Characterization of Synthetic Reference Material for ctDNA Detection Assays

Characterization of Synthetic Reference Material for ctDNA Detection Assays

Sarah Lynne Fortunata Simmons

A Thesis in the Field of Biology

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2022

Abstract

Precision medicine is a fast-growing field within the cancer biotechnology industry that tailors a patient's treatment to their specific biological makeup. Liquid biopsy offers a noninvasive, more heterogenous look into a patient's cancer compared to solid tissue biopsies. Liquid biopsy circulating tumor DNA (ctDNA) detection assays allow each patient's treatment to be informed by their individual cancer mutational status. The development and validation of these diagnostic tests are essential for them to be a clinically relevant tool - a tool that doctors and patients can use. Validation provides analytical accuracy and is what is used to build a case for clinical utility. A key part to validating a diagnostic test is to have a reliable reference material to which you can assess the diagnostic test against. Finding or making the correct reference material is difficult, and there are particular hurdles to overcome when designing a biologically similar reference material for ctDNA. There are preanalytical variables that can affect ctDNA measurements, from the way the DNA is fragmented to the background DNA that is incorporated and to the matrix it is suspended in. This thesis evaluates off-the-shelf reference materials for their suitability to be used for ctDNA assay validation. Reference material from 3 different vendors; Anchor Molecular, Horizon, and SeraCare, are evaluated. Each vendor is assessed on accuracy, precision, and ctDNA concordance. Overall Seracare has the best performance, next is Anchor Molecular and then Horizon. Choosing the correct reference material is imperative to validating a ctDNA detection assay and comparisons like these can aid in such imperative choices.

Dedication

To my dad who taught me that nothing is out of reach as long as you take the time to learn it. To my mom who is my biggest supporter and has never stopped believing in me. To both of them for their unconditional love. Thank you.

Acknowledgments

Thank you to my advisers, coworkers, friends and loved ones that have helped me on my way to completing this thesis. I could not have done it without the tremendous amount of support you have given me. Thank you from the bottom of my heart.

Table of Contents

List of Tables

List of Figures

Chapter I.

Background

Liquid Biopsies

Precision oncology's main objective is to improve the diagnosis and treatment of cancer (Heitzer, Haque, Roberts, & Speicher , 2018). Current precision oncology consists of a range of Next Generation Sequence (NGS) based approaches that are used to characterize a patient's tissue tumor sample (Heitzer, Haque, Roberts, & Speicher , 2018). In recent years, liquid biopsies for cancer mutation detection have become an important tool for precision oncology (Cescon, Bratman, Chan, & Siu, 2020) Liquid biopsies present a way of obtaining diagnostic information about a complex human disease from a simple blood draw or bodily fluid extraction, and to not require a tissue sample (Heitzer, Haque, Roberts, & Speicher , 2018). The non-invasive manner, and the potential to track a disease over time, make liquid biopsy an attractive approach in precision oncology (Heitzer, Haque, Roberts, & Speicher , 2018). Liquid biopsies bring forward new possibilities for cancer care, such as the monitoring of treatment response and the potential for early screening and diagnosis (Geeurickx & Hendrix, 2020). There are many analytes present in liquid biopsies that have the potential to tell us more about a patient's cancer. Researchers are examining analytes such as circulating tumor cells (CTC), extracellular vesicles (EV's), cell free RNA (cfRNA), micro RNA (miRNA), cell free DNA (cfDNA), and circulating tumor DNA (ctDNA) (Heitzer, Haque, Roberts, & Speicher , 2018). These analytes can represent the molecular status of heterogeneous

tumors or metastatic microenvironments found in cancer patients that can provide deep insight into a patient's cancer. Further, given the ability for repeat sampling, liquid biopsies provide the potential to capture multiple timepoints and track the changes of the patient's tumor (Geeurickx & Hendrix, 2020).

cfDNA

Circulating-free DNA/cell-free DNA (cfDNA) is one analyte that has been in the forefront of precision medicine. cfDNA is the fragmented DNA in the non-cellular component of blood (Geeurickx & Hendrix, 2020). cfDNA was first mentioned in a paper by Mendel and Metaisandin in 1948 and was first found to be mutated in cancer patients in 1994 (Geeurickx & Hendrix, 2020). More recently it has been revealed that cancer patients can have greater levels of plasma cfDNA than healthy individuals (Geeurickx & Hendrix, 2020). The origins of cfDNA are still somewhat elusive. Most evidence suggests that cfDNA in healthy and diseased individuals is from an apoptotic origin (Bronkhorst A. J., 2019). Enzymatic cleavage of DNA during apoptosis results in DNA fragments of approximately 166bp in length (Geeurickx & Hendrix, 2020). This fragment size corresponds to DNA wrapped around a single nucleosome which protects the DNA from degradation (Geeurickx & Hendrix, 2020). Larger fragment sizes exist in cfDNA and are derived from a necrotic origin. Necrotic-derived fragments start at 320bp, corresponding to DNA wrapped around two nucleosomes, and range up to 10,000bp (Geeurickx & Hendrix, 2020). In an attempt to elucidate the origins of cfDNA, Bronkhorst,et al.(2016) conducted a study with cancer cell lines and concluded that while there is apoptotic release of cfDNA there is also a large portion of actively released

cfDNA into the circulatory system. Active secretion of cfDNA is thought to be through EV's or DNA association with a lipid protein complex, released to help maintain homeostasis (Bronkhorst, et al., 2016). Secretion of cfDNA in the blood still may not be fully understood, but it is increasingly clear that DNA is released into the blood through a combination of mechanisms: necrosis, apoptosis, and active secretion (Bronkhorst, et al., 2016). Circulating-tumor DNA (ctDNA) is a subpopulation of cfDNA that is specifically shed from tumors (Geeurickx & Hendrix, 2020) and is an analyte that researchers are interested in isolating and using for diagnostics. ctDNA is distinguished physically from cfDNA in a few ways. In cancer patients, ctDNA fragments lie within a range of 90-150bp, slightly shorter than normal cfDNA fragments (Geeurickx & Hendrix, 2020). cfDNA found in blood plasma has a half-life of 2.5 hr and reported concentrations that range from 1.8-44ng/ml, whereas ctDNA concentrations are very low and the amounts vary from patient to patient depending on stage and type of cancer (Geeurickx & Hendrix, 2020). In general, cfDNA concentration is less than 100ng/ml, and only a fraction of all cfDNA is tumor derived (Khatami & Tavangar , 2018). ctDNA makes up a very low proportion of the total liquid biopsy and one major challenge in the analysis of ctDNA is its detection at low fractions with high sensitivity. However, with advances in molecular technology, ctDNA detection with high sensitivity has been made possible (Burgener, Rostami, De Carvalho, & Bratman, 2017). Technologies used to detect ctDNA include NGS, and digital PCR (dPCR) methods; microfluidic based droplet digital PCR (ddPCR) and BEAMing (beads, emulsion, amplification, magnetics, combination PCR method) (Bronkhorst A. J., 2019). These technologies usually allow mutant alleles to be identified down to 2% allele frequency (Fortuna & Dvir, 2020).

There are a few ultrasensitive methods that report a sensitivity of ctDNA detection below .01% (Johansson, et al., 2019). To do so requires improved sequencing with error correction that allows for improved sensitivity, or specific PCR methods that have sensitivity capabilities that are restricted to a few target mutations (Johansson, et al., 2019). Despite the growth seen in the ctDNA liquid biopsy field, there is still room for improvement and hope that liquid biopsy can potentially be utilized in screening for the early detection of cancer. Improving sensitivity and making a more standardized workflow is a critical next step for ctDNA liquid biopsies (Geeurickx & Hendrix, 2020). The full clinical application of ctDNA in liquid biopsies has been stalled as there is a need for clinical validity and utility of the liquid biopsy space (Geeurickx & Hendrix, 2020).

Advances have been made in the liquid biopsy space, with many new methods of measurement and analysis being developed for ctDNA (Cescon, Bratman, Chan, & Siu, 2020). As the field progresses there have been a growing number of tumor based tests, for the detection of therapeutic intervention, that include diagnostic, prognostic and predictive measures (Cescon, Bratman, Chan, & Siu, 2020). ctDNA liquid biopsies have begun to have some routine clinical uses in the metastatic setting, identifying actionable tumor mutations that can be treated with targeted therapy (Cescon, Bratman, Chan, & Siu, 2020). These analytes have also begun to be used in the evaluation of early treatment response and the discovery of the resistance mechanisms used by cancer cells (Cescon, Bratman, Chan, & Siu, 2020). ctDNA detection in liquid biopsies are an emerging diagnostic tool for clinicians, allowing for a novel way to examine cancer (Fortuna &

Dvir, 2020). As the field of liquid biopsy has grown, so has the need for validation. The validation of diagnostic tests is essential for assays to become a clinically relevant tool.

Validation

Validation of a diagnostic test is a complex multi-part process. It begins with the initial discovery of the analyte and biomarker through primary experimental work. Following this, it is validated, ideally, by independent studies, finally ending with clinical implementation (Goossens, Nakagawa, Sun, & Hoshida, 2015). A biomarker is defined in very broad terms by the World Health Organization as "...any substance, structure or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease." (Goossens, Nakagawa, Sun, & Hoshida, 2015). Cancer biomarkers like ctDNA can be put into 3 categories: predictive, prognostic, and diagnostic. Predictive biomarkers show a response to specific therapeutic interventions, prognostic biomarkers show overall disease progression, and diagnostic biomarkers allow for the identification of a patient-specific disease condition (Goossens, Nakagawa, Sun, & Hoshida, 2015). After biomarker discovery, clinically applicable assay platforms for the biomarker are validated. The assay platform undergoes two types of validation: analytical and clinical. Analytical validation refers to how accurately and reliably the assay measures the analyte of interest in the patient specimen (Goossens, Nakagawa, Sun, & Hoshida, 2015). Clinical validation refers to how robustly and reliably the assay results are tied with the clinical outcome of interest (Goossens, Nakagawa, Sun, & Hoshida, 2015). A biomarker assay that has been both analytically and clinically validated is ready to be implemented in clinical care. This implies that the assay is ready for regulatory

approval, commercialization, and incorporation into clinical practice guidelines and health insurance coverage (Goossens, Nakagawa, Sun, & Hoshida, 2015).

Analytical validation explores all aspects of an assay performance and is an evaluation of the data generated from the methods used to make said product (Valverde, 2005). There are many characteristics that should be considered when validating an assay (Valverde, 2005). Governing bodies such as the FDA provide guidance on the analytical validation process for differing biomarkers, regardless of the platform used for an assay (Kraus, 2018). In 2018, the FDA put out guidelines outlining key considerations for designing, developing, and establishing analytical validity of Whole Exome Sequencing and Targeted sequencing, NGS based DNA testing. The main components of assay validation can include any of the following: selectivity, accuracy, precision, recovery, sensitivity, reproducibility and stability (Kraus, 2018). The FDA guide suggests for NGS based DNA testing for diagnosis of germline diseases that the performance metrics used for analytical validation should include measurements of accuracy, precision, limit of detection, and analytical specificity (Food and Drug Administration, 2018). Measuring a test's analytical performance includes demonstrating whether the test can successfully measure and identify the predefined indications that will provide information on a disease (Food and Drug Administration, 2018). These analytical validation guidelines are for the diagnosis of germline diseases and not for tumor genome sequencing and cell free DNA (Food and Drug Administration, 2018). Analytical validation for cell free DNA is still being investigated; however these guidelines do explicitly say that even though the principles described in the guidelines may not include all the performance characteristics or risk considerations for cell free DNA and tumor

genome sequencing, it may be applicable to those types of tests depending on use (Food and Drug Administration, 2018). The FDA encourages the scientific community to discuss and apply these principles of these guidelines to other NGS tests (Food and Drug Administration, 2018). Geeurix et al. (2020) states: Analytical validity of a liquid biopsy detection assay includes an assessment of a test's accuracy, sensitivity, specificity and robustness (Geeurickx & Hendrix, 2020). The demonstration of accuracy, precision, and reproducibility of a test helps determine how well the test measures what it claims to measure (Lathrop, 2015). cfDNA detection assays, due to dependence on various pre-analytical variables and protocols for sample preparation, have been difficult to analytically validate (Geeurickx & Hendrix, 2020). The field is coming along with developing ways to analytically validate a cfDNA detection assay. A key component to validating a diagnostic test is to have a reliable reference material in which you can assess the diagnostic test.

