

Next-Generation Targeted Sequencing of Circulating Cell-Free DNA from Droplet Volumes of Blood

¹Jonathan Spurgin, ²Andrew Ford, ³John Athanasuleas, ⁴Chen-Hsiung Yeh

Circulo Gene Diagnostics, Birmingham, USA

Abstract: The non-invasive circulating cell-free DNA (cfDNA) approach – a liquid biopsy – is revolutionizing a paradigm shift in how cancer is detected, monitored and treated. In contrast to single-site, single time-point sampling by tissue biopsy, real-time and longitudinal mutation profile derived from tumor-specific cfDNA could potentially inform better and faster clinical decision-making, monitor tumor dynamics, assess response to treatment and identify mutations associated with acquired drug resistance. However, cfDNA analysis requires large volume of blood due to its relatively low amount in circulation and poor extraction efficiency of current methodologies. To overcome these major challenges, we have developed a proprietary cfDNA recovery technology with unique features of ultra-low input and ultra-high output. In this study, we evaluated our method side-by-side with the industry standard Qiagen kit, for the yield, cfDNA amplifiability and mutation detection from patient plasma.

Compared to Qiagen cfDNA extraction kit using different chemistry and different workflow, our approach allowed high-yield cfDNA enrichment directly from droplet volumes of unprocessed plasma, leading to >100-fold more recovery. NGS studies with cfDNA from 17 cancer plasma and 2 spiked samples further demonstrated the superiority of our protocol over Qiagen kit in generating more usable, on-target, high-quality \geq Q20 reads, and detecting more mutations.

Our cfDNA preparation breakthrough enables clinicians and laboratories to work with a sample volume as small as 20 microliters (via a finger-prick), in contrast to the current requirement of 10-20 milliliters, further expedite clinical decision-making and identify targeted therapies for eligible patients in a time- and cost-efficient manner.

Keywords: Cell-free DNA, liquid biopsy, blood-drop, next-generation sequencing.

I. INTRODUCTION

In the era of precision medicine, there is an enormous clinical need in determining the tumor genome to predict patients' response to certain anticancer agents. Interest in these “drug-test” combinations, also referred to as “companion diagnostics,” has increased exponentially laying a solid foundation for targeted therapy [1]. Targeted cancer therapies are drugs designed to interfere with tumor-specific molecules. Traditional cytotoxic chemotherapies or radiotherapies usually target dividing cells in the body including normal cells. A primary goal of targeted therapies is thus to fight cancer cells with more precision and potentially fewer side effects. First, the selection and intensity of treatment could be guided by the type of mutation(s) and mutant fraction existed at diagnosis; then, the response of patients could determine either extension of treatment in responders or early termination of therapy in non-responders; lastly, early detection of recurrence could allow salvage therapy to be instituted before complications of relapse develop.

Molecular analysis of genomic alterations are normally performed on archival tumor tissues, but there have been concerns that this approach does not truly and fully reflect the tumor genetic landscape at the time of targeted therapy, which is often months or even years after the primary diagnosis and/or surgery. Moreover, repeated biopsies are not feasible for

practical and ethical reasons. The advantage of plasma cell-free DNA (cfDNA) analysis to reveal de novo mutations allows monitoring of clonal heterogeneity at multiple time points, detecting resistant mechanisms early and tailoring treatment therapies accordingly [2-4]. Therefore, the ability of cfDNA as a liquid biopsy to detect those actionable tumor-specific mutations over time is warranted for better clinical outcome of targeted therapy.

Because of the small quantity and fragmented nature of cfDNA in circulation, a fast, efficient, and reliable isolation method is still a problem, and so far there is no consensus on a standardized protocol [5]. Most current existing methodologies utilized silica resin that selectively binds cfDNA fragments in the presence of salts. These commercial cfDNA extraction kits make the whole process much easier and faster than the conventional methods, but their recovery efficiency are extremely low due to sample loss during binding, washing and elution steps. Consequently, cfDNA-based analysis usually requires large volume of starting materials (>10 mL of blood) to get sufficient amount of DNA.

