We are grateful for the Editorial Board of Nature Medicine and the Reviewers for the constructive criticisms provided, which have guided us in the design of additional experiments and analyses. This provided us with the opportunity of substantially strengthening our manuscript. Below please find a point-by-point response addressing the Reviewers’ comments.

**Referee expertise:**

**Referee #1: computational biology/ctDNA**

**Referee #2: clinical application of liquid biopsy**

**Referee #3: genomics in cancer diagnostics**

**Reviewer #1:**

**Reviewers' Comments:**

Remarks to the Author:

In this manuscript, an outstanding team reports use of circulating tumour DNA by deep sequencing of plasma cell-free DNA for the noninvasive detection of mutant DNA in subjects with advanced forms of 3 human malignancies (carcinomas of the breast, prostate, and lung), by using a case-control study of ~40 subjects from each of the 3 malignancies along with ~47 healthy adults from a single centre. A significant portion of the study is focused on the derivation of mutant molecules as relates to tissue of origin with special focus on somatic variants from clonal hematopoiesis (CH) as inferred from sequencing of matches leukocytes. This latter effort aims to show that some variants of unknown source (VUSo) are somatic mutations deriving from tumours, and some from CH.

The manuscript is well-drafted and well-structured, and as additional notable strengths, the approach is notable for a significant degree of technical refinement, as well as a largely standardised workflow when considering pre-analytic and analytic variables relevant to ctDNA. However, when compared to recently published studies on cfDNA sequencing in patients with advanced malignancies (eg Bettegowda et al 2014 and others using for example the Guardant360 assay), this manuscript is clearly limited in terms of the breadth of tumour types analysed and disease stages considered. Separately, when considering the CH-dedicated portion of this study, the work is more limited than a prior study on the same topic (Ref #26). As such, despite the strengths noted above, the key described results seem to lack significant novelty, and it is difficult to appreciate in this study a significant advance at either a conceptual, biological, or technical level as might be expected for Nature Medicine. Separately, the authors do not seem to provide a useful set of source data or analytical tools as might represent a broadly accessible resource to the community at large. This latter impression is in part because the description of the methods, as well as the provided supporting data, were challenging to interpret as detailed below. Overall, these impressions along with the list of major and minor concerns outlined below significantly hamper enthusiasm for the manuscript for the current venue and its broad readership. The following list of critiques are provided as hopefully constructive feedback as might improve the manuscript for another journal.

Major comments:

1. The main new finding as highlighted within the abstract appears to be that the majority of cfDNA mutations are most likely derived from leukocytes due to clonal hematopoiesis (CH) and that CH is a "pervasive" biological phenomenon. While confirming prior findings is of significant value, this observation is not a new one especially because CH is very common with age, “trending towards inevitability” when considering somatic variants in circulating leukocytes (Zink et al 2017 Blood). Seeing that circulating leukocytes are the dominant source of cfDNA as demonstrated by several studies, finding CH-derived variants masquerading in cfDNA seems hardly surprising. More specifically, as cited by the authors (Ref #26), a prior study of 259 healthy adults (more than their 124+47 combined) previously found that the majority of these subjects those had ≥1 nonsynonymous mutation (>75% for those older than 50) in the plasma cfDNA, and that most of these somatic variants were present in matched blood leukocytes. That same study also showed that many mutations were passengers and not classic CHIP mutations. The authors themselves state twice that their findings are "consistent with" Ref 26 thus acknowledging that this result it is not novel.

Authors: We thank the Reviewer for the most pertinent of questions. The Reviewer is correct that clonal hematopoiesis (CH) has been described, and that non-synonymous mutations have been identified in circulating cell-free DNA (cfDNA) of >75% of cancer patients. Our study, however, provides several novel aspects to the characterization of CH in healthy individuals and cancer patients, given the unprecedented approaches employed, namely the depth of sequencing attained in the cfDNA and white blood cell (WBC) samples analyzed coupled with the use of unique molecular identifiers (UMIs), the large ‘genomic footprint’ assessed and the bespoke hierarchical Bayesian model to reduce errors in sequencing data. This novel approach allowed us to go above and beyond the observations made in Liu *et al.* (PMID: 30475948). Here, we demonstrate that although CH correlates with age, it can be detected in a substantial proportion of young patients and is almost invariably found in patients with advanced cancers, an aspect not investigated in Liu *et al.* given that in their study, the focus was only on healthy individuals. In addition, the average depth of WBC sequencing achieved in Liu *et al.* (~406X) does not allow for a fair comparison with our study, where both cfDNA and WBC samples were sequenced with the same approach at comparable depths (deduplicated raw target sequencing depth >60,000X). Second, our results demonstrate objectively and based on direct evidence, the importance of matched WBC sequencing as >50% of the mutations identified in cfDNA of cancer patients originate from CH. Third, we provide direct evidence to demonstrate that cfDNA sequencing without taking into account the results of WBC sequencing can be misleading; this is perhaps best exemplified by mutations affecting ‘cancer genes’ (e.g. *TP53*) which can be present in tumor-derived cfDNA and also be part of CH. Fourth, we provide evidence that the assessment of the repertoire of somatic mutations in cfDNA post-therapy (in the context of disease monitoring) can be confounded by CH-associated mutations, and that these post-therapy alterations preferentially affect specific genes.

