A research approach for the detection of somatic mutations at 0.5% frequency from cfDNA and cTc DNA using a multiplex sequencing assay targeting 2800 tumor mutations

Dumitru Brinza, Dalia Dhingra, Charles Scafe, Richard Chien, Fiona Hyland (Thermo Fisher Scientific, 180 Oyster Point Blvd., South San Francisco, CA 94080 USA)
Paul Dempsey (Cynvenio Biosystems, 2260 Townsgate Rd #2, Westlake Village, CA 91361 USA)

life technologies
A Thermo Fisher Scientific Brand

ABSTRACT

Availability of effective blood analysis process for tracking of recurrence and resistance of tumors may improve outcomes in the future. Research studies suggest that virtually all tumors carry somatic DNA mutations, and these may serve as biomarkers that may be tracked from blood. The two well-characterized sources of tumor DNA in blood are circulating tumor cells (cTc) and cell-free tumor DNA (ctDNA). The abundance of cTc and/or ctDNA in blood may be very low at critical stages such as early recurrence or development of resistance. Hence there is great interest in being able to detect biomarkers at very low frequency from blood, and in characterizing the relationship between somatic mutations present in the tumor and those in cTc or ctDNA.

Here, we present a research use only analysis workflow for peripheral monitoring that enables detection of mutations at frequency above 0.5% in normal blood, cTc enriched, and cell-free DNA (cfDNA) samples (Figure 1). We used lysis to isolate white blood cells (germline), centrifugation to extract plasma DNA (cfDNA), while cTc cells were isolated using Cynvenio LiquidBiopsy™ platform a fully automated antibody-based solution. We barcoded 3 sub-samples and run them on a single Ion PGM™ 318 sequencing chip using Ion AmpliSeq™ Cancer Hotspot Panel (CHPv2), that enables very deep (~10,000x coverage) and accurate sequencing. This panel allows interrogation of ~2800 relevant biomarkers from COSMIC and FDA actionable databases, and de-novo variant detection at ~20,000 genomic positions. Mutations were annotated using the Oncomine® database in Ion Reporter™ software. The research assay requires a small amount of input DNA (~10ng), and has a fast turn around time from extracted DNA to variants of less than 24 hr. An integrated analysis solution is released as “AmpliSeq CHPv2 peripheral/CTC/CF DNA single sample” workflow in Ion Reporter™ 4.4 Software.

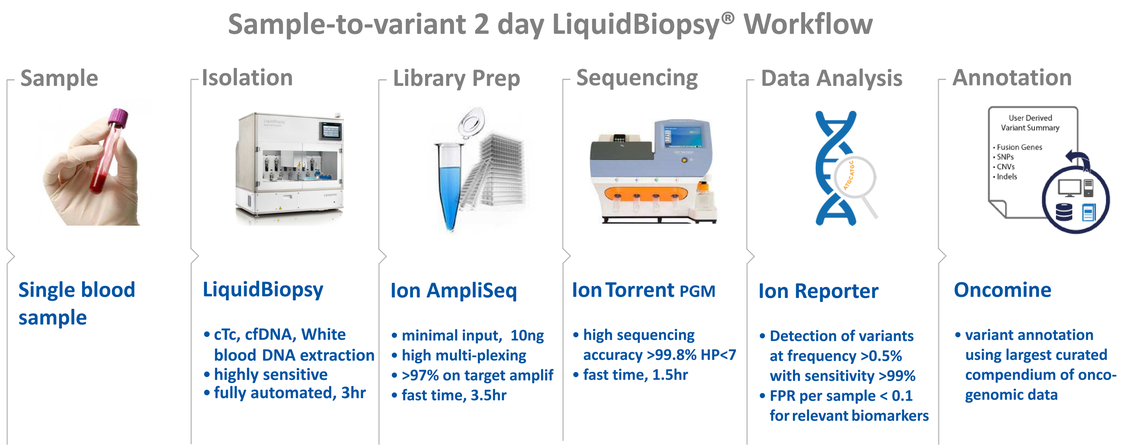

Figure 1. The LiquidBiopsy™ workflow enables cTc enumeration and analysis of cTc, cfDNA, and germline nucleic acids in a single workflow. Isolated cTc can be processed directly for downstream NGS analysis without whole genome amplification. Ion AmpliSeq™ technology allows low input DNA requirements at 10 ng or less. Ion Reporter™ Software reports variants at frequency 0.5% with sensitivity above 97% and false positive rate bellow 0.1 per sample at hot-spot positions and bellow 2 on the whole target (Figure 2).