Obtaining molecular information from a precious sample can be challenging. Any sample processing condition can impact the end result - from sample acquisition to storage, quantity to quality, and isolation to analysis. Accordingly, any loss during sample preparation will significantly affect downstream testing results. We have developed a proprietary cfDNA enrichment and recovery technology for high-efficiency recovery of DNA fragments from biofluids through enzymatic DNA modification and manipulation, eliminating extraction-associated loss and RNA contamination. This novel technology ensures near-full recovery of both small-molecular-weight (apoptotic cell death) and high-molecular-weight (necrotic cell death) cell-free DNA species from droplet volumes of plasma or serum (20-150 uL). This sample preparation breakthrough enables multiple analyses with droplet volumes of sample on a broad range of genomic platforms, including next-generation sequencing (NGS). Our technology allows clinicians to work with a sample volume as small as 20 microliters via a finger-prick (next-generation liquid biopsy), instead of putting patients through blood-draw into several tubes.

In this study, we compared two cfDNA recovery methods, a leading commercial kit (Qiagen) and our patent-pending method by quantitative and qualitative analyses. Our results revealed that our method yielded much more high-quality cfDNA from much less input. Furthermore, cfDNA recovered using our technology produced more usable, on-target, high-quality \geq Q20 sequencing reads by NGS analysis, and detecting more actionable mutations than that isolated by Qiagen kit.

II. MATERIALS AND METHODS

Preparation of plasma cell-free DNA:

Blood was collected in EDTA-containing tubes (Becton Dickinson, Franklin Lakes, NJ) and was centrifuged at 2500 rpm for 20 minutes. Plasma was transferred to cryovials, being careful to avoid the buffy coat, and was stored at -80°C until further analysis. Circulating cell-free DNA was recovered from 20 uL and 200 uL of plasma sample using CirculoGene's proprietary cfDNA enrichment and recovery technology (CGD method) and QIAamp Circulating Nucleic Acid kit (Qiagen, Valencia, CA), respectively.

Quantification and amplification of plasma cell-free DNA:

Quantification of cfDNA was performed using the Qubit 2.0 Fluorometer together with dsDNA BR and HS assay kits (Life Technologies, Carlsbad, CA). Amplifiability of cfDNA was carried out in duplicate for each sample using TaqMan real-time quantitative PCR, with primers designed specific for *KRAS*, *BRAF*, *PIK3CA*, and *NRAS* genes (Life Technologies, Carlsbad, CA). The amplification plots and Ct values were generated by build-in software of QuantStudio 12K instrument (Life Technologies, Carlsbad, CA). Appropriate blanks and positive controls were included in each run to control the accuracy of PCR reaction.

Deep targeted sequencing and data analysis by Ion Torrent NGS:

Briefly, the targeted sequencing libraries were generated using the Ion AmpliSeq Library kit 2.0 and Cancer Hotspot Panel v2 according to the manufacturer's instructions (Life Technologies, Carlsbad, CA). The starting material consisted of 1-10 ng cfDNA prepared from CGD or Qiagen method. Each sample was analyzed for the entire 50-gene panel interrogating total \sim 2,800 mutations. The primers used for library amplification were then partially digested by Pfu enzyme, and followed by ligation with corresponding barcoded adapters and purified using Ampure Beads. The quality of

the libraries was assessed using quantitative PCR. Twenty picomoles of each library were put on the Ion Chef system for emulsion PCR to clonally amplify sequencing templates. Based on the number of sample analyzed, chip 314, 316 or 318 was used. Deep sequencing was performed on Ion Torrent PGM with coverage ranges of 1000-4000X. Sequencing data were analyzed by the Variant Caller 4.0 software using the somatic high stringency parameters and the targeted and hotspot pipelines. All the variants identified were further confirmed by analyzing the data through GenePool (Station X, San Francisco, CA).