Our study is novel as it challenges one of the current paradigms in the field of cfDNA analysis. Cohen *et al.* (PMID: 29348365) stated “there must be a limit on the number of bases queried in the test because the more bases queried, the more likely that artifactual mutations would be identified, reducing the signal-to-noise ratio”. Here, we demonstrate and validate with orthogonal methods that this statement is not necessarily correct. If sequencing is performed with UMIs and a robust error correction method is applied to the sequencing, a relatively large genomic footprint (2Mb) can be employed for the detection of tumor-derived mutations, but this can only be accurately achieved at present if WBC sequencing is performed concurrently.

Given the technological advancements presented in our study and the shifts to the current paradigms in cfDNA sequencing and data analysis our manuscript provides, we would contend that although some elements of our manuscript are confirmatory, collectively the results are novel and would constitute an important contribution to both the deployment of cfDNA sequencing as a tool for early diagnosis and disease monitoring, as well as to the characterization of CH in WBC and cfDNA samples.

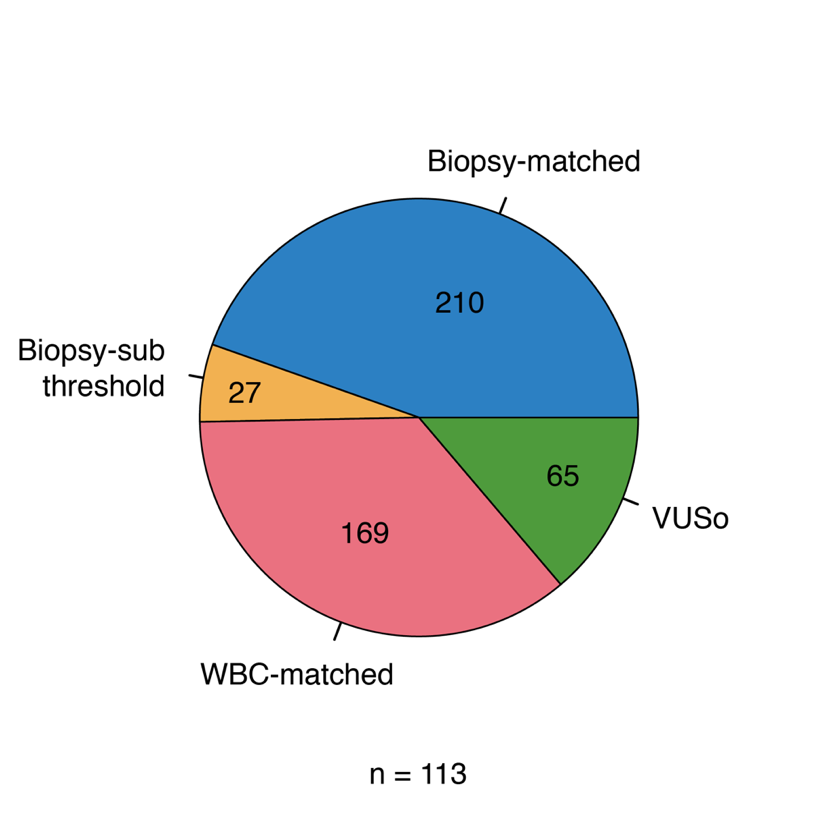
To clarify the novelty aspects of our manuscript in relation to Liu *et al.*, we have added the following passage to page xx, paragraph xx: “xxxx here add some of the differences between our study and Liu et al. xxxxx”.

2. If one were to ignore item #1 above, then sequencing of cfDNA from 124 patients with metastatic malignancies does not seem novel as there have been several papers presenting data from thousands of patients. More specifically, the clinical importance of this study as is not clear in the subjects profiled. Currently, in the metastatic setting, cfDNA sequencing assays such as Guardant 360 and Foundation One are used to genotype patients for clinically actionable mutations to aid in the selection of targeted therapies without the sequencing of matched leukocytes. None of the results presented here indicate that the sequencing of matched leukocytes is necessary for these assays. If the authors were to show that some portion of “actionable mutations” identified in the cfDNA of patients or controls arise as a result of CH that could be clinically relevant. It is worth noting that prior studies have already suggested this, such as the Oxnard group (Hu et al. 2018 Clinical Cancer Research), however a more systematic study across covering a more diverse panel of clinically actionable mutations would be useful.

Authors: We appreciate the Reviewer’s comment and sought to provide the systematic analysis of a more diverse panel of clinically actionable mutations in the patients analyzed in our study. We would like to emphasize, however, that to the best of our knowledge, this is the first study of concurrent ultra-deep sequencing of cfDNA and matched WBC; hence, the conclusions the Reviewer correctly drew based on previously published studies are correct, but we hope the Reviewer would concur that they are based on circumstantial rather than direct evidence. As mentioned in the response to Comment #1, here we demonstrate objectively and based on direct evidence, the importance of matched WBC sequencing as >50% of the mutations identified in cfDNA of cancer patients actually originate from CH. This novel and important observation is not of mere academic interest, as this finding illustrates one of the technical pitfalls of the majority of currently available, either commercial or academic, cfDNA assays such as Guardant G360 or FoundationOne Liquid, do not utilize matched WBC sequencing and despite the formidable computational and bioinformatics approaches employed to filter out the `noise`, somatic mutations related to CH or other forms of somatic mosaicism may still be included in the reports.

To highlight this issue, we restricted our analysis to the 73 genes included in the Guardant G360 assay (**Response to Reviewers Figure 1**). In 113 of the 114 non-hypermutated patients, at least 1 mutation detected in cfDNA is CH-derived whilst in total, 169 of 471 (35.9%) somatic mutations detected in the cfDNA of non-hypermutated patients were also found in WBC and were not detected in the matched tumor, indicating that these mutations could have been erroneously reported as tumor derived.