RESULTS

Figure 2. Ion Reporter™ variant calling performance with 0.5% limit of detection in CHPv2 hotspots and 1% in the whole target

	CHPv2 Hotspots	*De novo* whole target regions		
Sample	**Plasmid Controls 0.5% dilution**	**Plasmid Controls 1% dilution**	**Control cfDNA 1% dilution**	**Cell Line DNA 1% dilution**
Type	SNP / Indel	SNP / Indel	SNP	SNP
Number	275 / 15	390 / 15	12	14
Sensitivity	>99% / >99%	>97% / >99%	>99%	>99%
FP/per sample	0	1	1	1

Figure 2. Verification of the limit of detection at 0.5% allelic frequency for CHPv2 hotspots was performed with engineered plasmid controls (AcroMetrix® Oncology Hotspot Control) diluted in background GM24385 genomic DNA. Similarly, plasmid controls, control cfDNA, and tumor cell line DNA were diluted to 1% allelic frequency for *de novo* whole target variant detection. Ion AmpliSeq™ libraries were made with CHPv2 and sequenced on the Ion PGM™ Sequencer. Variant caller detected SNPs and Indels with sensitivity >99% at 0.5% frequency for hotspots and >97% at 1% frequency for *de novo* targets, provided average coverage >2000x.

Figure 3. Verification of the LiquidBiopsy™ Workflow

	Target Cell Recovery	Target Cell Purity	PIK3CA p.E545K Expected Allelic Freq.	PIK3CA p.E545K Variant Caller Allelic Freq.
Spiked Sample 1	115.9%	72.1%	36.1%	29.2%
Spiked Sample 2	119.1%	76.5%	38.3%	39.8%
Spiked Sample 3	115.9%	69.6%	34.8%	37.0%

Figure 3. Verification of the LiquidBiopsy™ workflow was performed with MCF-7 cells spiked into normal donor blood at 90 cells/mL (provided by Cynvenio). 7.5 mL of the sample was subjected to the LiquidBiopsy™ workflow. Recovery and purity were assessed by DAPI, cytokeratin (CK), and CD45 staining to distinguish target (CD45-, CK+, DAPI+) from non-target (CD45+, CK-, DAPI+) populations. Isolated target cells were subjected to library prep and sequencing. Expected frequency of the heterozygous MCF-7 PIK3CA mutation was based on the target cell purity.

Figure 4. LiquidBiopsy™ Workflow: Read Depth at CHPv2 Hotspots

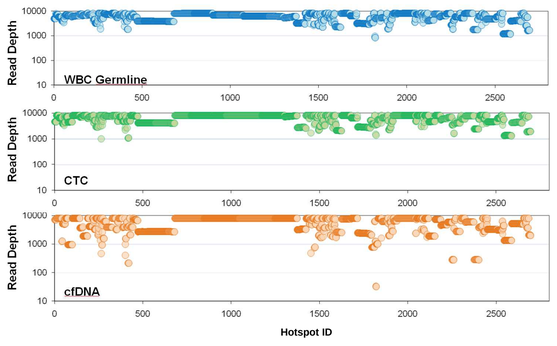

Figure 4. Sequencing data for WBC, CTC, and cfDNA from representative sample (provided by Cynvenio®) were obtained with the LiquidBiopsy™ workflow. The majority of the 2800 hotspots were covered with read depth >500x.

Figure 5. LiquidBiopsy™ Workflow: Ion Reporter™ Variant Calls

Location	Type	Ref	Alt	Gene	WBC Allelic Freq.	CTC Allelic Freq.	cfDNA Allelic Freq.	COSMIC ID
chr2:212812097	SNV	T	C	ERBB4	56.2%	63.8%	44.1%	-
chr3:178917005	SNV	A	G	PIK3CA	50.0%	50.1%	52.4%	-
chr3:178927410	SNV	A	G	PIK3CA	55.9%	55.3%	49.9%	-
chr3:178936082	SNV	G	A	PIK3CA	--	36.2%	7.1%	COSM760
chr4:1807894	SNV	G	A	FGFR3	99.7%	100.0%	99.8%	-
chr4:55152040	SNV	C	T	PDGFRA	49.1%	48.8%	44.6%	COSM22413
chr4:55972974	SNV	T	A	KDR	51.2%	49.9%	52.6%	-
chr7:55249063	SNV	G	A	EGFR	99.8%	99.8%	100.0%	-
chr10:43613843	SNV	G	T	RET	100.0%	100.0%	100.0%	-
chr11:534242	SNV	A	G	HRAS	48.3%	21.6%	47.1%	COSM249860
chr13:28610183	SNV	A	G	FLT3	49.6%	53.5%	50.6%	-
chr14:105241437	SNV	G	A	AKT1	--	34.8%	8.2%	-
chr17:7579472	SNV	G	C	TP53	100.0%	100.0%	99.7%	-
chr19:1220321	SNV	T	C	STK11	48.8%	37.7%	46.2%	-

Figure 5. The LiquidBiopsy™ NGS analysis utilizes Ion Torrent Suite™ and Ion Reporter™ Software for research reporting of and annotation of variants called in WBC, CTC, and cfDNA. Results from representative sample are shown. Two variants were called in the CTC and cfDNA which were absent in the WBC germline.

