IHC and NGS data to provide more comprehensive information about LKB1 status.

References

1. Chen Z, Cheng K, Walon Z, et. al., *Nature* 2012.
2. Liu W, Monahan KB, Pfefferle AD, et. al., *Cancer Cell* 2012.
3. Liu WH, Kaur M, Makrigiorgos GM. *Hum. Mutat.* 2003.
4. http://www.broadinstitute.org/igv/
5. Bamford S, Dawson E, Forbes S, et. A., *Br J Cancer.* 2004.

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

MolecularMD