Validation of an Automated Method for Library Preparation for a Next-Generation Sequencing-Based Assay for Oncology

Catherine I. Dumur, Paula Anderson, M. Fernanda Sábato, Celeste N. Powers, Andrea Ferreira-Gonzalez

Department of Pathology, Virginia Commonwealth University, Richmond, VA

BACKGROUND

Targeted Next Generation Sequencing (NGS) technology is rapidly being adopted to assess the mutational status of multiple genes on formalin-fixed, paraffin-embedded (FFPE) tumor specimens in clinical settings

Library preparation is a critical, hands-on and time-consuming step in the NGS workflow. During library preparation, each library is prepared in an independent well of a 96-well plate, encompassing several washes and magnetic bead-binding steps.

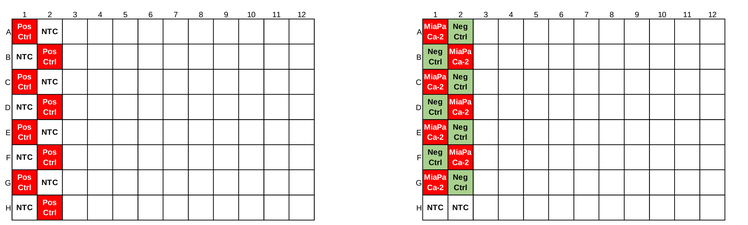
This format increases the number of technical hours as more samples/libraries are prepared, while increasing the risk of human-introduced error. Automation and scalability of library preparation is much needed to not only reduce these issues, but to allow for the laboratory to increase the sample throughput.

Here, we present the validation and implementation of an open liquid handling platform, the VERSA™ 1100 GENE (Aurora Biomed, Vancouver, BC) for medium to high-throughput library preparation for routine utilization with the Ion AmpliSeq™ Cancer Hotspot Panel v2 (CHP2) assay on FFPE clinical specimens, including FFPE Quality Control (QC) materials (1).

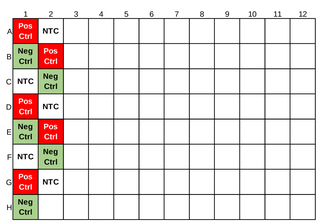
MATERIALS AND METHODS

Figure 1. Experimental Design

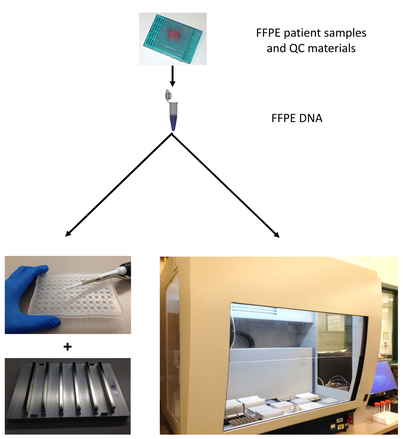
A – Cross-contamination (Checkerboard Experiments)



B – Reproducibility



C – Accuracy



Ion AmpliSeq™ Cancer Hotspot Panel v2 on PGM

RESULTS

Figure 2. Cross-contamination (Checkerboard Experiments)

A

Well	Sample ID	Library Conc. (ng/mL)	Well	Sample ID	Library Conc. (ng/mL)
1A	Pos Ctrl	196	2A	NTC	0
1B	NTC	0	2B	Pos Ctrl	782
1C	Pos Ctrl	318	2C	NTC	0
1D	NTC	0	2D	Pos Ctrl	598
1E	Pos Ctrl	280	2E	NTC	0
1F	NTC	0	2F	Pos Ctrl	572
1G	Pos Ctrl	288	2G	NTC	0
1H	NTC	0	2H	Pos Ctrl	582

Well	Sample ID	Library Conc. (ng/mL)	Well	Sample ID	Library Conc. (ng/mL)
1A	MiaPaCa-2	1356	2A	Neg Ctrl	2220
1B	Neg Ctrl	1872	2B	MiaPaCa-2	1284
1C	MiaPaCa-2	1062	2C	Neg Ctrl	1756
1D	Neg Ctrl	1684	2D	MiaPaCa-2	728
1E	MiaPaCa-2	906	2E	Neg Ctrl	1498
1F	Neg Ctrl	1218	2F	MiaPaCa-2	590
1G	MiaPaCa-2	1144	2G	Neg Ctrl	1464
1H	NTC	0	2H	NTC	0