III. RESULTS

We evaluated in parallel 2 approaches – CGD and Qiagen methods, on cfDNA recovered from 17 cancer patients. The experimental set-up involved DNA quantification by the fluorescent Qubit dsDNA BR or HS assay, and Taqman real-time PCR analysis on the *KRAS*, *BRAF*, *PIK3CA*, and *NRAS* genes [6]. Further, NGS was applied to detect mutations in these DNA sample sets. As shown in Fig. 1, CGD method recovered much more cfDNA than Qiagen kit by Qubit measurement. The average concentration of cfDNA recovered by CGD was 92.5 ng/uL compared to 0.42 ng/uL by Qiagen, with only one tenth input volume of Qiagen (N=17, P<0.0001). The higher recovery of cfDNA obtained with the CGD procedure was also revealed by agarose gel electrophoresis, which showed strong DNA intensities, in contrast, we can't visually detect cfDNA signals prepared by Qiagen. Fractionation of cfDNA revealed size distribution between 100- and 500-bp by the CGD method. Our CGD method also showed greater automation advantages, as it proved less time- and labor-consuming, and minimized the risk of contamination.

The amplifiability of cfDNA was then examined by TaqMan quantitative real-time PCR (qPCR) on four proto-oncogenes to evaluate the DNA quality by these two methods. The qPCR technique involves continuous monitoring of the progress of amplification and permits target quantification. As expected, amplification plots of *KRAS* showed ranges of threshold cycle (Ct) values between 17-24 and 28-33 for CGD and Qiagen cfDNA preparations, respectively (Fig. 2). In all the obtained samples, *KRAS* were amplified indicating the presence of sufficient cfDNA in the plasma samples. However, comparison of the Ct values indicates that our CGD protocol recovered at least 100-fold more amplifiable cfDNA than Qiagen kit ($\Delta Ct > 7$). We next determined whether similar differences existed with measurements made on the other 3 genes – *BRAF*, *PIK3CA*, and *NRAS*. The observed differences between the two methodologies were expressed graphically in representative amplification plots (Fig. 3). Consistently, much lower Ct values from CGD protocol compared to Qiagen method were observed across all 3 genes. The magnitude of ΔCt is in line with that of *KRAS*, confirming the superiority of our protocol over Qiagen in cfDNA recovery for both quantity and quality.

The ultimate evaluation of cfDNA preparation was performed using next-generation sequencing (NGS) analysis on the 50 cancer gene panel. We have obtained interpretable NGS results from 17 plasma samples (Table 1). Overall, eight cfDNA samples (8/17, 47.1%) extracted by Qiagen kit were QNS (quantity not sufficient), consistent with previous Qubit and real-time PCR results. Four concordant cases were determined to be wild-type (no mutation) by both methods. In another four cases, although Qiagen-extracted cfDNA showed no mutation, there is at least one mutation detected in the same samples prepared by CGD method. Another discordant case was wild-type when recovered by CGD protocol, but detected with a single mutation in the same sample by Qiagen extraction (low coverage, 463X). The coverage for those successful samples ranging 1006X - 3542X. After filtering out silent mutations and unconfirmed somatic mutations, 11 of 17 patients (64.7%) had at least one mutation by CGD protocol. All mutations identified were substitutions (Table 1). Together, our NGS results demonstrated that cfDNA extracted by Qiagen kit (200 uL plasma) could display high QNS and false negative rate due to significant loss in sample preparation step.

To further overcome the challenges surrounding variability between samples and potential failure to detect biomarkers, spike-and-recovery experiments were performed. Two NGS reference standards from Horizon Diagnostics were applied: standard 4 covers 10 mutations at 5% and standard 6 carries 20 mutations at 2.5%, both in genomic DNA format. We used two spiking concentrations 5 and 20 ng/mL in reference to 10-30 ng/mL of cfDNA usually found in healthy individuals. As shown in Table 2, sixty percent of mutations (6/10 and 12/20), either at 2.5% or 5%, can be detected in samples prepared by CGD method, in contrast, none was detectable in the same spiked samples extracted by Qiagen protocol. The results of spiking studies were consistent with previous NGS data and further confirmed that our CGD method recovered much more cfDNA from less plasma as compared to Qiagen kit, and allowed more mutations to be captured and identified.