CONCLUSIONS

- The LiquidBiopsy™ Workflow with the Ion Torrent™ platform is a comprehensive 2 days sample-to-variant solution that facilitates researchers to study biomarkers in DNA from germline, cTc, and cfDNA all available from a single blood sample.
- The “AmpliSeq CHPv2 peripheral/CTC/CF DNA single sample” workflow in Ion Reporter™ 4.4 Software provides a fully automated analysis solution for accurate detection of varianst at frequency >0.5% (Figure 2).
- High cTc purity enables detection of 1:1M cells and >99% sensitivity and specificity in variant calling from the blood sample without the need for whole-genome amplification (Figure 3).
- Low-input DNA capability is provided by Ion AmpliSeq™ technology. The CHPv2 panel can survey ~2800 hotspots and ~22K de novo regions of oncogenes and tumor suppressor genes.

REFERENCES

For more information, please visit www.lifetechnologies.com.
For more information on AmpliSeq™, please visit ampliseq.com.

For Research Use Only. Not for use in diagnostic procedures.

MATERIALS AND METHODS

cTc enrichment: cTcs were enriched using the LiquidBiopsy™ platform. The LiquidBiopsy™ Blood Collection Kit was used for blood collection. The blood sample was labeled with EpCAM ferrofluid and put on the LiquidBiopsy™ instrument for cTc isolation. cTcs were digested and DNA subjected to Ion AmpliSeq™ library prep for sequencing.

cfDNA isolation: Plasma from the same blood sample was obtained by centrifugation at 500 x g for 15 minutes at 4°C. cfDNA was extracted from the plasma fraction using Qiagen™ kit for Ion AmpliSeq™ library prep.

White blood cell (WBC) germline control: 0.2 mL of the same blood sample was transferred to a 50 mL conical tube. Red blood cells were removed by lysis. White blood cells were collected and digested for Ion AmpliSeq™ library prep.

Sequencing: Libraries were generated with the Ion AmpliSeq™ protocol using the Cancer Hotspot Panel v2 and the Ion AmpliSeq™ Library Kit 2.0 according to the product manual. Libraries were quantified by qPCR and templating was performed on the Ion OneTouch™ 2 instrument. Enriched Ion Sphere™ Particles were loaded onto a 318™ Chip v2 and sequenced using the Ion PGM™ Sequencing 200 Kit v2.

Analysis: Signal processing and base-calling was performed with the Ion Torrent Suite™ , while alignment, variant calling, and annotations were generated using “AmpliSeq CHPv2 peripheral/CTC/CF DNA single sample” workflow within the Ion Reporter™ Software. This workflow is designed to capture variants present at frequency above 1% in the sample, or above 0.4% across multiple DNA readings. To enable capture of variants down to 0.5% frequency in the sample at hotspots we set “hotspot_min_allele_freq": "0.003“ in the input parameters json file (Figure 2).

Limit of Detection with fixed Ion AmpliSeq™ panels

Next-generation sequencing platforms may provide sufficiently high sensitivity for detecting low frequency variants down to 0.5-1.0%. Discriminating low frequency variants from pseudo-random errors may be done by requiring higher read coverage or repetitive sequencing; thus, enriching for reads containing variants.

In a well defined fixed Ion AmpliSeq™ panel most errors are systematic, i.e., the strand, position, type, and frequency of the error in reads covering same genomic position are replicable across multiple runs and samples. This is mostly due to the fixed and robust nature of Ion AmpliSeq™ panels, where properties of DNA molecules amplified from the same target are almost identical, facilitating error correction without requiring re-sequencing or higher coverage. We discovered systematic errors using germline, cTc, and cfDNA from 20 distinct samples (Fig 6).

Figure 6. Multi-sample approach for indentifying systematic errors

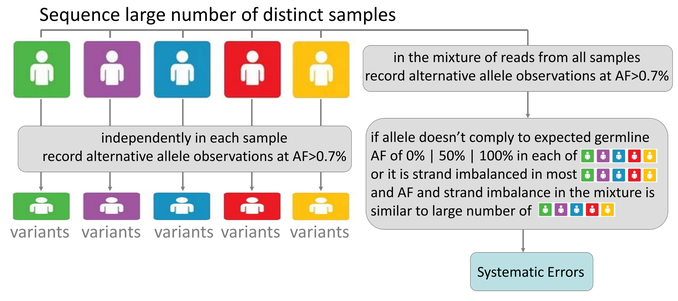

Figure 6. Majority of errors at frequency >0.7% are replicable across multiple runs and samples for a fixed Ion AmpliSeq™ panel. ~99.8% of errors are strand specific.

For the CHPv2 panel we built a model that allows accurate detection of variants at frequency >0.4% across reads (or >1% in the sample) at any position on the target excluding 10 positions that have >0.7% error rate on both strands. The set of detectable variants includes SNVs, Indels shorter than 10bp, and those that are not part of homopolymers of size 5bp or longer.