Reference Material

Reference material serves as a truth set that can be used to demonstrate assay performance (Ashford, 2017). They are used to test an assays' limits, ensure the assay is being performed correctly and allow for the precise comparison of the assay to other tests (Ashford, 2017). Reference material can be used to investigate an assay's selectivity, accuracy, precision, recovery, sensitivity, reproducibility and stability, and moreover is a tool used to analytically validate an assay. According to the ISO Council Committee on reference material, the definition of a reference material is a sufficiently homogeneous and stable material with respect to a specific property that is established to be fit for the

intended use in the measurement process (Geeurickx & Hendrix, 2020) (CLSI, 2010). The reference material definition is broken up further into different categories: certified reference materials, quality control materials, and calibrants (Geeurickx & Hendrix, 2020). Certified reference material is a material that is characterized through a metrologically validated process for one or more specified properties (Geeurickx & Hendrix, 2020). A metrologically characterized process is accompanied by a statement of metrological traceability and a certificate claiming the value and uncertainty of the specified property (Geeurickx & Hendrix, 2020). Quality control materials are not characterized enough to have metrological traceability, nor to be certified, they are often produced and used in-house (Geeurickx & Hendrix, 2020). Calibrants are ideally certified reference material that can provide a fixed property value - specifically a property value that has an appropriate uncertainty level that is tolerable for calibration and metrological traceability of that property value (Geeurickx & Hendrix, 2020). Calibrants are not always certified reference material and certified reference material are not always calibrants; the uncertainty of the certified reference material may not fall within a tolerable level for calibration (Geeurickx & Hendrix, 2020). There are few certified reference materials currently for liquid biopsy assays, so authors Geeurickx and Hendrix (2020) assign the terms “pre-analytical reference material” and "analytical reference material” when discussing liquid biopsy reference material. Pre-analytical reference material encompasses material used for calibration and quality control and only needs to match the physical characteristics of cfDNA. Analytical reference material refers to materials used to assign values to other material or LOD measurement method determination, and not only needs to match the physical characteristics but also the

biological characteristics of ctDNA. This thesis explores available analytical reference material for ctDNA detection assays.

For ctDNA detection assays, patient plasma samples, in terms of biochemical properties, would be the most ideal control material (Tsang & Chan, 2017). Unfortunately, not only would it be impractical to obtain the sufficient number and amount of patient plasma needed to facilitate a comparison of different measurement methods, but the varying fraction levels of mutations within the plasma, necessary for an LOD study, would be hard to obtain and adjust (Tsang & Chan, 2017) (He, et al., 2019). A good quality control material for a circulating cell-free DNA detection assay would be a material that biologically resembles authentic circulating DNA in cancer patients (Tsang & Chan, 2017). One way to judge a reference material is by looking at the commutability of the material. The recommended definition of the term commutability from the Clinical and Laboratory Standards Institute is when the reference material and a representative sample that is being measured have an equivalent mathematical relationship among many different procedural results (CLSI, 2010). There are two uses for the characterization of commutability that CLSI sets forth. One is to establish calibration traceability chains, and the other is to characterize reference materials that are used to assess the trueness of results from assay laboratory measurements (CLSI, 2010). The commutability of reference material defines the reference material's fitness for use as a truth set (CLSI, 2010). For a ctDNA detection assay, when comparing native samples and reference material, one would look to see if the extracted DNA profiles match, measuring extraction yield and size distribution, library yield, library complexity, and how often variants are found between the two.

Finding or making the correct reference material is difficult, and there are particular hurdles to overcome when designing a biologically similar reference material for ctDNA. Specifically, for ctDNA detection assays, there are many preanalytical variables that can affect ctDNA measurements, from the way the DNA is fragmented to the background DNA that is incorporated and to the matrix it is in. To mimic plasma DNA, these are the three main components that are most vital to mimic. As previously described, ctDNA is not randomly fragmented; the fragmentation is related to nucleosome positioning. Circulating tumor derived DNA fragments end up being shorter than the non tumoral-derived cell free DNA (Tsang & Chan, 2017). The performance of different analytical platforms is greatly affected by this property, making mimicking the fragment size distribution of circulating tumor DNA a key component to making a quality control material (Tsang & Chan, 2017). In addition to having a similar fragment distribution, these fragments also must contain mutations within them. A good quality control material should also allow for exploration into a variety of clinically relevant alterations, including small nucleotide polymorphisms (SNPs) as well as rearrangements and copy number variation (Tsang & Chan, 2017). Another property that is important to mimic is the genetic background of the cell free DNA. ctDNA makes up a very small portion of the cfDNA in the blood, meaning the somatic variant alleles present are at low concentrations relative to germline DNA (He, et al., 2019). Having a control with a defined cfDNA background that is isogenic to the mutant DNA allows for the analysis of the bioinformatic pipeline's variant identification performance (Tsang & Chan, 2017). It's critical to have a reference material where the total cfDNA amount and the mutant allele fraction is quantitatively defined and representative of the different disease stages (Tsang

& Chan, 2017). The last important component to mimic is the plasma that the cfDNA is in. To describe ctDNA in plasma as a simple mixture of double stranded non-interacting molecules is an oversimplification (Tsang & Chan, 2017). These ctDNA molecules associate with a whole host of other biological components that make up blood/plasma, such as nucleosome proteins, and immunoglobulins (Tsang & Chan, 2017). ctDNA exhibits clear differences in stability and configuration depending on the sample matrix composition, and it is a very important component to mimic (Tsang & Chan, 2017).

ctDNA Controls

When developing a ctDNA control, the aforementioned three components of mimicking native ctDNA samples must be kept in mind. In a quest to best mimic ctDNA, researchers have tried a multitude of combinations of fragmented DNA, background DNA, and matrix types, with the hopes of concocting an appropriate control to use for the analysis of ctDNA detection assays. Researchers have developed sets of synthetic cell-free DNA quality control materials. Zeng et al. (2017) describe a synthetic cfDNA quality control material (SCQCM) that they have developed. They used DNA from a cancer cell line that they fragmented using MNAse, a nuclease that is able to produce a size profile analogous to ctDNA, as it cuts DNA between nucleosomes. Using Crisper and PCR directed mutagenesis, they were able to create fractured DNA that had a variety of mutations from SNPs to rearrangements. From there, they combined their mutated and non-mutated DNA fragments, forming a mixture of mutated DNA in a background of non-mutated fragmented DNA. The background DNA and mutated DNA are genetically similar- they are from the same cell line and are fragmented in the same way. This spike-

in allows for examination of variant stability assessing if the variants were accurately detected at the proper allele frequency, and are distinguishable from the germline variants, across multiple assay types. Sequencing metrics were similar to that seen in native plasma samples. A major limitation for this reference material is the sample did not involve cfDNA extraction from a matrix or some sort of complex bio-fluid. Extracting reference material from a matrix is a key part to mimicking a native plasma sample. Ideally a reference material should be applicable to evaluate the complete cfDNA workflow starting with plasma preparation where many variables can enter, all the way to sequencing analysis (Geeurickx & Hendrix, 2020). This would be an example of a full reference material, versus reference materials that only evaluate one part of the process.

Synthetic controls that can be used as a full analytical reference material are of interest to the field of liquid biopsy. This research will specifically study the use of synthetic full reference material that allows for the measure of trueness that can be assessed due to its commutability with native clinical samples. A variety of different companies have tried their hand at making a full reference material. Anchor Molecular (AM), Horizon (HZ), and SeraCare (SC) are three that offer a full reference control.

Anchor Molecular (AM) has a reference material called “AM Multiplexed Mutation Reference Standard”, which is fragmented DNA with mutations at varying allele frequencies in a complex patient-like material and 125bp gDNA (genomic DNA) background. AM has made a plasma matrix that is depleted of DNA but mimics the proteins and other elements found in plasma. AM showed that their DNA depleted plasma compared to synthetic or TE (Tris – EDTA) matrices made their inserted DNA profile mimic native ctDNA more closely (Lu, Ooi, Shi, & Chan, 2020)**.** They spiked in

170bp fragments of BRAF mutated cell line and a gDNA background into the differing matrices then compared which scenario best mimicked native ctDNA (Lu, Ooi, Shi, & Chan, 2020). AM's reference material's advertised function is to quantify the relative number of mutant alleles present. This reference material looks at a variety of small nucleotide polymorphisms (SNP's), some insertions and deletions, and a few rearrangements. They also offer customizable reference material. Their reference material includes 37 different mutant genes, being offered at differing allelic frequencies: 0%, 0.1%, 0.3%, 1%, or 2%. The control material is sold in 5 ml plasma aliquots catalog number 60100202.

Horizon markets a full reference control material as well - the "Multiplex I cfDNA in Synthetic Matrix I", which is advertised to be an optimal synthetic matrix that can be used to access extracted ctDNA recovery and clinically relevant mutations. The synthetic matrix contains 8 cancer-relevant mutations at varying allele frequencies. The mutations include EGFR, KRAS, NRAS, PIK3CA genes at 5%, 1%, and 0.1% allelic frequencies (Horizon Discovery, 2021). The synthetic matrix contains 170bp fragmented DNA that originates from well characterized highly validated isogenic cancer cell lines that have been mixed together and diluted in a background wild type DNA (Amit, 2015) This allows for the detection of SNP's at varying allele frequencies.

SeraCare advertises the commutability of their reference material; that the size distribution, library yield and complexity are consistent with native ctDNA. SeraCare claims the use of their proprietary technology to fragment DNA allows for a more native cfDNA size distribution compared to other shearing methods (Ashford, 2017). Their reference material consists of 170bp fragmented mutated DNA in a background of well-

characterized GM24385 human genomic DNA. This is offered in a full-process plasma-like material, also known as a synthetic plasma matrix. The reference material is the "Seraseq ctDNA v2 reference material" and it offers a broad set of 40 actionable cancer mutations in 23 genes, along with a range of allele frequencies: 2%, 1%, 0.5%, 0.25%, 0.13%, 0%. Of the 40 mutations there are a variation of SNPs, Indels, and structural variants, allowing a variety of mutations to be assessed.

All three of these companies claim commutability and provide a variety of mutations at varying allele frequencies in a synthetic matrix to help researchers assess their ctDNA detection assays. Each is slightly different with the proprietary nature in which they are made and the variants that they allow you to assess (Table 1). It is important for laboratories to get an accurate measure of their assay's performance at low minor allele frequencies, which is why one wants their reference material to resemble native cfDNA in library yield and complexity. Reference material needs to be able to accurately reflect the number of unique molecules that may be found in a patient sample. Different fragmentation methods can compromise library complexity, making it critical to create reference material that reflects native ctDNA and not use DNA that is damaged from fragmentation (SeraCare, 2021). As previously mentioned, a good reference material is needed to validate an assay and make a diagnostic test a clinically relevant tool. ctDNA controls and finding a standard reference material is an emerging field, and there are a few different synthetic ctDNA controls that are now available for purchase. As reference materials for ctDNA detection become more available and frequent, it begs the question, which one should be used? Is one better than another? Does the reference material perform as advertised? Which allows for the most accurate assessment of your

assay? Choosing an appropriate reference material is an important decision to make when validating a ctDNA detection assay. This thesis will address and compare the aforementioned reference materials for their suitability to be used for ctDNA assay validation.

Full reference material from Anchor Molecular, Horizon, and SeraCare will be assessed for its ability to be used as a control material for ctDNA detection assays. Each vendor will be assessed on accuracy (whether all the introduced variants are detected at the allelic frequency expected), precision (reproducibility of the measured allelic frequency) and ctDNA concordance (library profiles match between control and ctDNA samples). Library profiles for each vendor will also be assessed based on yield and complexity. Library complexity is measured by the number of unique reads found through sequencing. Do the library yields and complexity closely resemble those of native cfDNA? How do they compare to each other? Characterizing these controls allows for informed decisions to be made on which control would be best for ctDNA detection test validation and provides insight into which controls can give clinical laboratories an accurate measure of the assay performance.

Chapter II.

Research Methods

Experimental Plan

This study evaluated three different full cfDNA reference materials. The accuracy and precision of the varying allele frequencies available from the three different vendors was compared. This study also included the examination of these reference materials from each vendor to native cfDNA libraries. This study is broken up into two parts; part 1 assessed each vendor (at four different percent AF's) on accuracy (whether all the introduced variants are detected at the allelic frequency expected), and precision (reproducibility of the measured allelic frequency), treating each as its own individual sample. Part 2 assessed the purchased controls' commutability with healthy cfDNA and compared DNA profiles of pooled healthy cfDNA to pooled purchased controls.