We also annotated the WBC-matched cfDNA somatic mutations using the MSK OncoKB knowledgebase (PMID: 28890946) to identify the pathogenic or likely pathogenic somatic mutations and provide the level of evidence for actionability of the WBC-matched alterations. As depicted below, in both cancer patients and healthy controls, we identified likely pathogenic WBC-matched cfDNA alterations. Considering the allele frequencies of these mutations, we expect almost all of these mutations to have been falsely reported as somatic tumor-derived cfDNA variants if matched WBC sequencing had not been performed.

****

**Response to Reviewers Figure 1: Distribution of the different categories of cfDNA somatic mutations.** The list of somatic cfDNA variants were restricted to those occurring in any exon of the 73 genes included in the Guardant G360 assay. The colors indicate WBC-matched, tumor biopsy-matched, biopsy-subthreshold variants, and VUSo. 114 non-hypermutated patients were included in this analysis. One breast cancer patient did not have any reported somatic cfDNA variant within the 73 genes.

**Response to Reviewers Table 1 (Supplementary Table xx of the Revised mutations)**: WBC-matched variants with highest level OncoKB annotation detected in cfDNA of cancer patients.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Patient ID | Gene | HGVSp | cfDNA depth | cfDNA alt count | cfDNA VAF (%) | WBC depth | WBC alt count | WBC VAF (%) | OncoKB highest level in the respective cancer type |
| MSK-VB-0058 | *ATM* | E503\* | 6705 | 11 | 0.164 | 3806 | 4 | 0.105 | 4 |
| MSK-VB-0058 | *ATM* | V1866Lfs\*54 | 4015 | 4 | 0.100 | 3005 | 2 | 0.067 | 4 |
| MSK-VB-0058 | *NF1* | L2395Ffs\*27 | 5116 | 17 | 0.332 | 3060 | 12 | 0.392 | 4 |
| MSK-VB-0063 | *NF1* | L925\* | 7942 | 9 | 0.113 | 4197 | 5 | 0.119 | 4 |
| MSK-VB-0067 | *PIK3CA* | C90R | 2548 | 18 | 0.706 | 3588 | 19 | 0.530 | 1 |
| MSK-VL-0028 | *NF1* | L2023\* | 7085 | 188 | 2.65 | 3778 | 76 | 2.01 | 4 |
| MSK-VL-0028 | *NF1* | I679Dfs\*21 | 6466 | 44 | 0.68 | 3821 | 43 | 1.13 | 4 |
| MSK-VL-0035 | *ATM* | D1278Tfs\*6 | 3050 | 6 | 0.197 | 2934 | 13 | 0.443 | 4 |
| MSK-VL-0064 | *NF1* | S637Vfs\*51 | 5731 | 8 | 0.140 | 3869 | 7 | 0.181 | 4 |
| MSK-VP-0001 | *NF1* | W1831\* | 6741 | 42 | 0.623 | 4881 | 32 | 0.656 | 4 |
| MSK-VP-0009 | *ATM* | C1502Afs\*5 | 3782 | 75 | 1.98 | 3676 | 98 | 2.67 | 4 |
| MSK-VP-0045 | *ATM* | Q1098Rfs\*11 | 3151 | 4 | 0.127 | 2503 | 6 | 0.240 | 4 |

**Response to Reviewers Table 2**: WBC-matched variants with highest level OncoKB annotation detected in cfDNA of healthy controls.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Patient ID | Gene | HGVSp | cfDNA depth | cfDNA alt count | cfDNA VAF (%) | WBC depth | WBC alt count | WBC VAF (%) | OncoKB highest level in any cancer type |
| W044216563529 | *ATM* | E713Gfs\*18 | 1582 | 8 | 0.506 | 2619 | 14 | 0.535 | 4 |
| W044216563537 | *NF1* | R1968\* | 4447 | 6 | 0.135 | 3795 | 2 | 0.053 | 4 |
| W044216563576 | *ATM* | R3008C | 5061 | 7 | 0.138 | 4190 | 3 | 0.072 | 4 |
| W044216563917 | *NF1* | E1266Dfs\*19 | 3324 | 7 | 0.211 | 4779 | 8 | 0.167 | 4 |
| W044216564621 | *KIT* | R634W | 3436 | 97 | 2.82 | 2104 | 72 | 3.42 | 1 |

To address these important points the Reviewer made, we have added the following passage to page xx, para xx, “In addition, without the filtering of WBC-matched mutations, CH-related mutations could have been possibly reported as somatic tumor-derived mutations (Supplementary Table X)”.

3. Although the elegant bioinformatics pipelines employed for error suppression and variant calling referenced in Figure 1 and associated Supplemental Figures seem reasonably described, the technical performance of these methods are insufficiently assessed, and overall metrics such as the global error rate of their assay seems conspicuously absent. More importantly, the authors do not report analyses or show figures illustrating how their error profiles uniquely inform the genotyping problems at hand, and how these compare with prior approaches such as TEC-Seq, Guardant, CAPP-Seq, SafeSeqS, etc. A flow chart showing the effect of the different steps of their error suppression pipeline, and how the error profile and total number of mutation calls is reduced with each step would seem important to better illustrate the utility of the pipeline.

Authors: We thank the Reviewer for highlighting this important omission in the original version of the manuscript, and for the opportunity of addressing these important points. Although we provided a full description of the probabilistic model used for joint calling of variants from cfDNA and WBC together with the procedures used for parameter estimation in the Online Methods under sections “Machine learning error model” (on page 30 of the original version of the manuscript) and “Joint variant analysis using the machine learning error model” (on page 33 of the original version of the manuscript), below we provide a more detailed description, including the relevant aspects of the analysis where the technical performance was assessed. In this discussion and unless otherwise stated, the terms `variants` and `mutation` refer to SNVs. The extensions brought to the model to account for indels are described in the Online Methods of the manuscript under the sections cited above (on pages 31 and 32 of the Online Methods). All empirical measurements of depth were obtained from raw pileups without base alignment quality (BAQ) filtering.

