Microbial Signatures and Genomic Profiling in Tumor Samples using Next Generation Sequencing.

Mei Qi Yee[1], Yik Lim Kok[1] Lina Nur Saleh[1], Tong Zhang[2], Pramila Ariyaratne[1], Yingnan Yu[2], Olivia Scully[2], Darwin Tay[2], Weng Kit Lye[1], Shi Hui Ang[1], Sirin Lee[1], Trang Nguyen[1], Claire Tang[1], Belinda Ling[1], Thaddaeus Ong[1], Harry Suhardi[1], Min Ko Ko Aye[1], Aung Win Kyaw[1], Jason Fischer[3], Eugene Wee[1], Ludovic Lacroix[4] & Charlie Lee[1,2]

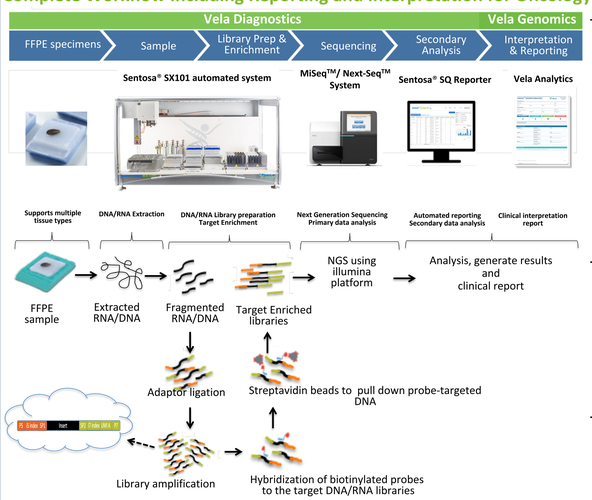
1 Vela Research, Vela Diagnostics Pte Ltd, Singapore, 2 Vela Genomics, Vela Diagnostics Pte Ltd, Singapore, 3 Vela Diagnostics Inc, USA, 4 Institut Gustave Roussy, France

Introduction

Precision medicine guides treatment options based on patients' genomic information. The advent of next-generation sequencing (NGS) has enabled the rapid acquisition of accurate genomic data. We have developed and validated a complete workflow (OncoKey SL Plus) comprising DNA and RNA extraction from the same formalin-fixed paraffin-embedded (FFPE) samples, library preparation, target enrichment, and high-throughput sequencing. Sequencing data is subsequently analyzed with an in-house bioinformatics pipeline to detect the presence of oncogenic bacteria and viruses, as well as an essential (OncoKey SL 60 Plus) or comprehensive (OncoKey SL 525 Plus) set of host's DNA mutations and RNA fusions associated with oncogenesis. Our workflows can be performed manually and have been optimised on the automated liquid handler Sentosa® SX 101 to prevent cross-contamination, with a 5 day turnaround time from 16 FFPE samples to lab report generation.

Materials and Methods

Complete Workflow Including Reporting and Interpretation for Oncology



Key features of OncoKey SL60 Plus & OncoKey SL525 Plus Panel

- ✓ Automation of 16 FFPE samples from sample preparation to results with 5 hours hands-on time and 5 day of turnaround time. (Including secondary analysis and interpretation with Sentosa SQ Reporter and Vela Analytics softwares, respectively.)
- ✓ Employs Unique Dual Indexes (UDI) and 10bp Unique Molecular Identifier (UMI) to enhance sensitivity and accuracy.
- ✓ Limit of detection: 40ng of DNA and RNA
- ✓ Detect SNP, Indels, CNV, Splice variants, Fusions, Microsatellite Instability (MSI) and Oncogenic pathogens
- ✓ Provide Tumor Mutational Burden value (OncoKey SL 525 Plus Panel)

- Standard reference materials used.