Materials and Methods

Ordering and Samples Preparation

Anchor Molecular (AM) sells 250ng/ml 5 ml aliquots of reference material at five different allele frequencies and includes both SNPs and structural variants. Horizon (HZ) sells 4 1ml vials (each 400ng/ml) of reference material at four different allele frequencies and only includes SNPs. SeraCare (SC) sells 125ng 5 ml aliquots of reference material at six different allele frequencies that include both SNPs and structural variants. AM

advertised controls at the varying allele frequencies (AF) were not available for purchase at the time. They did however have the materials to titrate out percent AF independently. For AM controls 3% AF tumor DNA in a matrix was ordered, along with wild type DNA in the same matrix. The AM 3% AF tumor DNA was diluted with wild type DNA to make each percent AF used. AM samples were also normalized to 125ng/ml using the AM matrix (DD plasma). Horizon reference materials came with each percent AF in 1ml vials with 400ng of DNA in each. In order to normalize the input of DNA for this control, a custom order was placed to receive vials of the matrix material. Horizon controls were normalize to 100ng/ml. See Table 2 for exact material and catalog numbers. Allele frequencies 0, 0.1, 1, and 2 were made/purchased from AM and SC. Allele frequencies 0, 0.1, 1, and 5 were purchased from Horizon. A list of all the variants included by these three vendors in their full reference materials can be found in Figure 1. All three vendors have EGFR (T790M), EGFR(L858R), EGFRDelE (746-A-750), and KRAS (G12D) mutations in their controls. Healthy DNA was from Research Blood Components and stored at –80 degrees Celsius in 10ml aliquots; all DNA used was healthy plasma from a single doner.

Workflow

Samples were broken up into part one and part two. See Figure 2 for workflow overview. For part one (Figure 2A), the samples for each allele frequency were prepped for extraction. Each sample was intended to be 125ng in 5ml – this is true for SC. Due to lab error, Horizon DNA input was 500ng in 5ml and AM's input was 625ng in to 5ml. For part two (Figure 2B), eight samples from one healthy donor were prepped for 5ml extractions and then pooled together right before library construction (LC). AM, SC, and

HZ samples for part two were pooled DNA from multiple different allele frequencies that were left over from part one extractions. For part one, each vendor had its own plate that was processed through sequencing. For part two, pooled healthy samples, pooled AM samples and pooled SC samples were all processed through to sequencing on one plate. It is to be noted HZ samples came in late and parts one and two were processed on one plate.

Extraction

Part1

For part one, three- 5ml extractions for all four allele frequencies were performed for each vendor. Samples were spun for 20 minutes at 7000 RPM. Supernatant was transferred to a new tube and the samples were digested with proteinase K. Next, the DNA was extracted using an automated bead-based extraction platform (MagMAX™ Cell-Free DNA Isolation Kit). Once the DNA was extracted, the DNA was quantified using cfDNA TapeStation from Agilent. Concentration and yield of each sample was recorded. Extracted DNA then went on to library construction. AM, SC, HZ and healthy DNA were each their own extraction, and extracted on separate days.

Part2

For part two, one healthy donor was broken up into eight- 5ml samples and run through the same extraction as the AM, SC, and HZ extractions in part one.

Library Creation

Part1

For part one, each vendor had its own library (LC) plate. Samples were normalized to 20ng inputs using elution buffer and split into two different LC repetitions. At this point, two process match controls were added to the plate, also at 20ng inputs. Once normalized, a cell free TapeStation was run and the actual amount that was put into LC was recorded. An automated NEBUltra II LC preparation method was used to make libraries. The adaptors and primers added included molecular barcodes. Next, constructed libraries went through 10 cycles of PCR and a generic SPRI clean-up. After clean-up, a PicoGreen assay was used to quantify the amount of DNA that was present. On top of the PicoGreen assay, a D1000 TapeStation from Agilent was run. Following quantification, hybrid capture (HC) was performed.

Part2

For part two, four pools of DNA were made, pooling together residual DNA from each vendor extraction. Pools consisted of multiple percent AF's. AM material, extracted SC material, extracted HZ samples, and extracted healthy DNA all made up their respective pool. Each pool of extracted DNA was split into nine LC samples that were normalized to varying inputs. Each pool had three reps at 30ng input, three reps at 20ng inputs, and three reps at 10ng inputs. The LC plate consisted of nine samples from each pool and three process match controls (PMC) that were at 30, 20, and 10ng inputs as well. The LC plate for part two went through the same LC methods as described for part one.

Hybrid Capture

Parts1 and 2

Plates for part one and plates for part two were normalized into HC bringing 2ug of DNA from LC. Samples were then concentrated, prepared for bait addition, and incubated for

24 hours with their baits. The baits used were a proprietary mix that targets a large range of relevant cancer genes. Once the DNA hybridized with the baits, the samples were washed using a series of buffers and put through 12 cycles of PCR. After PCR, the DNA was put through another generic SPRI bead clean-up and quantified again, using both a PicoGreen assay and D1000 TapeStation.

Sequencing

Parts1&2

Out of HC the samples from each LC plate were pooled together and made into a pool for sequencing. LC plates for both parts one and two went through the sample sequencing normalization and pooling. Before samples were pooled, each sample was normalized to 1nM. Once normalized and pooled, each pool was prepared for Illumina Nova Seq sequencing. A total of five different flow cells were run to achieve optimal coverage for the assay being performed. Sequencing data was processed using a custom bioinformatics pipeline. The pipeline outputs many Picard metrics that were used in analysis and also outputs the variants found in the samples and at what allele frequency they are found.

Chapter III

Results

Extraction

The first step of data collection in the process of creating libraries was extraction. The AM sample matrix was a yellow color and had a pellet after the first hard spin in extraction. SC and HZ sample matrices were both clear and without any sort of pellet or residue after the first hard spin in extraction. Mass input into and out of extraction was measured for all the vendors, not including the healthy samples and PMC's. The extraction efficiency (Figure 3) was calculated for each vendor. AM's and HZ's mean extraction efficiency were the highest at 57 and 58, respectively. AM and HZ also had very similar standard deviations of 4.8 and 4.3, respectively. SC had the most variation with a standard deviation of 29.5. No DNA could be extracted from the 0% AF for SC. The 0.1 AF had a large range of extraction efficiency. The mean for SC was 48.9% but individually the 2% allele frequency had the highest mean efficiency compared to the other vendors, with a mean around 78%. There was a slight increase in efficiency as percent AF increased for AM and SC.

Part 2

Library Construction and Hybrid Capture

In part 2, the extracted DNA was pooled together for each vendor and input into LC at 10, 20, and 30 ng inputs. Input mass into LC was measured and the percent cfDNA (Figure 4) was calculated for all three vendors, the healthy samples, and the PMC

samples. CfDNA was calculated using the Agilent cell free ScreenTapes that measures the DNA quantity between 50-200bp and compares it to the total amount of DNA in the samples. Almost all samples had a high amount of cfDNA at 85% or more. There was one set of samples that was lower than 85%; these were the 10ng input samples from AM. The vendors, PMC, and healthy cfDNA samples were all very similar in percent cfDNA. These samples being input into LC were run on a TapeStation showing the DNA profiles of each sample (Figure 5). Representative samples from the 20ng input group were used to represent the library profiles. SC (purple) and AM (orange) had traces that closely matched cfDNA (red) in terms of width of the distributions, whereas HZ (blue) trace more closely matched the PMC (black) samples in this regard. The mode of SC and HZ distributions most closely matched the mode of the cfDNA traces. AM had a slightly shorter mode. Libraries were put through PCR where the overall library profile was measured again (Figure 6). These traces similarly showed that SC and AM libraries most closely matched cfDNA libraries in terms of distribution width, whereas HZ libraries had similar width distribution to PMC libraries. The mode of the SC and HZ library peaks aligned closest to the cfDNA library mode. Overall yield for both LC and HC were measured (Figure 7) and no samples particularly stood out or had any problems.

Sequencing

All samples went through sequencing and Sequencing Analysis Viewer (SAV) metrics were assessed. SAV metrics look at the quality of the sequencing run. Common metrics are error rates and the percent of reads passing filter (%PF). Errors were below 1 and %PF was above 80%. Reads passing the SAV metric went into a bioinformatic pipeline. After, the pipeline samples were assessed in a few ways. First, the overall number of

reads was assessed (Figure 8). All samples had around or slightly under 200,000,000 reads. HZ and the PMCs obtained more reads due to being processed as a different pool and sequenced on a different flow cell. Insert size and the standard deviation of the insert size were also assessed (Figure 9a&b). AM expected fragment size was 125bp and HZ and SC expected fragment size was 170bp. SC samples seemed to be true to what was expected with a mean insert size of 173bp. AM were a little larger than expected, measuring at a mean insert size of 133bp and HZ measured lower than expected at 137bp. Healthy cfDNA mean insert size measured at 154bp. HZ, followed by SC and then AM, were the order of which was closest to the healthy cfDNA mean insert size. The standard deviation for each vendor's insert size was more closely matched for healthy cfDNA and SC, compared to AM and HZ. Healthy cfDNA and SC were close to identical with standard deviations for insert size at 34 and 35 respectively. AM's insert size standard deviations was 46, HZ was at 61 and PMC was at 81. The most variability in insert size was seen in the HZ samples. Unique coverage was the last metric that was assessed (Figure 10). SC, at the 3 different inputs, was the vendor that best mimicked the coverage seen in healthy cfDNA. AM and HZ were a bit lower and more similar to the unique coverage seen with the PMC.

Part 1

Variant Analysis

Out of the sequencing pipeline, variants were called for each sample. There were many variants found outside of the advertised incorporated variants in the controls. AM advertised ~52 specific mutations; in their sample only 45 were detected-the other 7 were not baited. SC advertised 40 and all 40 were present in the samples. HZ advertised 8

specific mutations and only 7 were present; the one missing was within the baits used. Only advertised variants were looked at for measuring accuracy and precision. Those variants were further filtered, taking out all variants that were below the LOD (Figure 11). The majority of the variants for the 0.1% AF for all the vendors were under the LOD and taken out for analysis, leaving SC with no variant calls at 0.1% AF. HZ was left with 1 variant call at 0.1% AF, and AM was left with 9 variant calls at 0.1 %AF. Only blue variants from Figure 11 were used. The mean for all the variants at each expected percent AF in ascending order were 0.18, 0.92, 1.80 for AM; 0.22, 1.16, 5.67 for HZ; and 0.17, 1.03, 2.03 for SC. Accuracy was assessed by plotting the observed percent AF against the expected certificate of analysis (CoA) percent AF for four common variants for each vendor (Figure 12**).** The CoA is a certificate from the manufacturers stating their product conforms to the requested requirements; in this case the CoA percent AF is the actual testing of each reference material lot that was purchased, saying the exact measure %AF for each prepared %AF for every variant. Pearson's coefficient of correlation values was obtained for each vendor using a standard y=x line. The Pearson's coefficient value for HZ was 0.98, SC was 0.98, and AM was 0.88, listed in descending order (Table 3). In part 1, precision was also assessed (Table 4)**.** The percent CV was calculated using the mean and standard deviation of all 6 repetitions for each vendor's percent AF. Each vendor and percent AF had three extraction repetitions, each with 2 LC repetitions. Percent CVs were only calculated with the four shared variants for each vendor. The SC percent AF's had the lowest range of CV's of all vendors, ranging from 7-17% across each of the four common variants. HZ had a very broad spread of CVs at 1% AF but a very narrow spread of CVs at 5% AF. At 1% AF, HZ's CVs ranged from 16-39% and at

5% AF, HZ's CVs ranged from 8-15%. AM had a similar profile to SC but a slightly higher range of CV's, ranging from 9-25%.

Chapter IV

Discussion

cfDNA and ctDNA are biomarkers with a tremendous potential in the diagnostic & prognostic world. With the increased utilization of cfDNA in a clinical setting, having adequate controls is vital. It is challenging to produce controls that mimic cfDNA in an end-to-end workflow (sample prep to sequencing). Anchor Molecular, SeraCare and Horizon controls were put through extraction, library construction, hybrid capture, and sequencing in an attempt to make a comparison of the controls and see if there is one that best models cfDNA behavior and preforms as advertised. Full reference material from Anchor Molecular, SeraCare, and Horizon were assessed for their ability to be used as a control material for cfDNA detection assays. Accuracy (whether all the introduced variants are detected at the allelic frequency expected), precision (reproducibility of the measured allelic frequency), and cfDNA concordance (library profiles match between control and cfDNA samples) were assessed for all samples.

When it came to ordering, Anchor Molecular and SeraCare were the most responsive, whereas Horizon took a little more effort due to needing to order a custom background matrix to dilute samples to desired DNA input. SeraCare was by far the easiest when it came to sample preparation; AM required more sample preparation because the advertised percent AF controls were not ready, so percent AF had to be made in-house. This unfortunately meant that there was also no certificate of analysis for each percent AF. Instead, the certificate of analysis for the stock 3% controls was used to calculate the expected values for the AM percent AF. When it came to extraction, no DNA was able to be extracted from the SC wild type (WT) control. Two rounds of the

WT control were ordered and both times no DNA was able to be extracted. Along with that, the 0.1 percent AF SC control had a few samples where little to no DNA extracted, but other samples did have enough DNA material to go into LC. Overall, the efficiencies out of extractions were similar – the means all hovered around 60% when outliers were removed.