At each genomic coordinate, , the expected alternate allele depth has a baseline noise rate where is the total depth at and is the allele frequency. The model assumes that is drawn from a Poisson distribution such that:

Each was assumed to be drawn from a mixture of Gamma distributions indexed by , the parameters of which varied based on discrete covariates such that:

The covariates included were: (1) whether was uniquely alignable, (2) whether was located within a known segmental duplication and (3) the trinucleotide context. These covariates were found to be indispensable given that the error rate can vary as a function of the coordinate , the type of mutation and its mutational context. **Response to Reviewers Figures 2a** and **2b** below show the distribution of the mean posterior estimates of () across a representative subset of sites by type of substitution and trinucleotide context.

Typical workflows for detecting variants from single tumor samples and/or matched tumor normal pairs seldom include site-specific priors and if they do, it is usually through ad hoc filters for putative mutations detected in panels of normal samples (PMID: 23396013). Here instead, the posterior distribution of obtained through Markov chain Monte-Carlo resampling provides site-specific error rates with non-zero estimates at sites where no actual alternate allele counts were observed across samples in the training set. **Response to Reviewers Figure 2c** of this response shows the scatter plot of site-specific mean posterior estimates against the observed for samples in the training set.

The posterior distribution of each was approximated to a Gamma distribution and translated to a negative Binomial predictive posterior for the observed number of alternate allele counts at each and each alternate allele such that:

where and are the rate and dispersion parameters. This provides the basis for recalibrated quality scores for each observed alternate allele at each such that:

where is the Phred-scaled quality score and is the alternate allele depth **Response to Reviewers Figure 2d** below compares the estimated to empirical probability of observing an event. The recalibrated quality scores provide an objective measure to define thresholds for calling variants. Using the mean and variance estimates of obtained by training the model on 43 healthy control individuals, we compute the site- and allele-specific quality scores comparing the variants thus detected against the matched tumor biopsy. **Response to Reviewers Figure 2e** shows the recall rate per cancer type against the mean number of variants detected in healthy control individuals at different thresholds. By definition, scores are expected to allow one false positive per million bases and retained almost all biopsy-matched variants.

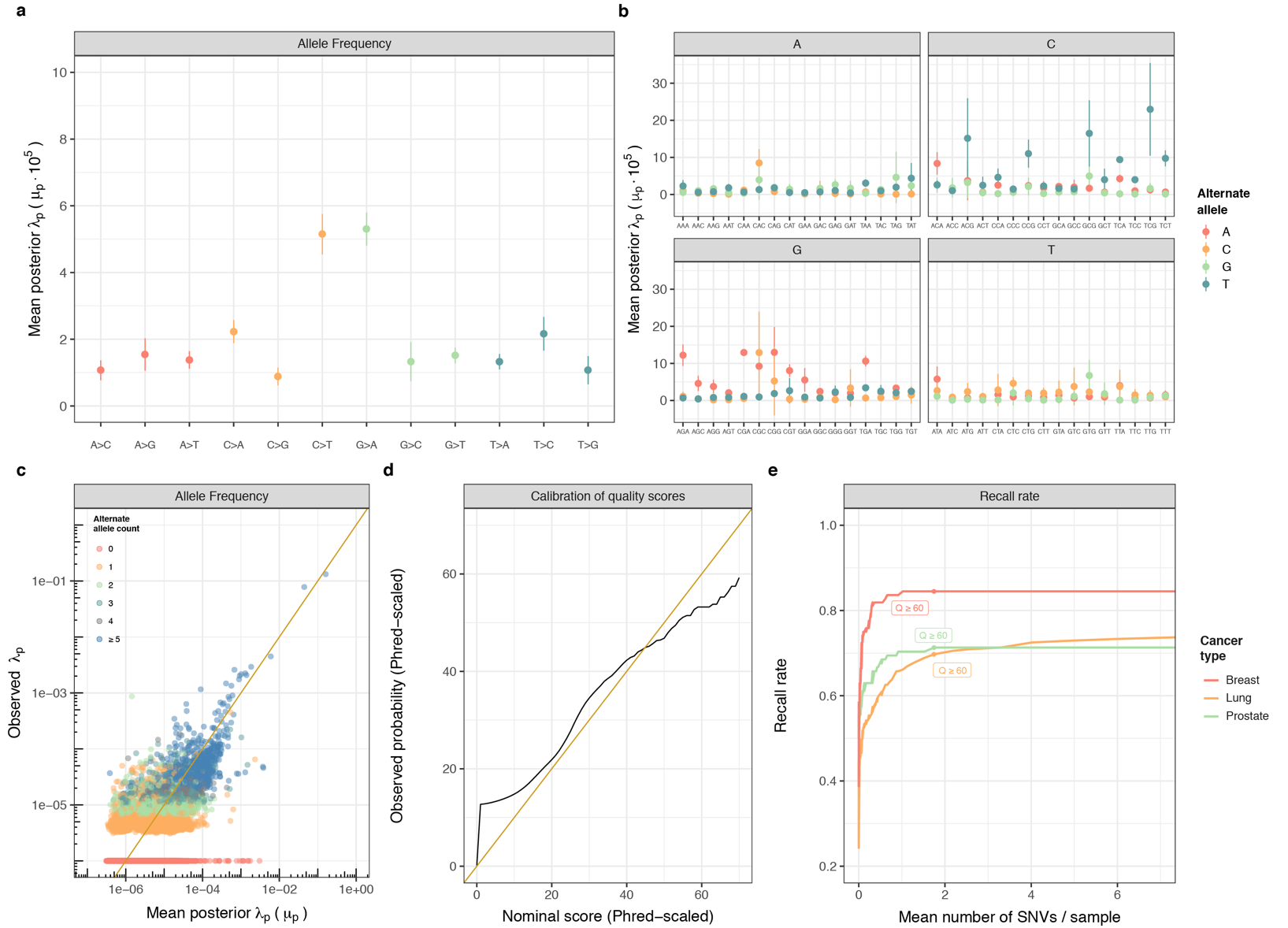
The hierarchical Bayesian model described above was applied to the raw variants scored both in cfDNA and WBC where the observed alternate allele depths, were assumed to correspond to the sum of negative Binomial noise and Poisson distributed signal. Formally:

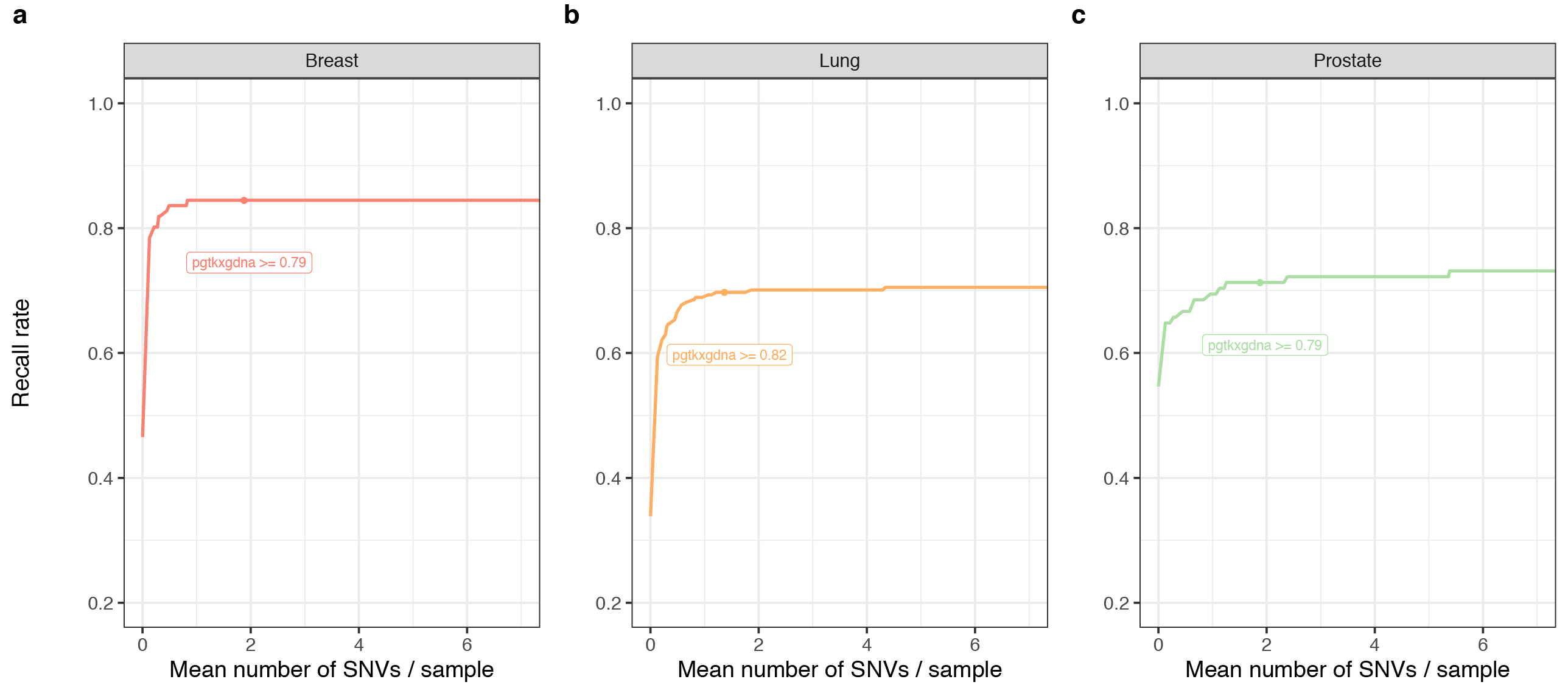
with an equivalent model for WBC. The likelihood of the observed pair of and was computed conditional on the cfDNA and WBC total depths, noise parameters and alternate allele fractions. Combining the joint likelihood with a uniform prior on the alternate allele fractions allows one to assign the source of origin of the variant such that:

where describes the posterior probability of assigning a detected variant to cfDNA. The procedure used for choosing is provided in the supplementary methods under section “Joint variant analysis using the machine learning error model”. The posterior probability allows the model to be further optimized. **Response to Reviewers Figure 3** illustrates the cancer specific trade-offs between the recall rate of biopsy-matched variants and allowing additional SNVs to be assigned to cfDNA in healthy control individuals at fixed . The particular thresholds were obtained through cross-validation by holding out each cancer type and selecting the threshold retaining most biopsy-matched variants whilst filtering out most variants of potential hematopoietic origin.