Customized FFPE reference standards	Horizon -HD-C331B
FFPE Tumor Fusion RNA reference materials	SeraSeq-0710-0496
Hepatitis B Virus (HBV) in PLC/PRF/5 cell line	ATCC-CRL-8024
Human papillomavirus (HPV)	Seracare-2400-0161
Merkel Cell Polyomavirus (MCPyV) in MKL-1 cell line	Sigma-9111801
HHV8 in BC-1 cell line	ATCC-CRL-2230
EBV in Raji cell line	ATCC-CCL-86
HCV Genotype 1a	Boca Biolistics

- 16 Horizon reference standards were used in the validation of the automated OncoKey SL assays. Sequencing was carried out using illumina MiSeq™/ NextSeq™ systems.
- 16 SeraSeq® reference materials were used to assess the effectiveness of the assay in identifying splice variants and fusions. The success rate of detection was determined.
- Clinical samples, HBV, HPV, MCPyV, HHV8, EBV and HCV samples were used to assess the effectiveness of the assay to detect Oncoviruses.

Results

The tables below showed the validation results of the automated assays. Each run using 16 FFPE reference standards with LoD 40ng were performed. Data are representative of 10 repeated runs and showed a 100% pass in QC criteria for 8 DNA samples (left) and 8 RNA samples (right) in both assays.

OncoKey SL60 Plus		OncoKey SL525 Plus	
QC Performance	DNA Sample Barcode D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8	QC Performance	DNA Sample Barcode D9 | D10 | D11 | D12 | D13 | D14 | D15 | D16
Mean Sample Coverage	689.89 - 851.01 (>200x)	Mean Sample Coverage	937.25 - 1269.22 (>200x)
Coverage Uniformity	99.22% - 99.46% (>90% targets with >50x coverage)	Coverage Uniformity	99.07% - 99.13% (>90% targets with >50x coverage)
Base Quality	89.66% - 91.41% (>80% bases with >30 quality score)	Base Quality	87.86% - 89.33% (>80% bases with >30 quality score)
Pass-Filter Reads	91,476 - 92,086 (>80% reads)	Pass-Filter Reads	100% (>80% reads)
Control Fragment Reads	87894 - 102666 (>1000 reads)	Control Fragment Reads	93858 - 126048 (>1000 reads)
Pass/Fail	All Pass	Pass/Fail	All Pass
Total Pass/Run	(100%, 95% confidence interval)	Total Pass/Run	(100%, 95% confidence interval)

Table 1 - OncoKey SL 60 Plus & OncoKey SL 525 Plus Quality Control Performance using Horizon DNA reference standards.

OncoKey SL60 Plus		OncoKey SL525 Plus	
QC Performance	RNA Sample Barcode R121 | R122 | R123 | R124 | R125 | R126 | R127 | R128	QC Performance	RNA Sample Barcode R9 | R10 | R11 | R12 | R13 | R14 | R15 | R16
Read Counts	616123 - 921262 (>10000 reads)	Read Counts	8155110 - 11182074 (>10000 reads)
Base Quality	87.43% - 90.50% (>80% bases with >30 quality score)	Base Quality	89.39% - 90.51% (>80% bases with >30 quality score)
Pass-Filter Reads	90.82% - 91.89% (>80% reads)	Pass-Filter Reads	100% (>80% reads)
Control Fragment Reads	9205 - 11780 (>1000 reads)	Control Fragment Reads	62190 - 94539 (>1000 reads)
Pass/Fail	All Pass	Housekeep Gene -ACTB	179779 - 336724 (>1000 reads)
Total Pass/Run	(100%, 95% confidence interval)	Housekeep Gene -GAPDH	94359 - 183553 (>1000 reads)
		Pass/Fail	All Pass
		Total Pass/Run	(100%, 95% confidence interval)

Table 2 - OncoKey SL 60 Plus & OncoKey SL 525 Plus Quality Control Performance using SeraSeq® Fusion RNA reference materials.