B

MiaPaCa-2

Gene ID	CDS_mut_syntax	AA_mut_syntax	Chrom	hg19 Position	Ref	Variant	Frequency	Quality	Coverage	Allele Cov	Strand Bias
APC	Not a HotSpot	Not a HotSpot	chr5	112175770	G	A	74.7	19084.4	1997	1491	0.50
FGFR3	Not a HotSpot	Not a HotSpot	chr4	1807894	G	A	100.0	11210.6	699	699	0.50
FLT3	Not a HotSpot	Not a HotSpot	chr13	28610183	A	G	64.0	14997.7	2000	1281	0.52
HRAS	c.81T>C	p.His27His	chr11	534242	A	G	53.1	5543.6	993	527	0.51
KDR	Not a HotSpot	Not a HotSpot	chr4	55980239	C	T	100.0	7165.1	455	455	0.50
KRAS	c.34G>T	p.Gly12Cys	chr12	25398285	C	A	100.0	31607.9	1986	1986	0.50
MET	Not a HotSpot	Not a HotSpot	chr7	116339672	C	T	68.3	16596.2	1999	1366	0.51
NOTCH1	Not a HotSpot	Not a HotSpot	chr9	139390822	G	C	100.0	21173.3	1325	1325	0.50
PDGFRA	Not a HotSpot	Not a HotSpot	chr4	55141055	A	G	100.0	14758.1	924	924	0.50
RET	Not a HotSpot	Not a HotSpot	chr10	43613843	G	T	67.6	15035.5	1846	1247	0.50
RET	Not a HotSpot	Not a HotSpot	chr10	43615633	C	G	65.9	12191.1	1564	1030	0.52
STK11	Not a HotSpot	Not a HotSpot	chr19	1220321	T	C	67.2	8842.7	1094	735	0.51
TP53	c.742C>T	p.Arg248Trp	chr17	7577539	G	A	100.0	31556.4	1983	1983	0.50

Negative Control

Gene ID	CDS_mut_syntax	AA_mut_syntax	Chrom	hg19 Position	Ref	Variant	Frequency	Quality	Coverage	Allele Cov	Strand Bias
APC	Not a HotSpot	Not a HotSpot	chr5	112175770	G	A	48.8	9248.7	1890	923	0.51
ATM	c.2572T>C	p.Phe858Leu	chr11	108138003	T	C	51.4	10550.2	1996	1025	0.50
FGFR3	Not a HotSpot	Not a HotSpot	chr4	1807894	G	A	100.0	16809.8	1048	1048	0.50
FLT3	Not a HotSpot	Not a HotSpot	chr13	28602292	T	C	47.6	9375.2	1996	950	0.52
FLT3	Not a HotSpot	Not a HotSpot	chr13	28610183	A	G	100.0	32079.3	2000	2000	0.50
HRAS	c.81T>C	p.His27His	chr11	534242	A	G	47.9	5761.1	1220	584	0.50
KDR	Not a HotSpot	Not a HotSpot	chr4	55972974	T	A	50.3	7480.7	1465	737	0.51
KDR	Not a HotSpot	Not a HotSpot	chr4	55980239	C	T	100.0	9211.1	580	580	0.50
PDGFRA	Not a HotSpot	Not a HotSpot	chr4	55141055	A	G	100.0	19542.4	1226	1226	0.50
RET	Not a HotSpot	Not a HotSpot	chr10	43613843	G	T	100.0	30332.6	1892	1892	0.50
STK11	Not a HotSpot	Not a HotSpot	chr19	1220321	T	C	51.0	4307.5	828	422	0.53
STK11	c.1062C>G	p.Phe354Leu	chr19	1223125	C	G	49.5	8451.3	1704	843	0.51
TP53	Not a HotSpot	Not a HotSpot	chr17	7578210	T	C	51.3	10551.7	1999	1026	0.51
TP53	Not a HotSpot	Not a HotSpot	chr17	7579472	G	C	92.2	18215.1	1337	1233	0.52

Fig. 2: **Checkerboard library preparations**
A- Library concentrations measured by the Qubit dsDNA HS Assay for the two checkerboard experiments. **B-** Representative variants called for the KRAS homozygous mutant pancreatic cancer-derived cell line, MiaPaCa-2, and the Negative Control libraries from the second checkerboard experiment. The expected p.Gly12Cys *KRAS* mutation in the red box was systematically detected in the MiaPaCa-2 libraries at 100% frequency, whereas it was not detected on any of the Negative Control libraries prepared by the VERSA™ 1100 GENE.