Part 1:

Many variants were found for all three vendors. HZ was the only vendor that had one advertised variant that did not show up in any of the percent AF samples. Samples were pared down to ones that were advertised and above the limit of detection (Figure 11). Figure 11 is observed vs. expected but used the advertised expected values and not the measured CoA expected values. In general, all the mean percent AF hovered around the advertised expected percent AF but there were variable standard deviations with each percent AF. The advertised expected values were not used in vendor specific calculations; CoA expected values were used instead. Vendor accuracy was compared using the four variants that were in common across all three vendors (Figure 12). Using the Pearson's coefficient of correlation, all vendors had high correlation coefficients (Table 3). SC and HZ were the most accurate controls – Pearson correlation co-efficient of 0.98, whereas AM had a Pearson correlation co-efficient of 0.88. In assessing precision for those same shared variants, SC had the most precise extraction replications at both 2% AF and 1% AF (Table 4). However, HZ had the lowest percent CV values for their sample repetitions at 5% AF (compared to all other percent AFs across vendors). HZ 1% AF sample replicates had the highest CV values, making it the least precise. SC was the one with the lowest CV values for both 2% AF and 1% AF. AM had a similar

profile to SC but had just slightly higher CVs. As such, SC had the most precision across all percent AF, followed by AM and then HZ. HZ was only most precise at the 5% AF. When picking a control there needs to be consistent precisions across the different percent AFs. When looking at both precision and accuracy together, SC stands out the most, with a relatively strong accuracy and the best all-around precision.

Part 2:

To round out the comparison of these controls, a pooled set of AFs from each vendor was compared with its' likeness to native cfDNA. In part 2 we observed many different metrics to assess likeness. The first metric that was looked at was percent cfDNA, looking to see if there were any glaring differences between native and control samples when it came to their extracted DNA samples. Potentially here we would see the different backgrounds of the vendors be distinguished based on whether or not backgrounds were made with genomic DNA. We should be able to assess if the spiked in DNA mimics that of cfDNA. Horizon background is made with a mixture of variant cell lines that are fragmented and then put into a matrix (Horizon Discovery's, 2020). AM background is WT cell genomic DNA fragmented into cfDNA size. SC uses the well characterized cell line GM24385 DNA that is fragmented and sized along with the spiked in variants at the same size into the matrix. In all three vendor's extracted material one would not expect large amounts of genomic DNA since all vendors fragment their DNA to cfDNA size. In Figure 4 we see that the mean percent cfDNA for all vendors at various LC inputs is around 90% and matches what we see in our healthy sample. The 10ng AM input is an outlier and seems to have a slightly higher genomic DNA content than all the other vendors, as its' mean is around 83% cfDNA. Between all the vendors and the

healthy samples there was a slight increase in percent cfDNA as input increased. These extracted DNA samples were run on a TapeStation to look at the DNA profile (Figure 5). AM and SC mimicked healthy DNA closely. Whereas Horizon and the process match control (PMC), which is a mix of cell line like Horizon is, had very similar profiles that were distinctly different than the healthy cfDNA sample. SC and AM were not just a mix of cell lines – SC was a mix of a single cell line that makes up the background and synthetically altered DNA. AM has variants that are made from cell lines and synthetically made in a WT fragmented genomic background. Library construction and HC looked normal for all samples. There was nothing in particular that stood out in this part of the assay (Figure 7). Post library prep samples were profiled on a Tape Station (figure 6)**.** Similar to the post extraction profiles, AM and SC profiles matched the healthy cfDNA profile, looking almost identical, and the HZ and PMC profiles were similar and both very different from the healthy cfDNA profile.

The last three metrics that were assessed for part 2 were Picard metrics that come out of sequencing. Figure 8 shows the total reads that went to each sample. This is important to look at before looking at library uniqueness metrics because unique reads are directly affected by input amount, and in turn are affected by the number of reads that go to each sample. In general, every 1ng of sample leads to around 300x coverage, so with more material comes an increase in the number of molecules and more chance that they could be unique and with the more coverage there is more of a chance to have unique reads covered in sequencing. Only when all things are equal one can make a statement about unique coverage. In figure 8 HZ has more reads than the other vendors, and SC and cell free healthy DNA have the least number of reads, which we need to

consider when looking at our unique reads. Figure 10 shows the mean unique reads for each sample, and SC and healthy were the most similar with means of 5107 and 4641 reads. AM and HZ were lower with 2427 and 1934 reads, respectively. Despite the higher total reads for HZ, the unique coverage was the lowest. Whereas SC and healthy DNA had the lowest number of total reads but the highest unique coverage of these vendors. SC matched healthy cfDNA most closely at each of the three inputs, implying that SC libraries matched the complexity found in cfDNA libraries best. The last way libraries were assessed was looking at if the expected library insert size was obtained and how close it was to what we saw for healthy cfDNA samples**.** Figure 9a is the observed mean insert size plotted against the expected advertised insert size, with a dotted line that is y=x. AM and SC were closest to the expected than HZ was, with HZ being very far from expected. The top half of Figure 9b was very similar to 9a but in 9b each vendor was broken down by input, there were two reference lines showing expected insert values and healthy cfDNA was included. HZ was the closest to cfDNA size. The bottom half of the graph is showing the standard deviation of the mean insert size; SC and healthy again lined up having the least amount of spread in insert size, and AM had a slightly higher spread and HZ had the largest spread of insert size.

Conclusion

Looking at Table 5 it seems the SeraCare most consistently mimics healthy cfDNA overall, with more green in the SC vendor column – color darkness indicating closeness to healthy DNA samples for part 2. For part one of Table 5, SC also seems to have all around better accuracy and precision. For the accuracy row, the darker the color, the higher the correlation coefficient/the higher the accuracy. For the precision row, the

darker the color the lower the cv/the more precise. The AM vendor column in Table 5 is the next most colored column, making it the 2nd overall best control. HZ, with the least amount of color, comes in last of the controls when considering overall cfDNA commutability and overall accuracy and precision. With that being said, depending on the assay purpose and what aspects one is trying to control for, a different conclusion may be determined. While SC seems to have done best, there is something to be said for extraction of the WT DNA was not possible and is something that should be considered before making a choice on which controls to use for your assay.

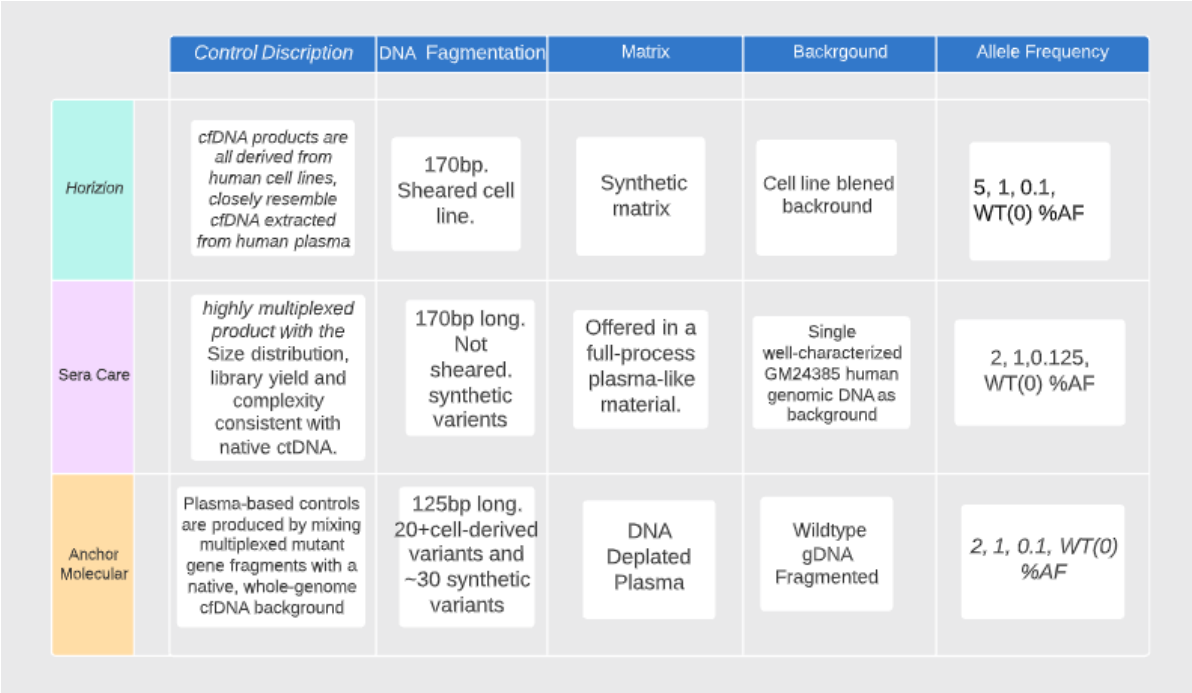
Limitations and Future Directions

This study had a few limitations, one being its size. To truly evaluate these companies' controls, a study with many more repetitions and matching samples run through other labs' NGS process would be ideal. It would also be good to perform a more in-depth analysis of the variants that each vendor has made. This study looked at the common variants in hopes that it would be a more "apples to apples" comparison but depending on what one is interested it might be beneficial to look into the accuracy of other variants more closely. In this study assessing commutability was not perfect – limitations come with making the native cfDNA profile itself. cfDNA yields are unique and can vary greatly person to person (Geeurickx & Hendrix, 2020), therefore yields are difficult to compare, whereas things like fragment size remain relatively consistent. But to make a good native profile with which to compare the reference material would require a very large data set allowing for a representative average. It would be valuable to also investigate the commutability of these controls with unhealthy, real patient samples. Lastly the expansion of different matrix types for cfDNA controls could add to this

comparison, for example adding human plasma as a matrix for DNA spike in. Despite this study's limitations, this is a good first step in assessing what is available on the market for full cfDNA reference material. There are many other areas of research to explore for cfDNA reference controls. Zhang et al (2017), as previously mentioned, has made spike in controls that can help distinguish germline alterations from somatic alterations; an interesting next step might be to explore if that innovation can be incorporated into a matrix, allowing the controls to be used all the way through the NGS process. The field of liquid biopsy and precision medicine is expanding quickly and well-characterized controls like these are critical to liquid assays' further use in diagnostics.

Appendix

Table 1. Advertised cfDNA Controls



	Control Discription	DNA Fagmentation	Matrix	Backrgound	Allele Frequency
Horizion	cfDNA products are all derived from human cell lines, closely resemble cfDNA extracted from human plasma	170bp. Sheared cell line.	Synthetic matrix	Cell line blened backround	5, 1, 0.1, WT(0) %AF
Sera Care	highly multiplexed product with the Size distribution, library yield and complexity consistent with native ctDNA.	170bp long. Not sheared. synthetic varients	Offered in a full-process plasma-like material.	Single well-characterized GM24385 human genomic DNA as background	2, 1,0.125, WT(0) %AF
Anchor Molecular	Plasma-based controls are produced by mixing multiplexed mutant gene fragments with a native, whole-genome cfDNA background	125bp long. 20+cell-derived variants and ~30 synthetic variants	DNA Deplated Plasma	Wildtype gDNA Fragmented	2, 1, 0.1, WT(0) %AF

Advertised description of purchased controls.

Table 2. Ordering

Vendor	Item Ordered	Supplier Part Number
Seracare	Seraseq ctDNA Ref Material v2 AF0.125%	0710-0207
	Seraseq ctDNA Ref Material v2 AF1%	0710-0204
	Seraseq ctDNA Ref Material v2 AF2%	0710-0203
	Seraseq ctDNA Ref Material v2 WT	0710-0208
Anchor Molecular	AM Normal female cfgDNA (250ng Spiked in 5ml Plasma, WT background))	60133502
	AM Multiplexed 50 ctDNA Standard 3%AF (250ng Sipked in 5ml Plasma)	60100102
	DD Plasma	60601001
Horizion	Multiplex I cfDNA in Synthetic Matrix II	HD917
	Synthetic Matrix II	HDX-CUST-RefStd

Table 2- list of all the parts order from the vendors

Table 3. Part 1 Correlation Coefficients

	AM	SC	HZ
Pearson's correlation coefficient	0.88	0.98	0.98

Table of Pearson's correlation coefficient for each vendor

Table 4. Sample Precision

%CV for each vendor and each %AF 6 LC replications												
%AF	0.1	1	2	5	0.1	1	2	5	0.1	1	2	5
EGFR_2369C>T	NA	25	10	NA	NA	16	NA	15	NA	12	7	NA
EGFR_2573T>G	NA	15	9	NA	NA	23	NA	8	NA	11	11	NA
KRAS_35G>A	NA	19	19	NA	NA	35	NA	7	NA	15	8	NA
PIK3CA_1633G>A	NA	12	18	NA	NA	39	NA	8	NA	13	17	NA
Vendor	AM				HZ				SC			

%CV were calculated using the 6LC reps for each percent AF. %VC were only calculated for each of the shared variants.