Results demonstrated that the mutation detection sensitivity at 5% variant frequency for OncoKey SL60 Plus and SL525 Plus were 100% and 99%, respectively (left, using DNA reference standards). The mutation detection sensitivity at 5% variant frequency for OncoKey SL60 Plus and SL525 Plus were 96.4% and 97.9%, respectively (right, using RNA reference materials).

OncoKey SL60 Plus				OncoKey SL525 Plus			
Gene	Expected Mutation	Expected Allelic Frequency (%)	Allelic Frequency for each DNA Sample Barcode D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8	Gene	Expected Mutation	Expected Allelic Frequency (%)	Allelic Frequency for each DNA Sample Barcode D9 | D10 | D11 | D12 | D13 | D14 | D15 | D16
BRAF	V600E	7.7	4.08 - 8.87	BRAF	V600E	7.7	3.72 - 7.36
	V600M*	3.8	8.42 - 14.91		V600M*	3.8	6.06 - 9.69
EGFR	G719S	20.0	16.32 - 23.13	EGFR	G719S	20.0	15 - 18.68
	delT746-A750*	4.0	2.79 - 7.93		delT746-A750*	4.0	Not detected
GNA11	Q209L	5.0	4.81 - 6.60	GNA11	Q209L	5.0	4.21 - 7.87
	A166T	5.0	7.36 - 17.62		A166T	5.0	5.39 - 14.56
KRAS	G12S	5.0	3.60 - 7.08	KRAS	G12S	5.0	2.56 - 8.51
	G13D	20.0	15.27 - 23.31		G13D	20.0	11.91 - 18.76
NRAS	Q61H	5.0	2.79 - 9.17	NRAS	Q61H	5.0	2.51 - 6.71 (not detectable in D11)
PIK3CA	E542K	5.0	5.81 - 10.34	PIK3CA	E542K	5.0	2.94 - 8.16
	H1047R	25.0	17.44 - 28.34		H1047R	25.0	19.39 - 26.58
	E545K	5.0	3.95 - 9.29		E545K	5.0	4.35 - 6.66
KIT	D816V	5.0	8.85 - 20.00	KIT	D816V	5.0	6.99 - 16.64
GNAQ	Q209L	5.0	6.44 - 13.65	GNAQ	Q209L	5.0	4.27 - 8.83
Total Mutations (AF≥5%) Detected per Sample (excluding V600M* and delT746-A750*)		12	12 | 12 | 12 | 12 | 12 | 12 | 12 | 12	Total Mutations (AF≥5%) Detected per Sample (excluding V600M* and delT746-A750*)		12	12 | 12 | 11 | 12 | 12 | 12 | 12 | 12
Total Mutations (AF≥5%) Detected per Run		96	96/96 (100%, 95% confidence interval)	Total Mutations (AF≥5%) Detected per Run		96	95/96 (99%, 95% confidence interval)

Table 3 - The mutation detection efficiency of OncoKey SL 60 Plus & OncoKey SL 525 Plus using Horizon DNA reference standards.
Read count with ≤1 are indicated as non detectable.
* These samples with mutations have an expected allelic frequency of <5%, thus, they are not included in the acceptance criteria. Nonetheless, both OncoKey SL60 Plus and OncoKey SL525 Plus assay detected them.