IV. DISCUSSION

Traditional Sanger sequencing analysis on tissue biopsy has been widely used to guide therapy for cancer patients. Nevertheless, next-generation sequencing (NGS) technology holds a number of advantages over traditional methods, including the ability to fully interrogate large numbers of samples/genes/mutations in a single run, higher throughput, sensitivity and specificity, and automation-friendly [7, 8]. The rapid advances in NGS technology will further lower the overall cost, speed the turnaround time, increase the breadth and accuracy of genome sequencing, detect important genomic parameters, and most importantly, become applicable to lower-quantity and poor-quality specimens [9]. NGS technology brought immense opportunities for molecular diagnosis in clinical settings. The employment of non-invasive plasma cfDNA, as sources of genetic materials (liquid biopsy), coupled with NGS, offers an unprecedented platform for unmet clinical needs and improved cancer management.

In clinical setting, the quantity and quality of circulating cfDNA is critical, could heavily impact the testing outcome. Standardized sample collection and processing conditions are well conceived and easy to implement, e.g., using EDTA as anticoagulant, reducing cell lysis, complete removal of cells at plasma isolation, using fresh blood samples, discarding hemolytic samples, transferring plasma to -80°C timely, and avoiding repeated freezing and thawing etc [10]. On the other hand, the loss of starting materials during cfDNA recovery process is often overlooked and ignored. It should be noted that any loss during this step could dramatically reduce test sensitivity and accuracy. Unfortunately, current existing cfDNA isolation methodologies are inherited with unavoidable sample loss, low recovery efficiency, leading to the requirement of large sample volume and thus are far from satisfactory [5]. Our novel cfDNA recovery technology in this regard provides an ultimate "one-for-all" solution to three major unsolved challenges in the clinical application of cfDNA-based liquid biopsy, i.e., input volume, output quantity and quality. In this report, we have demonstrated that the key performance characteristics (quantity, quality and amplifiability) of our technology is far superior over the leading commercially available cfDNA extraction kit. This makes our method an attractive technology for use in liquid biopsy diagnostic applications more so than other conventional silica matrix-based protocols: (1) minimal input - droplet volumes vs. 10-20 mL of blood; (2) maximal output - more than 100-fold yield of Qiagen kit; (3) superior quality - capture more mutations than Qiagen extraction (the current industrial standard) by NGS. The study with our limited sample set showed encouraging findings. Studies with much larger cohorts including tissue biopsies and result comparison with other laboratories are necessary in future to technically and clinically validate our innovative blood-drop cfDNA approach to guide targeted therapy of cancer.

As shown in our experiments here, sequencing negative by Qiagen method does not mean no mutations, it is highly likely that portions of the mutant cfDNA was lost during sample preparation. Contaminants may also be responsible for poor yield and low sensitivity because of "carry-over" of ethanol or salts that could inhibit or reduce the efficiency of downstream applications such as PCR, sequencing and generate false-negatives [11, 12]. Moreover, the Qiagen method is expensive, time-consuming and labor-intensive. Although with these limitations, many laboratories still use Qiagen kit for cfDNA preparation, the problems in sample loss, low extraction efficiency and yield will, in time, support a transition to a higher throughput, cost-effective and near-full recovery technology.

The spiking studies firmly indicated that the decrease in mutation detection is the result of loss of cfDNA fragments during Qiagen extraction step, and that the resulting high-quality cfDNA by our method is suitable for ultra-deep targeted sequence analysis.

V. CONCLUSION

In the era of precision and personalized medicine, by generation of "actionable" gene sequence abnormalities, our blood-drop cfDNA-based NGS-driven clinical service will have a custom report that is focused on the abnormalities that can lead directly to the best available targeted therapy, direct a patient to an ongoing clinical trial, or result in some other actionable response from the treating clinicians.

REFERENCES

- [1] Papadopoulos N, Kinzler KW, Vogelstein B (2006) The role of companion diagnostics in the development and use of mutation-targeted cancer therapies. *Nat Biotechnol* 24:985–995.