Table 5. Overall Conclusions

	Healthy	AM	HZ	SC	Conclusion
Part 1					
Accuracy		0.88	0.98	0.98	SC most accurate across % AF, HZ only accurate at its highest AF
Precision		16	19	12	These are the mean of precision; SC most precise, then AM, then Horizon
Part 2					
Extraction Eff	na	57.6	58.3	48.9	Unable to extract WT DNA from SC, otherwise SC was slightly more efficient
Extraction Traces		Similar to healthy cfDNA	Not similar in profile to cfDNA	Similar to healthy cfDNA	HZ did not match healthy CfDNA profile
%cfDNA	92.2	88.8	93.5	92.9	SC is the most similar to healthy; AM is low because the
Library Traces		Similar to healthy cfDNA	Not similar in profile to cfDNA	Similar to healthy cfDNA	HZ did not match healthy CfDNA profile
LC yelids(ng)	4382	3961	3409	4740	Yields are similar across vendors and input
HC yelids	329	342	273	316	Yields are similar across vendors and input
Total Reads(M)	166	180	223	164	AM and HZ have less total reads than healthy and SC; SC very similar to healthy
Unique Coverage	5108	2427	1934	4641	Despite lower coverage, SC and healthy have the most unique coverage and the means are very close. HZ has the least number of unique reads out of the 3
Insert size	154	134	137	174	HZ closest in size to healthy
Insert sizeStdDev	34	47	61	35	SC almost identical std dev to healthy; AM a little higher than healthy. HZ insert size std dev much higher than healthy

All results collected into one table. Part 2 color based on likeness to healthy cfDNA and part one colored based on the most accurate and the most precise.

Figure 1. Variant List

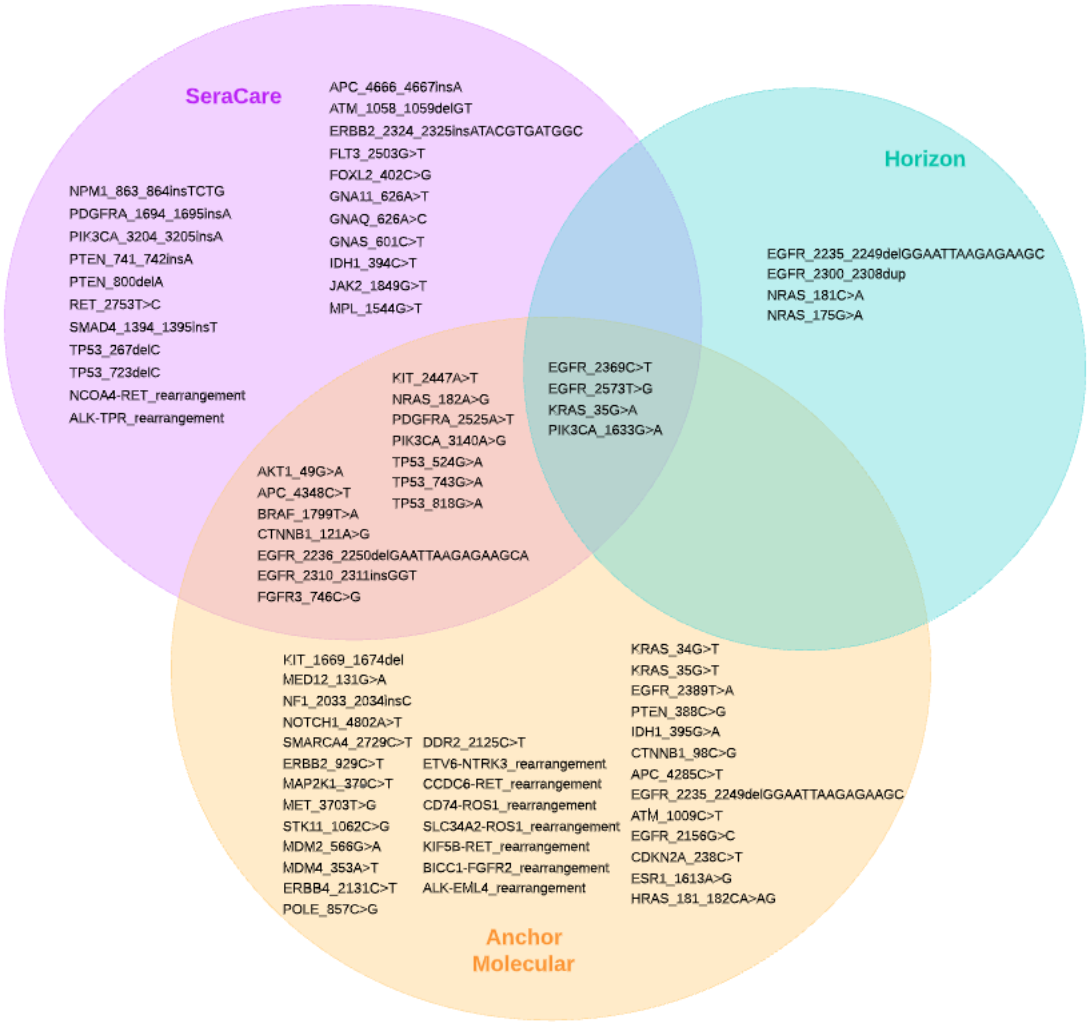
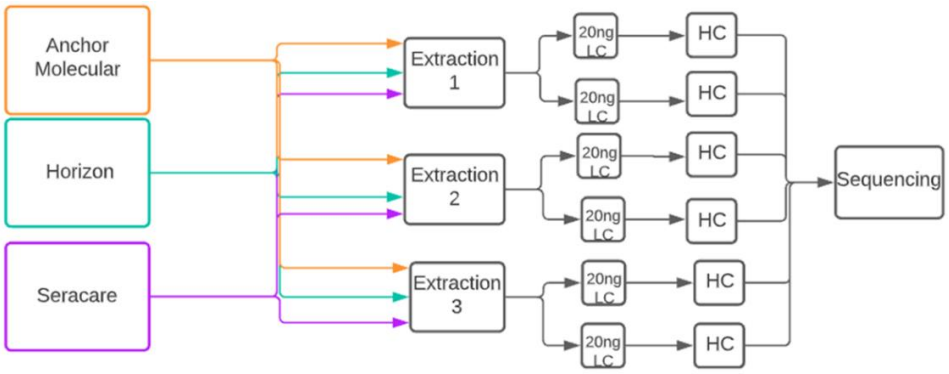


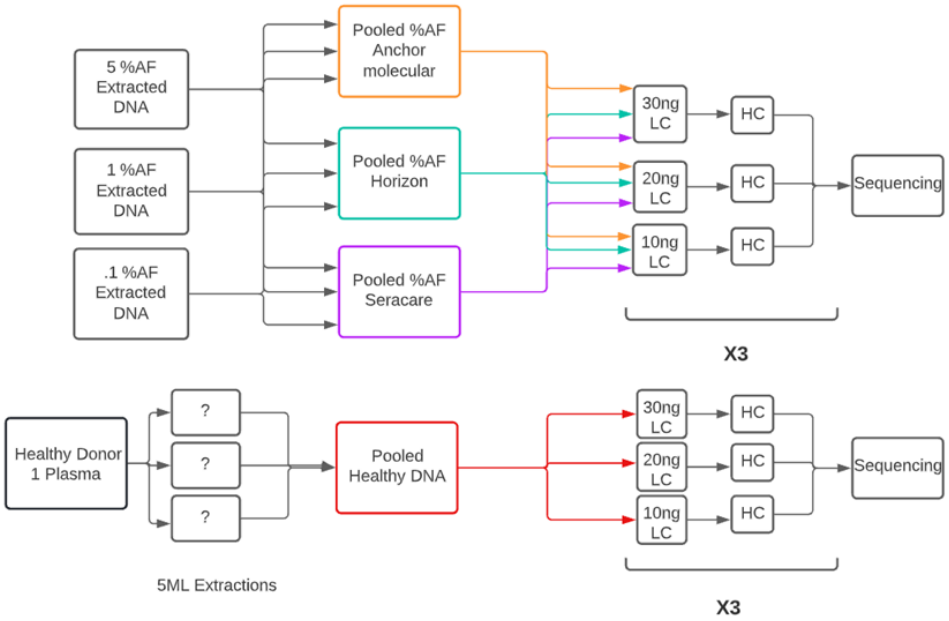
Figure 1 is a Venn diagram of all the variants that each vendor has in their controls. These variant lists are the protein changes.

Figure 2a. Part 1 Workflow



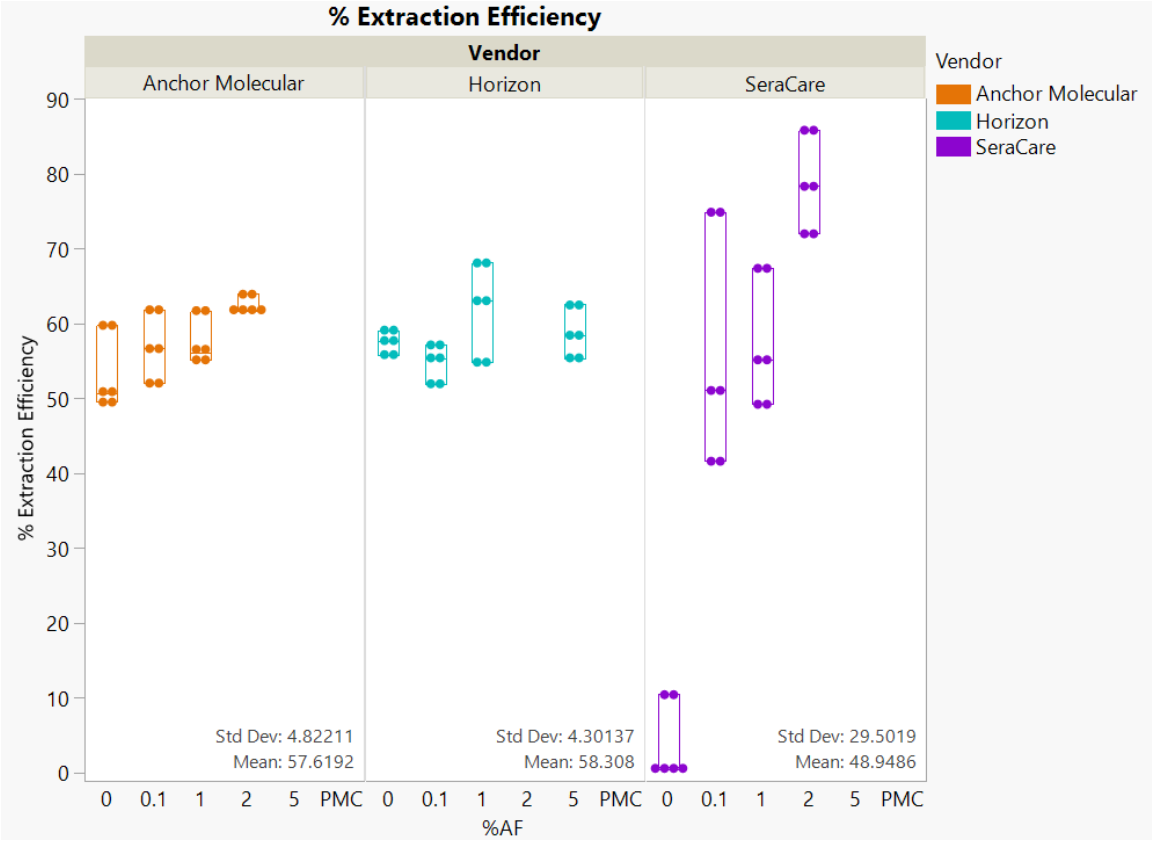
Part 1 workflow, each vendor has three extractions at each % AF making 6 LC reps for each %AF for each vendor.