Figure 3. Reproducibility

A

Well	Sample ID	Library Conc. (ng/mL)	Well	Sample ID	Library Conc. (ng/mL)
1A	Pos Ctrl	1330	2A	NTC	0
1B	Neg Ctrl	2120	2B	Pos Ctrl	1302
1C	NTC	0	2C	Neg Ctrl	1530
1D	Pos Ctrl	930	2D	NTC	11.2
1E	Neg Ctrl	1704	2E	Pos Ctrl	764
1F	NTC	0	2F	Neg Ctrl	1564
1G	Pos Ctrl	1140	2G	NTC	0
1H	Neg Ctrl	2700	2H		

Sample ID	Library prep Method	No. of Variants	Pearson's r (against Manual library prep)
Pos Ctrl	Manual	36	N/A
Pos Ctrl	VERSA 1100	36	0.997
Pos Ctrl	VERSA 1100	36	0.995
Pos Ctrl	VERSA 1100	36	0.995
Pos Ctrl	VERSA 1100	36	0.993
Pos Ctrl	VERSA 1100	36	0.994
Neg Ctrl	Manual	14	N/A
Neg Ctrl	VERSA 1100	14	0.998
Neg Ctrl	VERSA 1100	14	0.998
Neg Ctrl	VERSA 1100	14	0.999
Neg Ctrl	VERSA 1100	14	0.992
Neg Ctrl	VERSA 1100	14	0.995

B

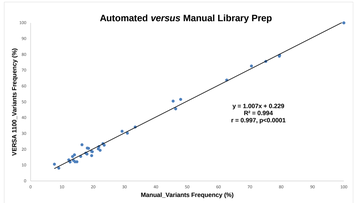


Fig. 3: **Reproducibility of Control Samples**
A- Library concentrations measured by the Qubit dsDNA HS Assay for five Positive and Negative control samples each (Left Panel) and number of variants and Pearson's correlations of variant frequencies with those obtained from manual library preparations (Right Panel). **B-** Representative curve showing Pearson correlation of the 36 variants frequency identified in the Positive Control sample by both library preparation methods.

Figure 4. Accuracy

Sample ID	No. of PCR cycles	Library prep Method	No. of Variants	Pearson's r (against Manual library prep)
Case_1	20	Manual	N/A (library failed)	N/A
Case_1	23	Manual	19	N/A
Case_1	20	VERSA 1100	19	0.992
Case_1	23	VERSA 1100	19	0.992
Case_2	20	Manual	N/A (library failed)	N/A
Case_2	23	Manual	17	N/A
Case_2	20	VERSA 1100	17	0.996
Case_2	23	VERSA 1100	17	0.997
Case_3	23	Manual	12	N/A
Case_3	23	VERSA 1100	12	0.995

Fig. 4: **Accuracy in the variants called on FFPE patient samples**
Difficult to amplify samples were chosen to compare the library yields and variants called from automatic *versus* manual library preparation protocols were used. Cases 1 and 2 failed to generate libraries using the manual protocol, so they were subjected to higher number of PCR cycles to generate libraries. For those samples, the VERSA™ 1100 GENE was used under both conditions, obtaining liberates even at fewer PCR cycles. The number and frequency of the variants found in every case were highly correlated.

CONCLUSIONS

- **From the checkerboard experiments, we concluded that this automated liquid handling system shows no evidence of cross-contamination, by either no library on the no template control (NTC) wells, or no variants called on negative samples after sequencing using the CHP2 assay.**
- **Also, high reproducibility was observed in both, library yields and variants called across all technical replicates of the Quality Control materials.**
- **All patient DNA samples yield good quality libraries, including those difficult samples that had previously failed using the manual library preparation method, and variants were called with highly correlated (Pearson's r>0.990) frequencies to those obtained with the manual method.**
- **Altogether, our results show that the performance of the VERSA™ 1100 Gene automated liquid handling workstation is very robust and might eliminate human-introduced errors, when compared to the manual library preparation method for the CHP2 assay.**

REFERENCES

1- Dumur CI *et al.* Quality control material for the detection of somatic mutations in fixed clinical specimens by next-generation sequencing. *Diagn Pathol*. 2015;10(1):169. PMID: 26376646, PMCID: PMC4573924