In summary, candidate variants in cfDNA and WBC generated through *de novo* assembly using the de Bruijn graph are assigned a quality score based on site-specific error rates and a posterior probability of originating from cfDNA based on the joint likelihood of the alternate allele depth in WBC. Using cross-validated cancer specific thresholds on the quality score and the posterior probability of assignment to cfDNA, the number of putative somatic variants is reduced to <2 non-synonymous variants per healthy control individual. **Response to Reviewers Tables 3-6** summarize the final number of variants whereby driver variants are defined as cancer-specific hotspot mutations and mutations considered to be pathogenic or likely pathogenic (PMID: 28890946).

Several groups have explored the space of high accuracy error corrected sequencing. Integrated Digital Error Suppression (iDES) (Newman *et al.*, PMID: 27018799) computed a per-base error rate which optimally balanced error suppression with molecular depth as 2x10-5. The Safe-Sequencing System (Safe-SeqS) (Kinde *et al.* PMID: 21586637) reported 0.9x10-5 supermutants (likely errors) per base pair. Both reports are similar to the per base error rate of 1x10-5 to 3x10-5 reported in our study (See Online Methods). Additionally, Lanman *et al.* (PMID: 26474073) and Phallen *et al.* (PMID: 28814544) both point out that beyond per base error rates, effective filtering for false positives is necessary; both report stringent filtering on small, highly curated panels to produce no false positive mutation calls in 1.56x106 bases (20 samples with a panel size of ~78Kb) attempted, and fewer than 1 false positive mutation call per 3x106 bases (38 samples with a panel size of ~80kb) attempted, respectively, in a population of relatively young individuals. This report included a panel that was more than ten-fold larger than the reports noted above, with 1.5 variant calls per age-matched control sample in approximately 1x106 bases per experiment.

**Response to Reviewers Figure 2: Estimation of error rates and performance assessment of the hierarchical Bayesian model.** The posterior distributions of site-specific were summarized by its mean and displayed for a subset of representative sites in (a) by type of mutation and (b) by trinucleotide context. Panel (c) shows the estimated against the observed allele frequencies for samples in the training set. Note the data points at the bottom are all with non-zero mean posterior and zero observed alternate allele counts. Panel (d) compares the estimated probability of observing an event (x-axis) with the actual empirical probability of observing such an event (y-axis). The plot is calibrated based on estimates of on chromosome 21. Note the initial sharp rise reflects the number of sites with zero alternate allele counts observed whilst the excess low probability events at the other end reflects the difficulty of stringently filtering out rare biological events such as clonal hematopoiesis. Panel (e) shows the mean number of variants detected in healthy control individuals (x-axis) against the recall rate of biopsy-matched variants (y-axis) for the different cancer types. At , one expects one false positive per million bases. Here, to exclude potentially CH-derived variants, we use a fixed threshold of 0.8 on the posterior probability of detected variants originating from cfDNA (i.e. ).

**Response to Reviewers Figure 3: Performance characteristics of WBC filtering.** The panels show the mean number of variants detected in healthy control individuals (x-axis) against the recall rate of biopsy-matched variants (y-axis) at different probabilities for allowing variants to be assigned to cfDNA for (a) breast, (b) lung and (c) prostate cancers. The thresholds displayed were obtained by cross-validation holding out each cancer type and selecting a threshold which retains most of the biopsy-matched variants whilst still filtering out variants of potential hematopoietic origin. Here, to exclude variants potentially due to noise, we use a fixed threshold of .

**Response to Reviewers Table 3**: Mean number of variants per sample (hypermutators included)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cohort | Mean no. of candidate SNVs | Mean no. of cfDNA SNVs ≥ | Mean no. of cfDNA SNVs ≥ WBC-filtered | Mean no. of cfDNA non-synonymous SNVs ≥ WBC-filtered |
| Breast | 295.4 | 77.08 | 54.48 | 28.44 |
| Lung | 233.2 | 34.15 | 18.17 | 9.81 |
| Prostate | 204.3 | 29.16 | 12.73 | 7.58 |
| Healthy | 222.9 | 14.13 | 2.47 | 1.43 |

**Response to Reviewers Table 4**: Mean number of variants per sample (hypermutators omitted)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cohort | Mean no. of candidate SNVs | Mean no. of cfDNA SNVs ≥ | Mean no. of cfDNA SNVs ≥ WBC-filtered | Mean no. of cfDNA non-synonymous SNVs ≥ WBC-filtered |
| Breast | 191.3 | 23.28 | 9.91 | 5.7 |
| Lung | 233.2 | 34.15 | 18.17 | 9.81 |
| Prostate | 201.3 | 28.09 | 11.85 | 7.02 |
| Healthy | 222.9 | 14.13 | 2.47 | 1.43 |

**Response to Reviewers Table 5**: Positive percent agreement of biopsy-matched variants in cfDNA (hypermutators omitted)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cohort | Candidate SNVs | cfDNA SNVs ≥ | cfDNA SNVs ≥ WBC-filtered | Non-synonymous cfDNA SNVs ≥ WBC-filtered |
| Breast | 0.84 | 0.84 | 0.84 | 0.84 |
| Lung | 0.79 | 0.73 | 0.72 | 0.72 |
| Prostate | 0.77 | 0.76 | 0.74 | 0.74 |

**Response to Reviewers Table 6**: Positive percent agreement of biopsy-matched driver variants in cfDNA

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cohort | Candidate SNVs | cfDNA SNVs ≥ | cfDNA SNVs ≥ WBC-filtered | Non-synonymous cfDNA SNVs ≥ WBC-filtered |
| Breast | 1 | 1 | 1 | 1 |
| Lung | 0.73 | 0.64 | 0.64 | 0.64 |
| Prostate | 1 | 1 | 1 | 1 |

To reflect the discussion above, we have added on page xx, paragraph xx of the revised manuscript, “Our ‘high intensity’ sequencing assay was found to have a favorable error rate (1x10-6 base-pairs sequenced) as compared to other high-fidelity sequencing assays applied to cfDNA analysis (i.e. Digital Error Suppression (iDES), 2x10-5; Safe-Sequencing System (Safe-SeqS), 0.9x10-5 supermutants (likely errors); GUARDANT G360, <1 error per 1.56x106 base-pairs sequenced; targeted error correction sequencing (TEC-Seq), <1 false positive in 3x106 base-pairs sequenced)”.