Figure 2b. Part 2 Workflow



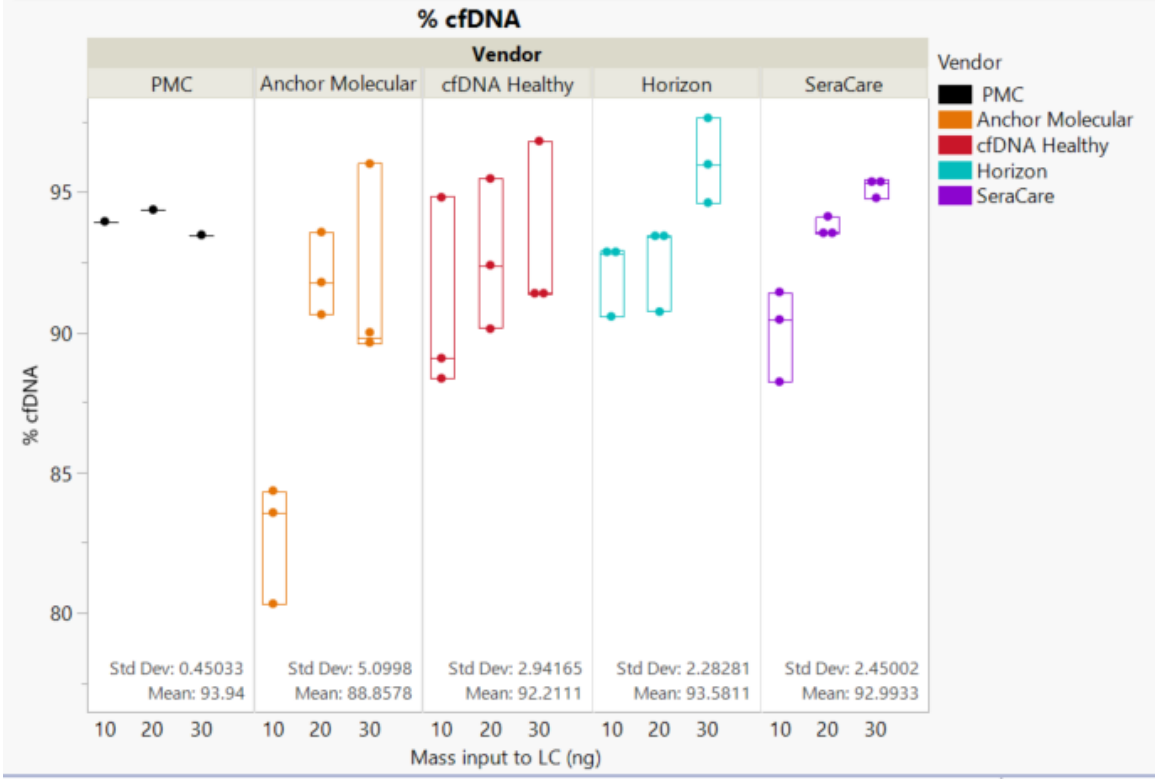
Part 2 workflow. Vendor DNA extracted in part one was pooled and put into LC at three inputs three times. Healthy DNA was extracted, pooled, and input into LC in the same way. 9 LC reps for each vendor and healthy samples.

Figure 3. Extraction Efficiency



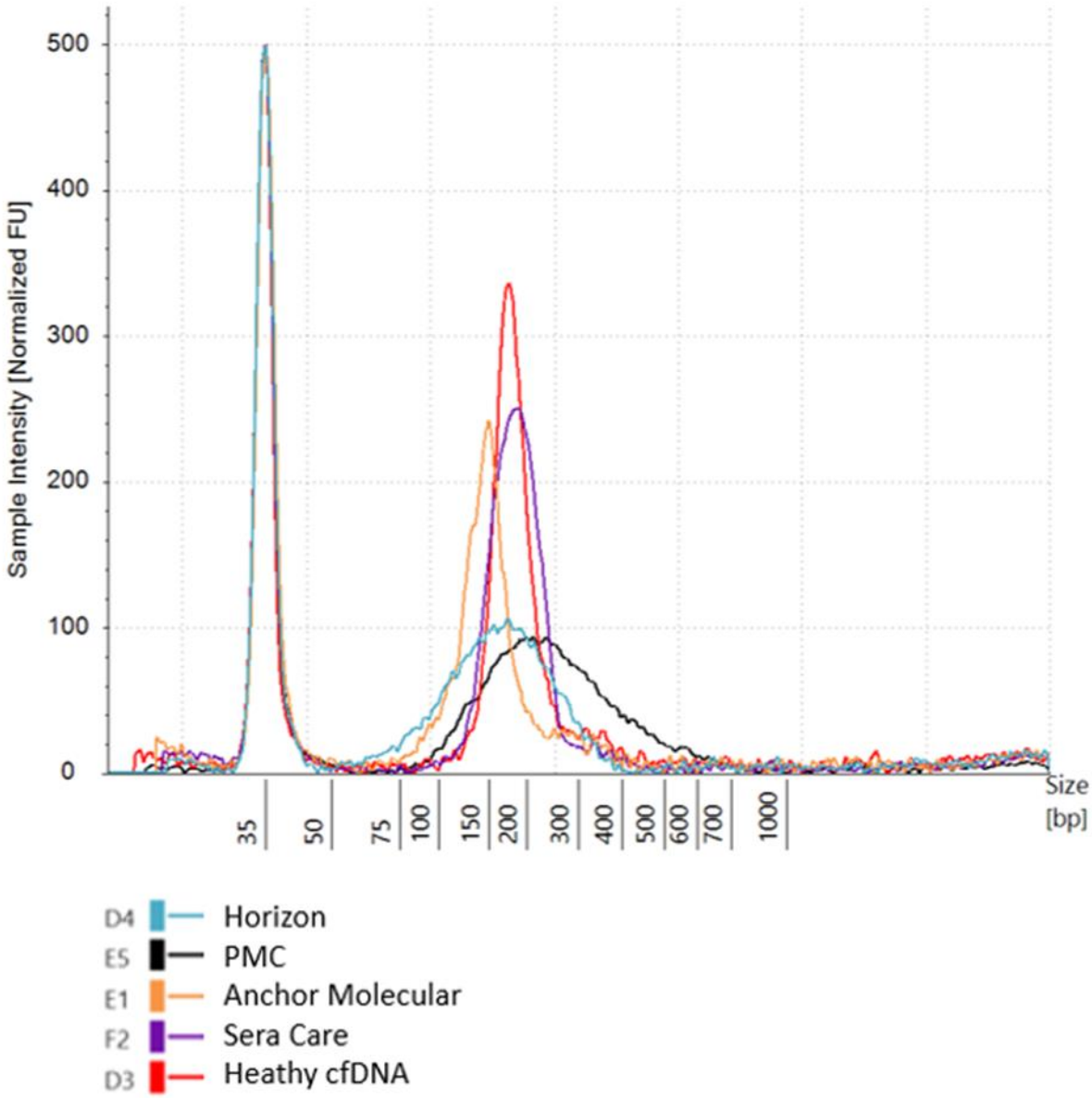
Extraction efficiency for each vendor by %AF

Figure 4. Percent cfDNA of Extracted DNA



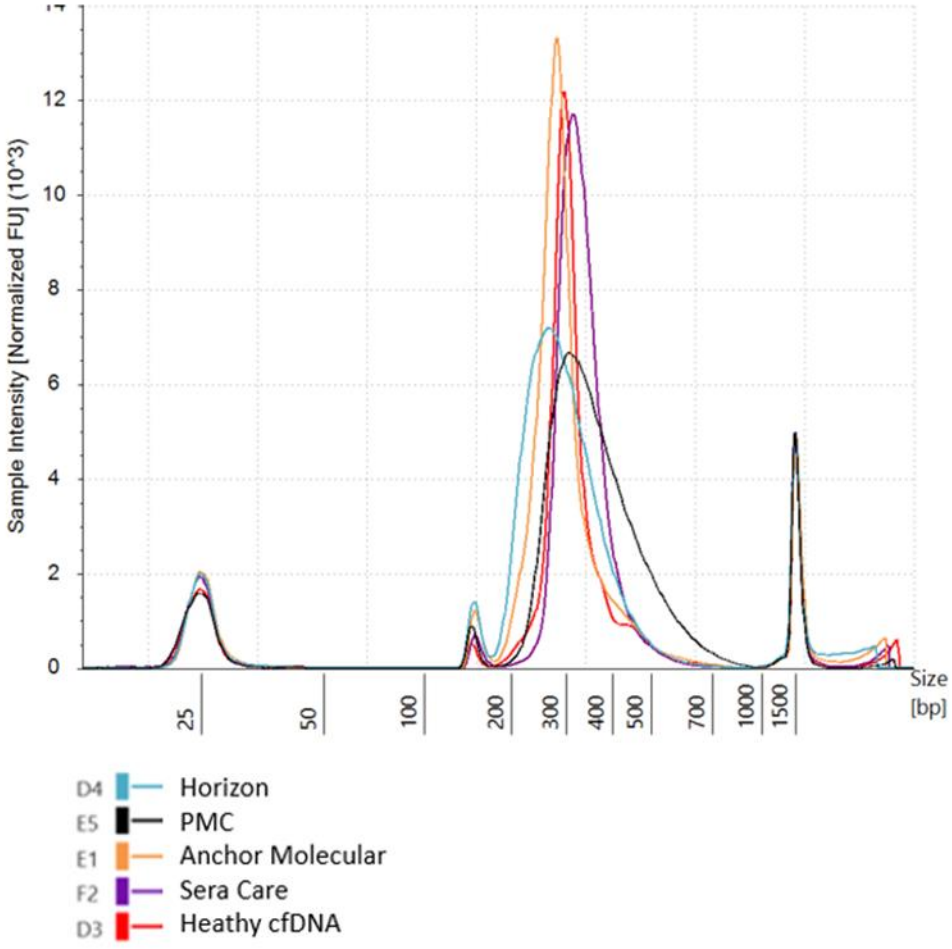
% cfDNA for each vendor at each input with mean and standard deviation calculated for each vendor.

Figure 5. Extracted DNA Profile



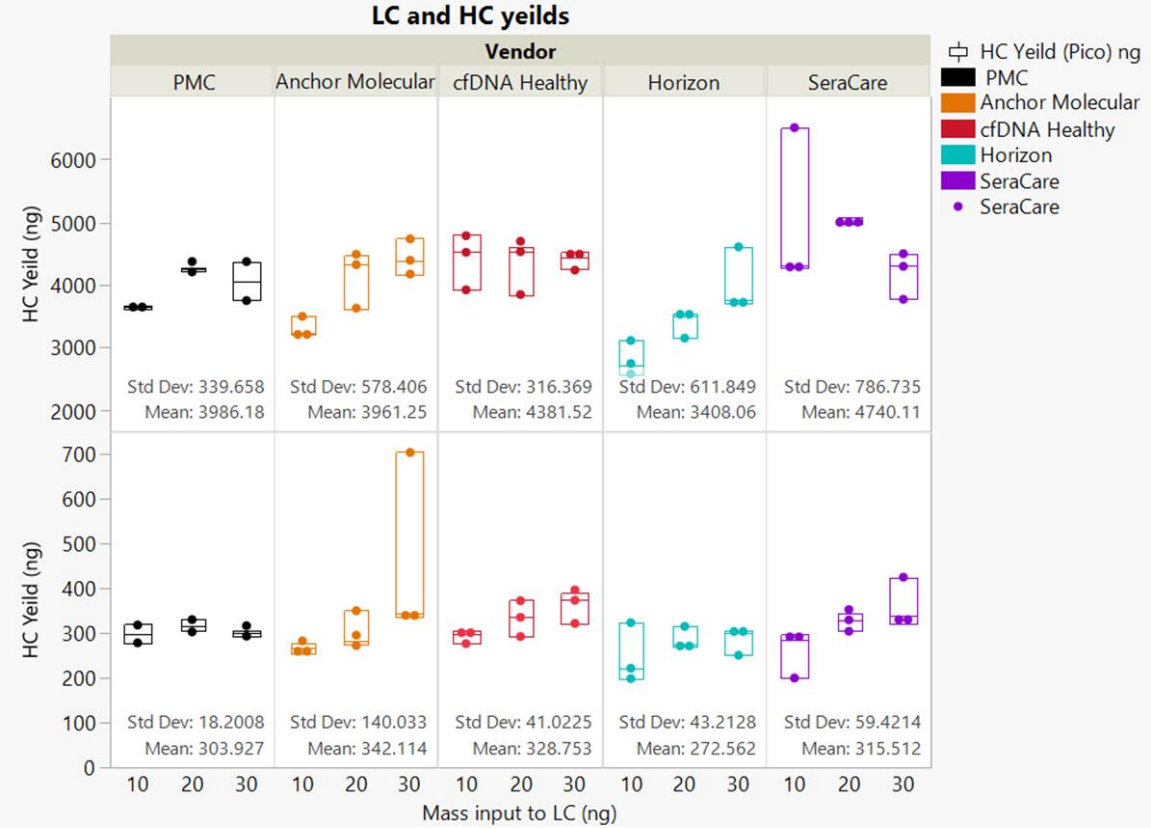
TapeStation trace for extracted libraries