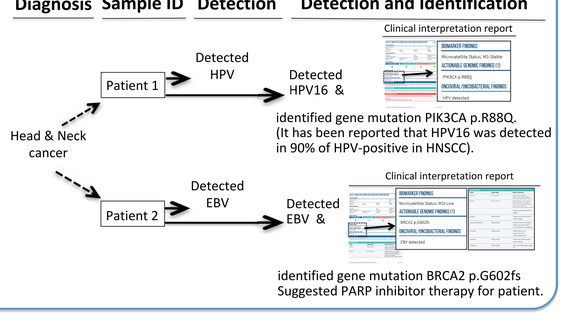
OncoKey SL60 Plus								OncoKey SL525 Plus							
Expected RNA Fusion	Supporting RNA Fusion Read Count for each RNA Sample Barcode R121 | R122 | R123 | R124 | R125 | R126 | R127 | R128							Expected RNA Fusion	Supporting RNA Fusion Read Count for each RNA Sample Barcode R9 | R10 | R11 | R12 | R13 | R14 | R15 | R16						
CCDC6-RET	2 - 6							CCDC6-RET	31 - 649						
CD74-ROS1	3 - 29							CD74-ROS1	45 - 448						
EGFR variant II*	Not applicable							EGFR variant II*	All detected (except for R103)						
EGFR-SEPT14*	Not applicable							EGFR-SEPT14*	12 - 70						
EML4-ALK	2 - 12							EML4-ALK	93 - 205						
ETV6-NTRK3	5 - 22							ETV6-NTRK3	244 - 1341						
FGFR3-BAIAP2L1	5 - 17							FGFR3-BAIAP2L1	204 - 610						
FGFR3-TACC3	9 - 42							FGFR3-TACC3	1033 - 3027						
KIF5B-RET	3 - 12							KIF5B-RET	528 - 2522						
LMNA-NTRK1	3 - 4 (not detectable in R128)							LMNA-NTRK1	82 - 357						
MET Exon 14 Skipping	18 - 99							MET Exon 14 Skipping	All detected						
NCOA4-RET	3 - 12							NCOA4-RET	18 - 392 (not detected in R97)						
PAX8-PPARG1*	Not applicable							PAX8-PPARG1*	132 - 1206						
SLC34A2-ROS1	3 - 9 (not detectable in R128)							SLC34A2-ROS1	57 - 145						
SLC45A3-BRAF	2 - 26 (not detectable in R125)							SLC45A3-BRAF	2 - 228						
TPR-NTRK1	2 - 8 (not detectable in R128)							TPR-NTRK1	63 - 233						
TPM3#ROS1-ERG#	Not applicable							TPM3#ROS1-ERG#	2 - 90 (not detected in R102)						
TPM3-NTRK1	3 - 18							TPM3-NTRK1	44 - 527						
Total RNA Fusions Detected per Sample	14	14	14	14	13	14	14	11	Total RNA Fusions Detected per Sample	17	18	18	18	17	17	18	
Total RNA Fusions Detected per Run	100/112 (96.4%, 95% confidence interval)								Total RNA Fusions Detected per Run	141/144 (97.9%, 95% confidence interval)							

Table 4 - The mutation detection efficiency of OncoKey SL 60 Plus & OncoKey SL 525 Plus using SeraSeq® Fusion RNA reference materials.
Read count with ≤1 are indicated as non detectable
These RNA fusions are not targeted in the OncoKey SL 60 Plus assay.
Probes targeting this specific RNA fusion were not included in the particular run.

Results showed assay ability to detect oncogenic viruses- HBV, MCPyV, HPV51, HPV18 & HPV16. Previously the presence and the copy number of the extracted virus was determined using ddPCR.

Virus	Detected
HBV	✓
MCPyV	✓
HPV16,18,51	✓
HHV8	✓
EBV	✓
HCV	✓

32 FFPE clinical samples were analysed and HPV16 and EBV were identified in 2 head and neck samples. The presence of HPV and EBV was confirmed using immunohistochemistry (provided by Institut Gustave Roussy). OncoKey specifically identified different gene mutation PIK3CA p.R88Q and BRCA2 p.G602fs in the samples and provided detailed information on targeted therapy.



Conclusion.

Identifying virus and bacteria as etiological agents of specific cancers aids our understanding of the interplay between the microbiome and cancer, diagnosis and treatment. Robust automation coupled with UMI-based hybridisation capture workflow and NGS, we demonstrated that OncoKey SL60 and SL525 Plus assays can accurately identify over 60 clinically relevant DNA and RNA variants in FFPE samples, as well as the microbial signatures in various tumor types. Hence, we introduced a cost-effective strategy that enables high-throughput and rapid detection with accuracy and precision for comprehensive microbial and genomic profiling.