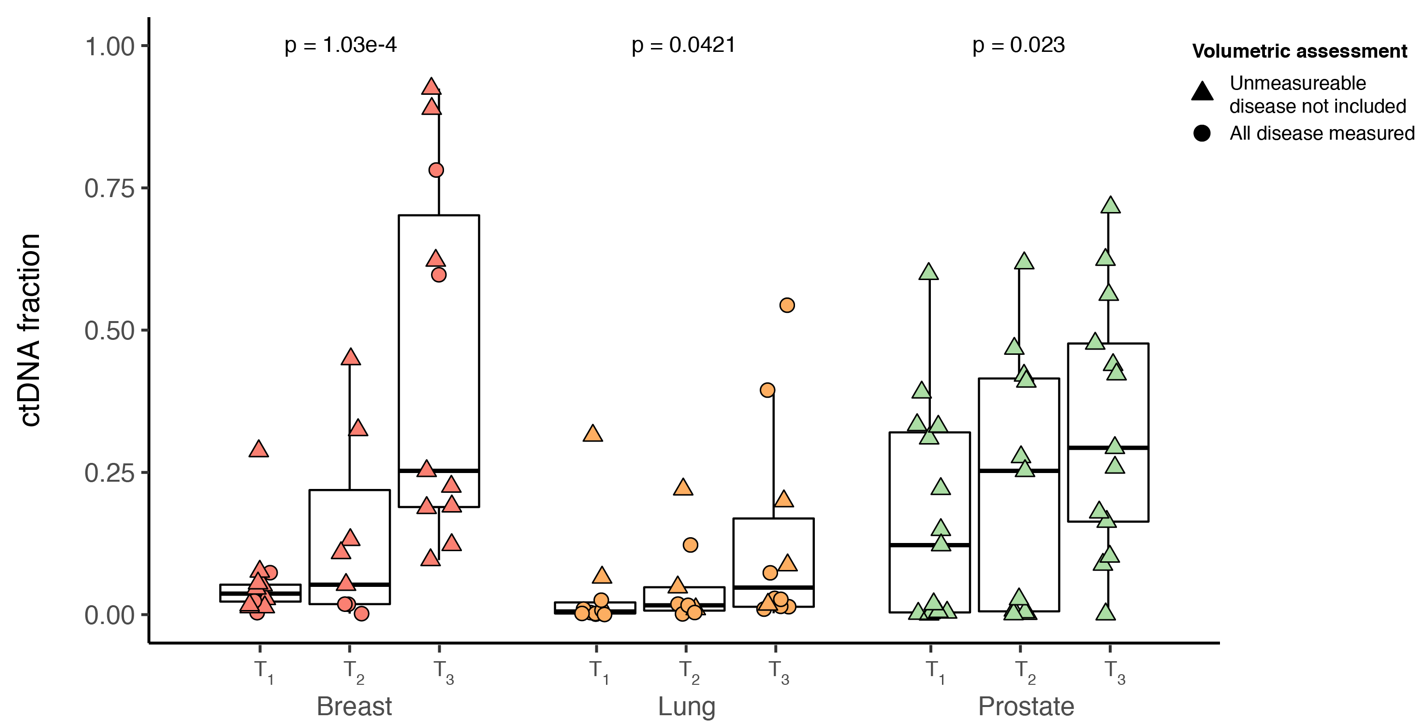
4. Figure 2 depicts the concordance of tumour and plasma and the correlation between ctDNA and body disease burden (# mets), both of which have been the subject of several prior papers in variety of malignancies and thus lacks novelty. Unless the authors are trying to make the case for a better correlation between ctDNA levels and breadth of anatomic spread (than total tumour burden), then better measures of burden than the number of metastatic sites are the more conventional measures of burden including MTV, GTV, SLD, etc.

Authors: We thank the Reviewer for suggesting these additional analyses, which have resulted in a substantial improvement of this aspect of our study. Following the Reviewer’s excellent suggestion, we performed volumetric analysis of the pre-cfDNA collection CT scans of all the breast and lung cancer cases. 77 of the 80 patients in the lung and breast cohorts had CT exams available from which volumetric tumor measurements could be obtained. 34 of the exams were CTs of the chest, abdomen, and pelvis without IV contrast, obtained as part of a PET/CT exam. 32 exams were CTs of the chest, abdomen, and pelvis with IV contrast. 5 exams were CTs of the chest only with IV contrast, 4 exams were CTs of the chest only without IV contrast, and 2 exams were CTs of the chest and abdomen with IV contrast. Exams were acquired on several different scanners at slice thicknesses ranging from 3.75 - 5 mm.