Figure 6. LC DNA Profile



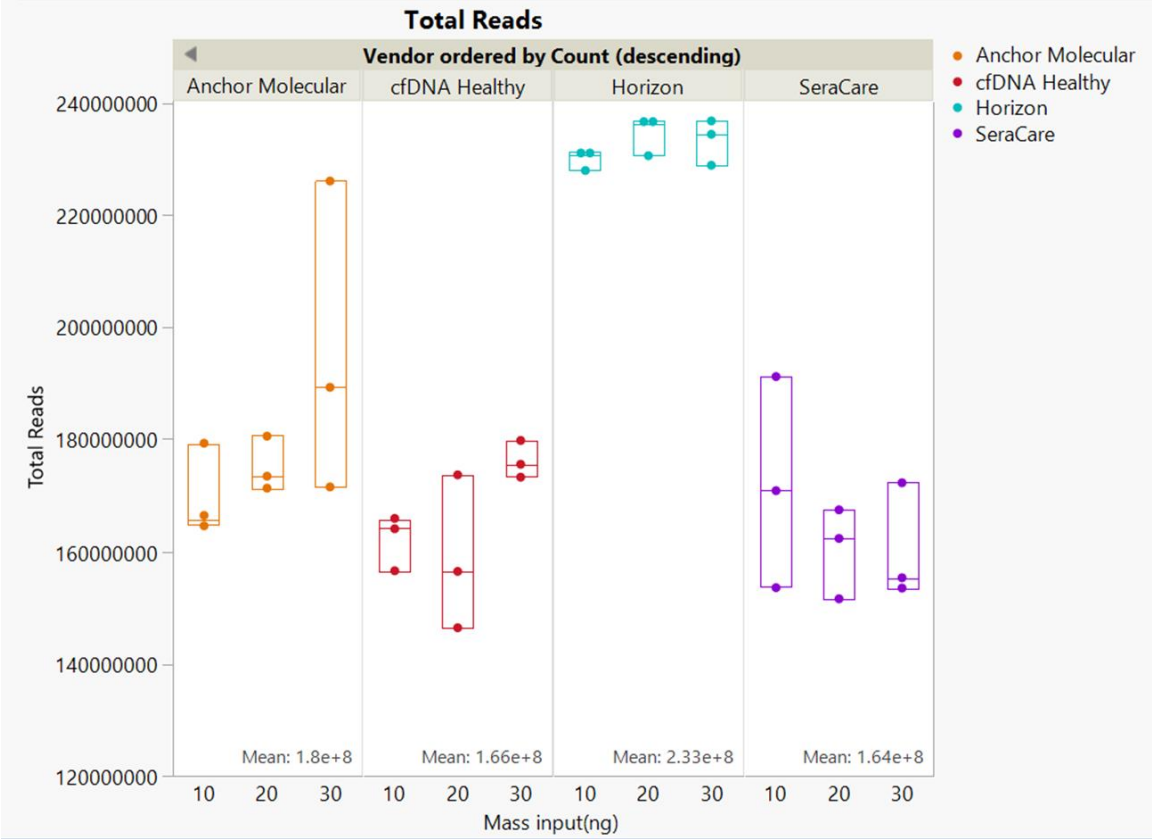
TapeStation trace for PCR'd libraries

Figure 7. LC and HC



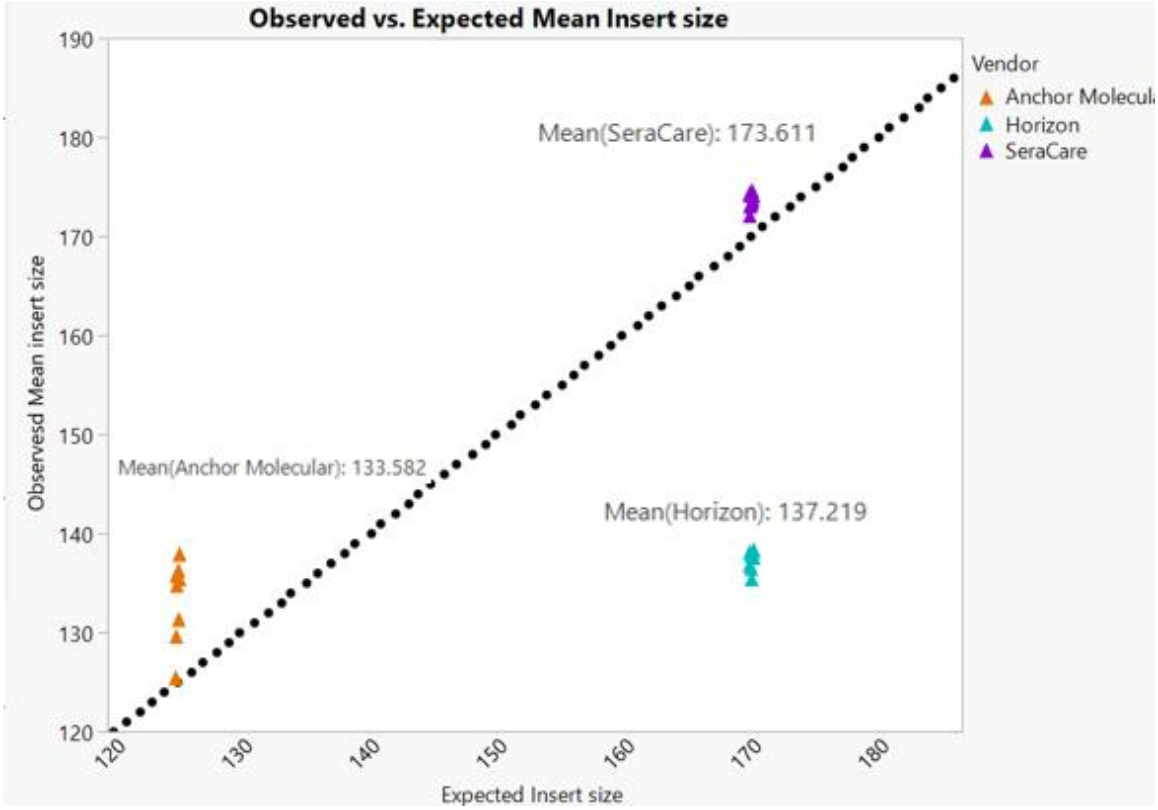
Top: LC yields for each vendor at every input. Bottom: HC yields for each vendor at every input. Mean and standard deviation per vendor calculated for both LC and HC.

Figure 8. Total Reads



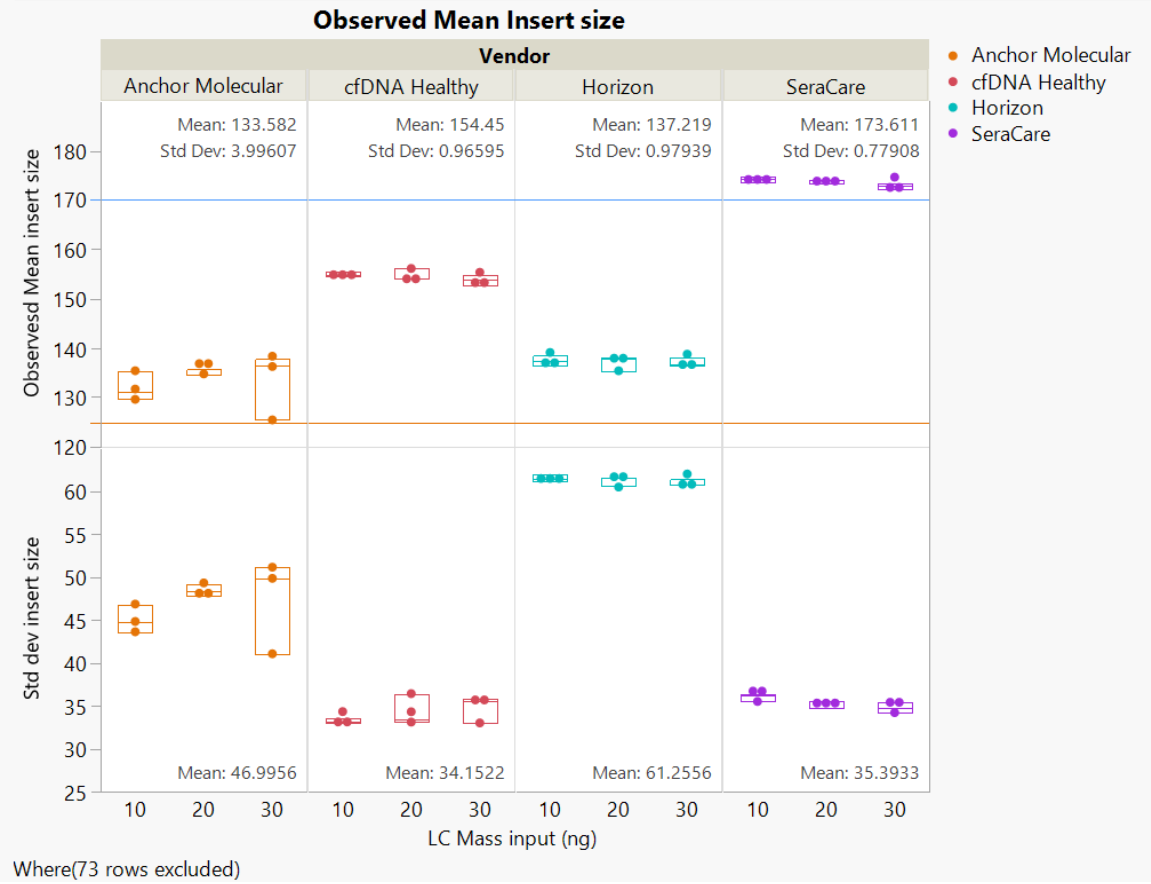
Total number of reads for each sample during sequencing

Figure 9a. Insert Size



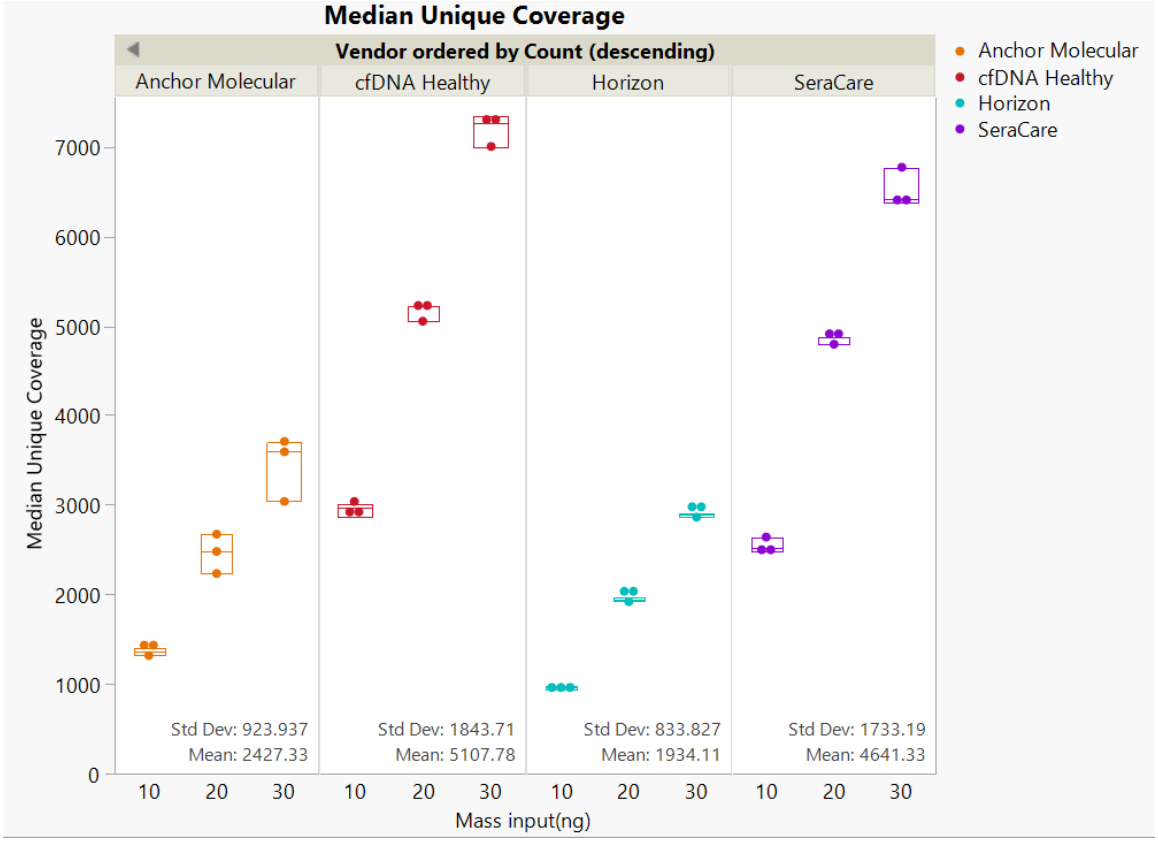
Observed mean insert size plotted against expected insert size for the three vendors