- [2] Dawson SJ, Tsui DW, Murtaza M et al (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med* 368:1199-1209.
- [3] Diaz LA Jr, Williams RT, Wu J et al (2012) The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486:537–540.
- [4] Bettegowda C, Sausen M, Leary RJ et al (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 6:224ra24.
- [5] Fleischhacker M, Schmidt B, Weickmann S et al (2011) Methods for isolation of cell-free plasma DNA strongly affect DNA yield. *Clin Chim Acta* 412:2085–2088.
- [6] De Roock W, Claes B, Bernasconi D et al (2010) Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* 11:753-762.
- [7] Metzker ML (2010) Sequencing technologies: the next generation. *Nat Rev Genet* 11:31–46.
- [8] Ross JS, Wang K, Gay L et al (2014) New routes to targeted therapy of intrahepatic cholangiocarcinomas revealed by next-generation sequencing. *Oncol* 19:235-242.
- [9] Newman AM, Bratman SV, To J et al (2014) An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nature Med* 20:548-554.
- [10] Hung EC, Chiu RWK, Lo YM (2009) Detection of circulating fetal nucleic acids: A review of methods and applications. *J Clin Pathol* 62:308–313.
- [11] Breitbach S, Tug S, Helmig S et al. (2014) Direct Quantification of Cell-Free, Circulating DNA from Unpurified Plasma. *PLoS ONE* 9:e87838. doi:10.1371/journal.pone.0087838
- [12] Fong SL, Zhang JT, Lim CK et al (2009) Comparison of 7 methods for extracting cell-free DNA from serum samples of colorectal cancer patients. *Clin Chem* 55:587–589.

APPENDIX – A

List of figures:

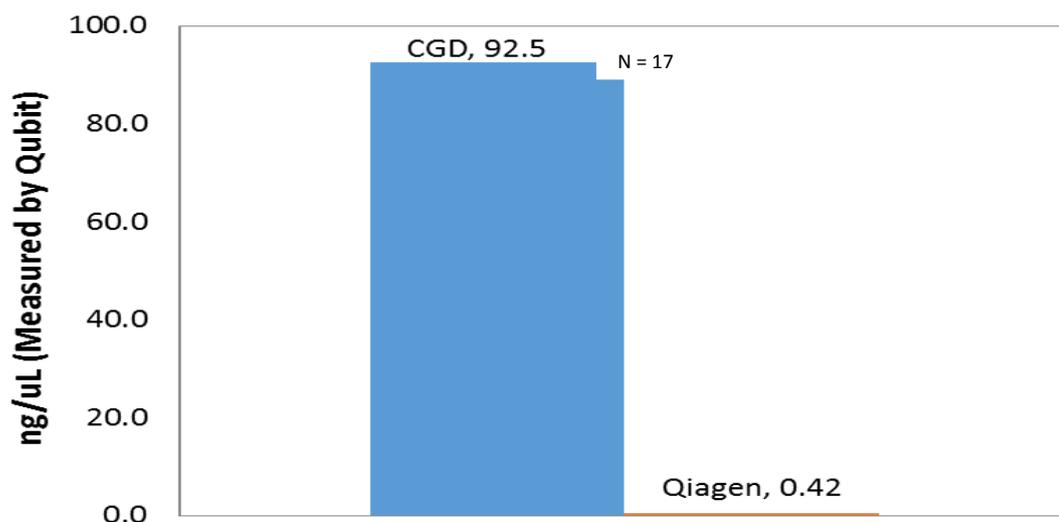


Fig. 1. Determination of plasma cfDNA concentration recovered by CGD or Qiagen method. Plasma cfDNA amounts were measured by Qubit fluorometer with dsDNA BR and HS assay kits. Data shown here are average concentrations from total 17 samples (P<0.0001 by Student's t-test)