Figure 9b. Insert Size



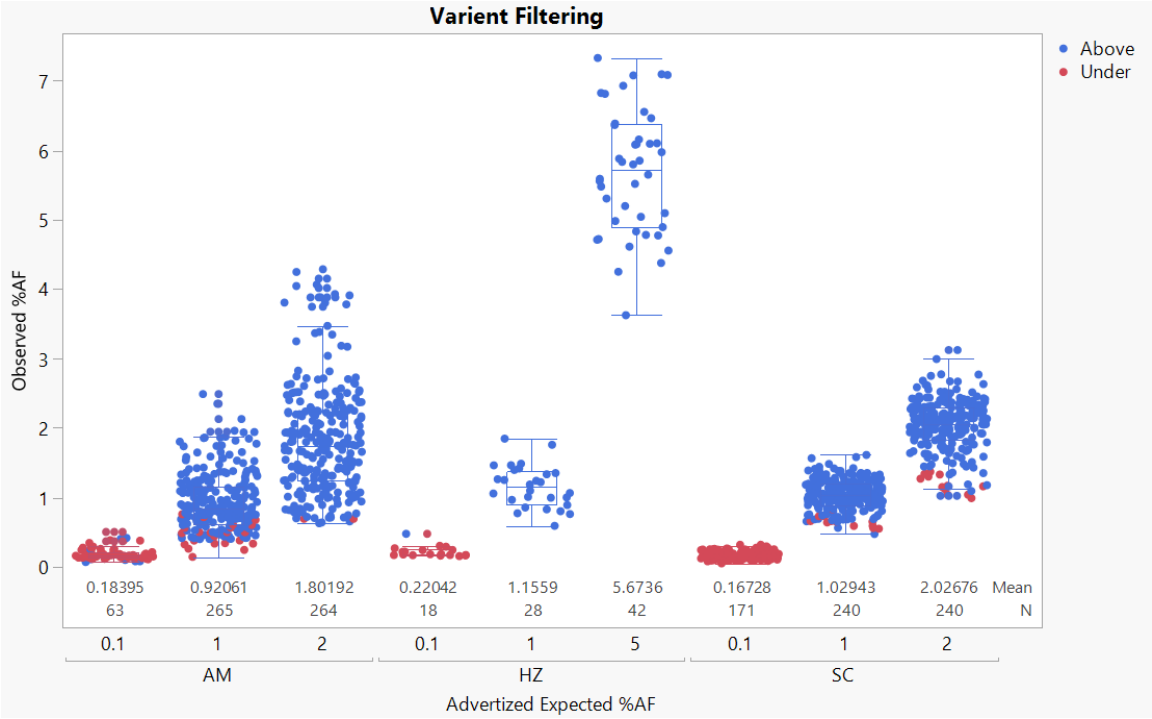
Top: Mean insert size for all three vendors and healthy cfDNA at each input, with reference lines at vendor expected insert size. Bottom: standard deviation of mean insert sizes for all samples at each input.

Figure 10. Unique Coverage



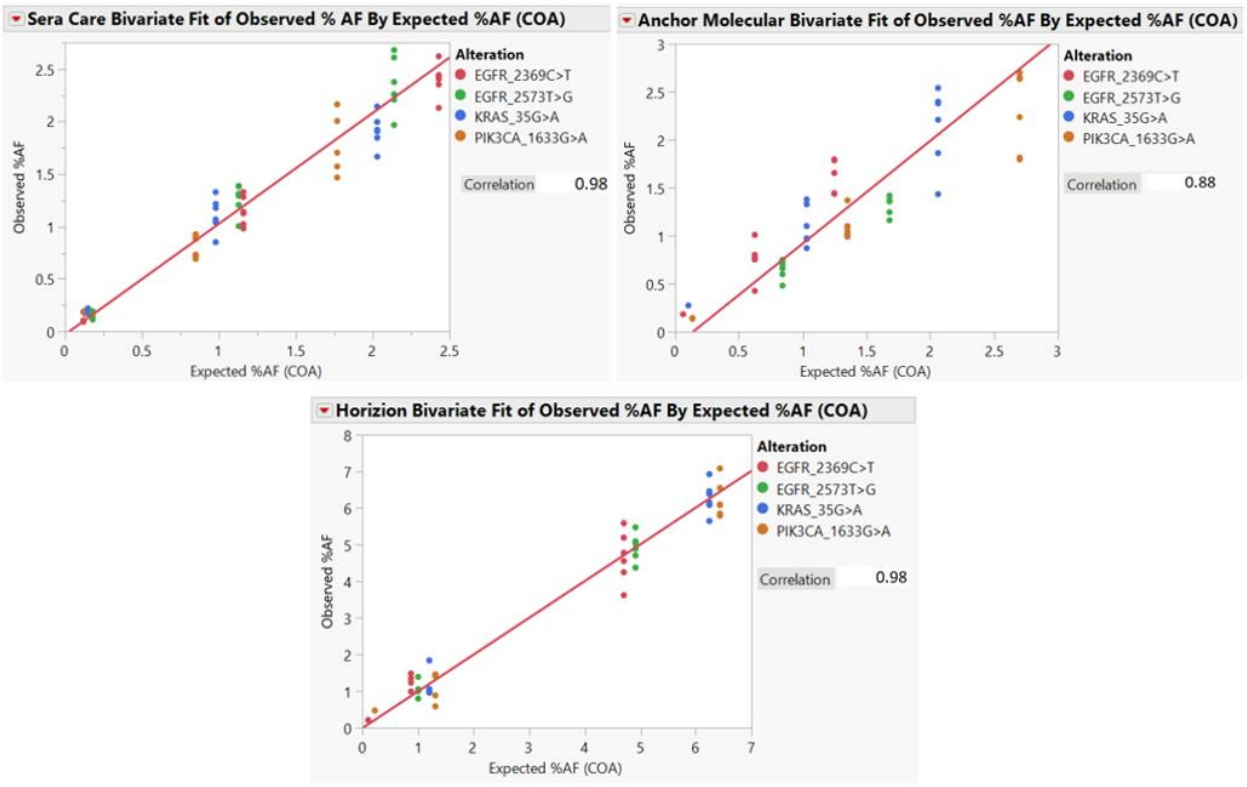
Median unique coverage for each vendor at every input

Figure 11. Variant Filtering



All observed advertised variants for each vendor at each % AF. Red are variants below LOD, blue are variants above the LOD.

Figure 12. Control Accuracy



Each graph is one vendor, each with the same set of variants. They compare the observed %AF vs. the expected CoA %AF. Each have a y=x line and its calculated Pearson's correlation coefficient.

References

Amit, H. W.-G. (2015). Using Cell Free DNA Reference Standards to Evaluate the Analytical Performance of Circulating Tumor DNA Testing and Solid Organ Transplant Health Surveillance. *BioTechniques, 4*, 59. doi:10.2144/000114346

Ashford, M. (2017). As Need for Liquid Biopsy Reference Materials Grows, Commercial Vendors See Opportunity. *genomeweb*. Retrieved November 09, 2020, from https://www.genomeweb.com/cancer/need-liquid-biopsy-reference-materials-grows-commercial-vendors-see-opportunity

Bronkhorst, A. J. (2019). The emerging role of cell-free dna as a molecular marker for cancer management. *Biomolecular Detection and Quantification, 100087*, 17. doi:10.1016/j.bdq.2019.100087

Bronkhorst, A., Wentzel, J., Aucamp, J., Dyk, E., Plessis, L., & Pretorius, P. (2016). Characterization of the cell-free DNA released by cultured cancer cells. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1863*(1), 157-165. Retrieved from https://doi.org/10.1016/j.bbamcr.2015.10.022

Burgener, J. M., Rostami, A., De Carvalho, D. D., & Bratman, S. V. (2017). Cell-free DNA as a post-treatment surveillance strategy: current status. *Seminars in Oncology, 44*(5), 330-346. doi:10.1053/j.seminoncol.2018.01.009

Cescon, D. W., Bratman, S., Chan, S. M., & Siu, L. L. (2020). Circulating tumor DNA and liquid biopsy in oncology. *Nature Cancer, 1*, 276–290. doi:https://doi.org/10.1038/s43018-020-0043-5

CLSI. (2010). *Characterization and Qualification of Commutable References Materials for Laboratory Medicine; Approved guideline.* CLSI Document EP30-A, Clinical and Laboratory standards Institute, Wayne.

Food and Drug Administration. (2018). *Considerations for Design, Development, and Analytical Validation of Next Generation Sequencing (NGS) – Based In Vitro Diagnostics (IVDs) Intended to Aid in the Diagnosis of Suspected Germline Diseases.* Retrieved March 14, 2021, from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-development-and-analytical-validation-next-generation-sequencing-ngs-based

Fortuna, G. M., & Dvir, K. (2020). Circulating tumor DNA: Where are we now? A mini review of the literature. *World Journal of Clinical Oncology, 11*(9), 723-731. doi:10.5306/wjco.v11.i9.723

Geeurickx , E., & Hendrix, A. (2020). Targets, pitfalls and reference materials for liquid biopsy tests in cancer diagnostics. *Molecular Aspects of Medicine, 100828*, 72. doi:10.1016/j.mam.2019.10.005

Goossens, N., Nakagawa, S., Sun, X., & Hoshida, Y. (2015). Cancer biomarker discovery and validation. Retrieved October 25, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4511498/

He, H., Stein, E. V., Konigshofer, Y., Forbes, T., Tomson, F. L., Garlick, R., & Cole, K. D. (2019). Multilaboratory Assessment of a New Reference Material for Quality Assurance of Cell-Free Tumor DNA Measurements. *The Journal of Molecular Diagnostics, 21*(4), 658-676 . doi:10.1016/j.jmoldx.2019.03.006

Heitzer, E., Haque, I. S., Roberts, C. E., & Speicher , M. R. (2018). Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nature Reviews Genetics, 20*(2), 71–88. doi:10.1038/s41576-018-0071-5

Horizon Discovery. (2021). *Multiplex I cfDNA in synthetic Matrix II certificate of analysis*. Retrieved from horizondiscovery: https://horizondiscovery.com/en/reference-standards/products/multiplex-i-cfdna-in-synthetic-matrix-ii#resources

Horizon Discovery's. (2020). *How Horizon Discovery's reference standards are made (and how they are right for your assay)*. Retrieved from Horizon Discovery's: https://horizondiscovery.com/en/blog/how-our-reference-standards-are-made#:~:text=Cells%20with%20genetic%20variations%20such,to%20produce%20our%20reference%20standards

Johansson, G., Andersson, D., Filges, S., Li , J., Muth, A., Godfrey, T., & Ståhlberg , A. (2019). Considerations and quality controls when analyzing cell-free tumor DNA. *Biomolecular Detection and Quantification, 17*, 100091. Retrieved October 04, 2020, from https://www.sciencedirect.com/science/article/pii/S2214753518300184

Khatami, F., & Tavangar , S. (2018). Circulating tumor DNA (ctDNA) in the era of personalized cancer therapy. *J Diabetes Metab Disord., 17*(1), 19–30. Retrieved September 28, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6154523/

Kraus, V. (2018). Biomarkers as drug development tools: discovery, validation, qualification and use. *Nature Reviews Rheumatology, 14*, 354–362. Retrieved from https://doi.org/10.1038/s41584-018-0005-9

Lathrop, T. J. (2015). Analytical Validation and Points for Discussion. *U.S. Food and Drug Administration*. Retrieved March 15, 2021, from https://www.fda.gov/media/88823/download

Lu, Y., Ooi, K., Shi, K., & Chan, E. (2020). Improving measurement of assay variation and Limit of Detection (LOD) of ctDNA assay by using validation samples contrived with DNA-free human plasma [abstract]. *American Association for Cancer Research 2020. 80(16 Suppl)*, p. Abstract nr 4272. Philadelphia: AACR. doi:10.1158/1538-7445.AM2020-4272

Newman, A. M., Bratman, S. V., To, J., Wynne, J. F., Eclov , N. C., Modlin, L. A., . . . Diehn, M. (2014). An ultrasensitive method for quantitating circulating tumor dna with broad patient coverage. *Nature Medicine, 20*(5), 548-554. doi:10.1038/nm.3519

Picard. (2021, March 16). Retrieved from Broadinstitute: https://broadinstitute.github.io/picard/#:~:text=Picard%20is%20a%20set%20of,%2FBAM%2FCRAM%20and%20VCF.&text=The%20Picard%20toolkit%20is%20open,and%20free%20for%20all%20uses

SeraCare. (2021). *Seraseq ctdna: A Breakthrough QC Technology*. Retrieved from SeraCare: https://digital.seracare.com/ctdna

Tsang, J. C., & Chan, K. C. (2017). Quality Materials for Quality Assurance in the Analysis of Liquid Biopsy Samples. *Clinical Chemistry, 63*(9), 1431-1432. doi:10.1373/clinchem.2017.276014

Valverde, J. L. (2005). Validation of analytical procedures: text and methodology Q2(R1). In *ICH harmonised tripartite guideline* (pp. 1-13). Madrid: Centre for Studies on European Pharmaceutical Law: International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use.

Zhang , R., Peng, R., Li, Z., Gao, P., Jia, S., Yang, X., & Li, J. (2017). Synthetic Circulating Cell-free DNA as Quality Control Materials for Somatic Mutation Detection in Liquid Biopsy for Cancer. *Clinical Chemistry, 63*(9), 1465-1475. doi:10.1373/clinchem.2017.272559