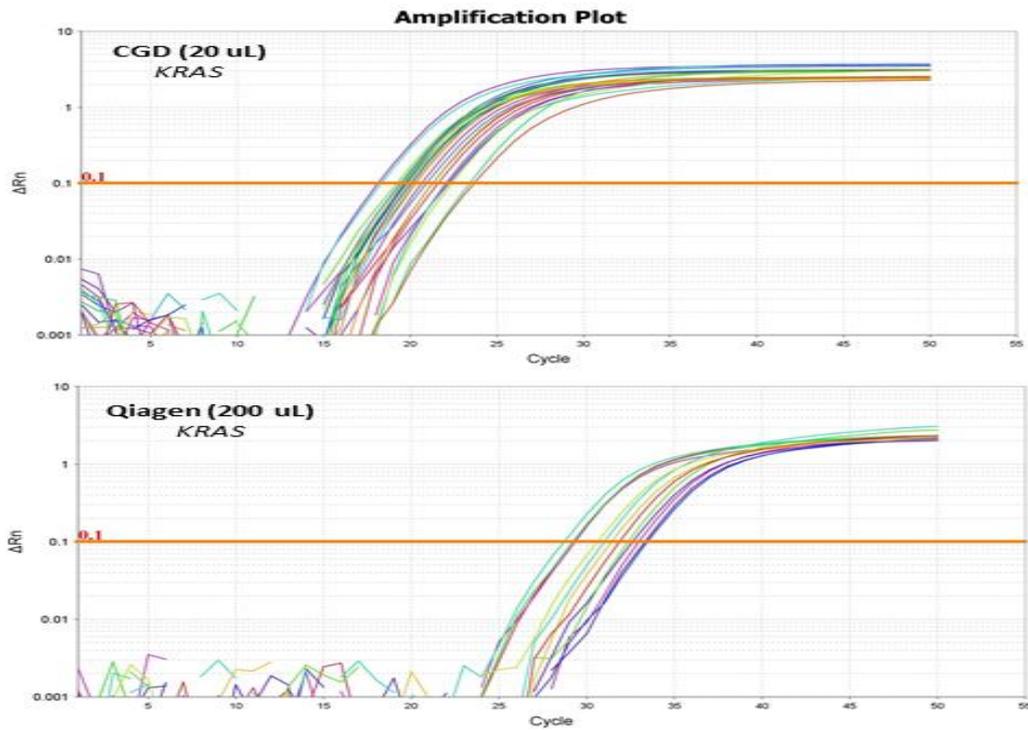


Fig. 2. Amplification plots of *KRAS* gene using cfDNA recovered by CGD (from 20 uL plasma) or Qiagen (from 200 uL plasma) method. Ct values were determined by setting threshold at 0.1

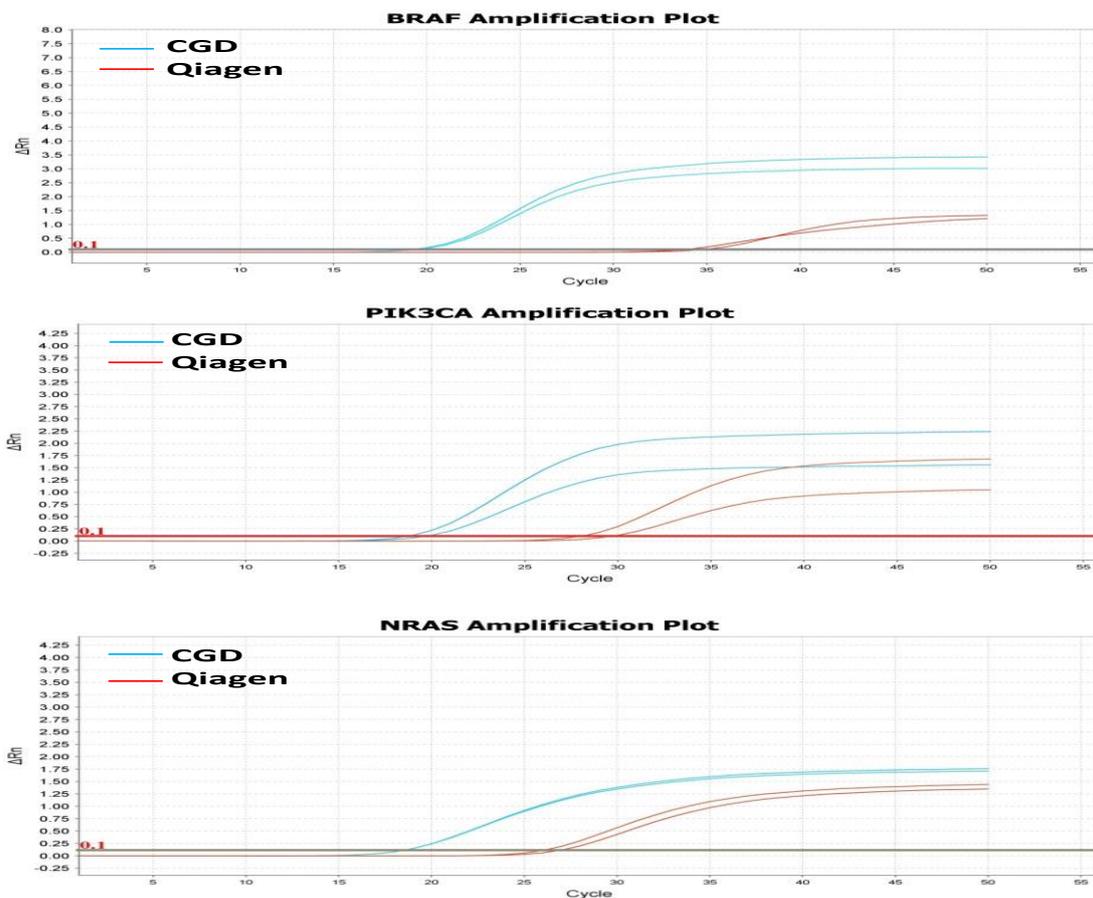


Fig. 3. Amplification plots of *BRAF*, *PIK3CA*, *NRAS* genes using cfDNA recovered by CGD (from 20 uL plasma) or Qiagen (from 200 uL plasma) method. Ct values were determined by setting threshold at 0.1.

APPENDIX - B

List of Tables:

Table 1. Summary of mutation detection by NGS on cfDNA prepared by CGD or Qiagen method

Sample	Mutation Detected by CGD Method (Mutant %, Read Depth)	Mutation Detected by Qiagen Method (Mutant %, Read Depth)
1	EGFR, G719D (6.5%, 1358)	QNS
2	PIK3CA, K111R (2.1%, 1189); ABL1, Y253H (2.4%, 1189); SMAD4, C499Y (2.1%, 1189)	QNS
3	None	QNS
4	None	None
5	HNF1A, R272H (3.0%, 1191); TP53, E221K (4.8%, 1191)	None
6	GNAS, R844C (3.2%, 1338); GNAS, R201C (3.2%, 1338)	QNS
7	PIK3CA, K111R (5.1%, 1023); EGFR, E734K (7.6%, 1023); EGFR, R776H (2.0%, 1023); HNF1A, T260M (13.1%, 1023)	None
8	PIK3CA, K111R (2.9%, 1783); EGFR, I821T (2.4%, 1783); TP53, R273C (3.7%, 1783)	QNS
9	VHL, C162R (1.9%, 1006); CTNNB1, T40I (1.8%, 1006)	None
10	None	None
11	PTEN, Q171E (3.1%, 1507)	QNS
12	TP53, P72S (3.3%, 1108)	None
13	None	None
14	None	TP53, S215G (2.2%, 463)
15	None	None
16	BRAF, V600M (3.8%, 2412); BRAF, I592V (3.3%, 2412); TP53, Y205H (2.2%, 2412); SMAD4, D355G (11.8%, 2412)	QNS
17	TP53, R249G (8.1%, 3542)	QNS

QNS: Quantity Not Sufficient

Table 2. Summary of spiking studies on cfDNA prepared by CGD or Qiagen method

	CGD Method	Qiagen Method
Spiked 5 ng/mL Std 4 Reference DNA (10 mutations, 5% mutant fractions, Horizon Dx)	6/10	0/10
Spiked 20 ng/mL Std 6 Reference DNA (20 mutations, 2.5% mutant fractions, Horizon Dx)	12/20	0/20