All exams were reviewed by a board-certified radiologist specializing in imaging of the chest, abdomen, and pelvis (Dr. Krishna Juluru, Director of MSK’s Advanced Imaging Laboratory). All metastatic lesions exceeding 1 cm in diameter were identified. Volumes were measured on all lesions except bone lesions. Bone lesions often have poorly defined borders and overlap the findings in active metastasis vs treated disease. Volumes were measured using the Aquarius iNtuition advanced visualization software, version 4.4.13.P3 (TeraRecon, Inc, Foster City, CA).

Of the 77 patients with available volumetric assessment, 34 breast and 29 lung cancer patients had evaluable ctDNA fraction and included in the analysis. We have now updated the correlative analysis (see **Response to Reviewers Figure 4** below) and added the results as **Figure 2g in the revised version of the manuscript**. The text has been modified accordingly. In brief and consistent with our previous observations and our working hypothesis, we found a significant association between the estimated disease volume (tertile of disease volume in ml) and ctDNA fraction in both breast and lung cancer patients (p = 1.03e-4 and 0.042, respectively).

As for prostate cancer, given their distinct pattern of metastasis, the approach employed for the volumetric assessment of disease was distinct from that employed for breast and lung cancers. In fact, the majority of the castration resistance prostate cancer patients included in this study had extensive bone disease and had undergone bone scans prior to enrolment in the study. Therefore, we obtained the automated bone scan index (aBSI) as a proxy for bone disease burden. aBSI is a fully quantitative assessment of a patient’s bony disease on a bone scan that reports the number of lesions, area, and the fraction of the total skeleton weight that is involved by tumor. We used the aBSI platform version 3.3, developed by EXINI Diagnostics AB (Lund, Sweden) on the available bone scans. The methodology of the automated platform has been described in detail in previous studies (PMID: 22306323). In brief, a neural network automatically segments the different anatomical regions of the skeleton followed by detection and classification of the abnormal hotspots. The weight fraction of the skeleton for each metastatic hotspot was calculated and the BSI was calculated as the sum of all such fractions. The aBSI method utilized in this study has undergone rigorous pre-analytical and analytical validation as an objective measure of the quantitative change in disease burden bone scans and has been shown to be a prognostic biomarker in patients with metastatic prostate cancer (PMID: 29799999). We assessed the association between the aBSI tertiles and ctDNA fraction in 39 CRPC patients with evaluable bone scans and ctDNA fractions (see **Response to Reviewers Figure 4** below). Our analysis revealed a significant association between aBSI and ctDNA faction (p = 0.023).



**Response to Reviewers Figure 4: Association of disease burden and ctDNA fraction.** ctDNA fraction estimates (y-axis) as a function of disease burden (x-axis) for breast, lung and prostate cancers. In breast and lung cancer, estimated disease volume and in prostate the automated bone scan index (aBSI) were utilized to estimate disease burden. The association between tertitles of disease burden for each cohort and ctDNA fraction was estimated using a one-sided Jonckheere-Terpstra test for increasing ctDNA fraction. Triangles indicate patients from whom some distant metastases could not be measured and were not included in the volumetric assessment.

We have now amended the manuscript to reflect the more precise assessment of tumor burden (please see page xx, paragraph xx) and **Figure 2g** of the revised manuscript**.**

5. Figure 3 depicts the relationship between inferred TMB from ctDNA and tumour biopsies. Again, this seems not to have novelty since TMB estimation from blood plasma has been published on by several groups in much larger cohorts. This section seems especially under developed in the current manuscript. The analysis of mutational signatures in the 10 hypermutated cases (seems an unusually high rate in 124 cases) also does not seem to represent an advance as no I cannot appreciate new biology, etc. The analysis of tumour mutational burden, MSI, and mutational signatures is interesting as an assessment of the ctDNA detection method but it is not clear that in the scope of this work as it does not ‘reveal the sources of plasma circulating cfDNA’ nor does it quantitate the accuracy of the method.

Authors: We agree with the Reviewer in that previous studies have reported on TMB estimations utilizing the results from cfDNA analysis. Our study, however, provides information above and beyond the mere estimation of the TMB or the characterization of the signatures, given that we could ascertain whether the mutations were biopsy-matched, detectable in the biopsy sample or CH-related. We respectfully differ with the Reviewer in that the signature analysis “does not ‘reveal the sources of plasma circulating cfDNA’ nor does it quantitate the accuracy of the method”. We would content that the signature analyses have provided strong circumstantial evidence to support the notion that a large proportion of the mutations in the hypermutant cases are tumor-derived rather than sequencing errors and/or CH, given that the exposures matched those expected in each cancer type (e.g. APOBEC signatures in metastatic breast and in lung cancers, as elegantly shown in Bertucci *et al.* (PMID: 31118521), de Bruin *et al.* (PMID: 25301630) and Alexandrov *et al.* (PMID: 23945592), smoking signature in a metastatic lung cancer, and MMR deficiency signatures in a prostate cancer with MSI). Importantly, in the hypermutated cases, a large subset of the mutations detected in cfDNA, although not biopsy-matched, were still present in the biopsy samples under the limit of detection of the clinical pipeline for mutation detection ( i.e. biopsy-subthreshold, 89 of 216 tumor-derived, 41%, **Figure 4**); the mutations present in the hypermutated cases displayed a significant correlation with gene size (p = 4.44e-16**, Supplementary Figure S9**), and were not enriched for canonical CH-related genes (**Supplementary Figure S8 of the revised version**).

Hence, the signature analyses have added another line of evidence in understanding the sources of somatic variants detected in cfDNA samples and given further biological credence to the notion that a large proportion of the mutations in hypermutated cases indeed stemmed from cancer cells. These findings also indirectly support the accuracy of the high-intensity sequence method coupled with the hierarchical Bayesian model for error correction employed in this study.

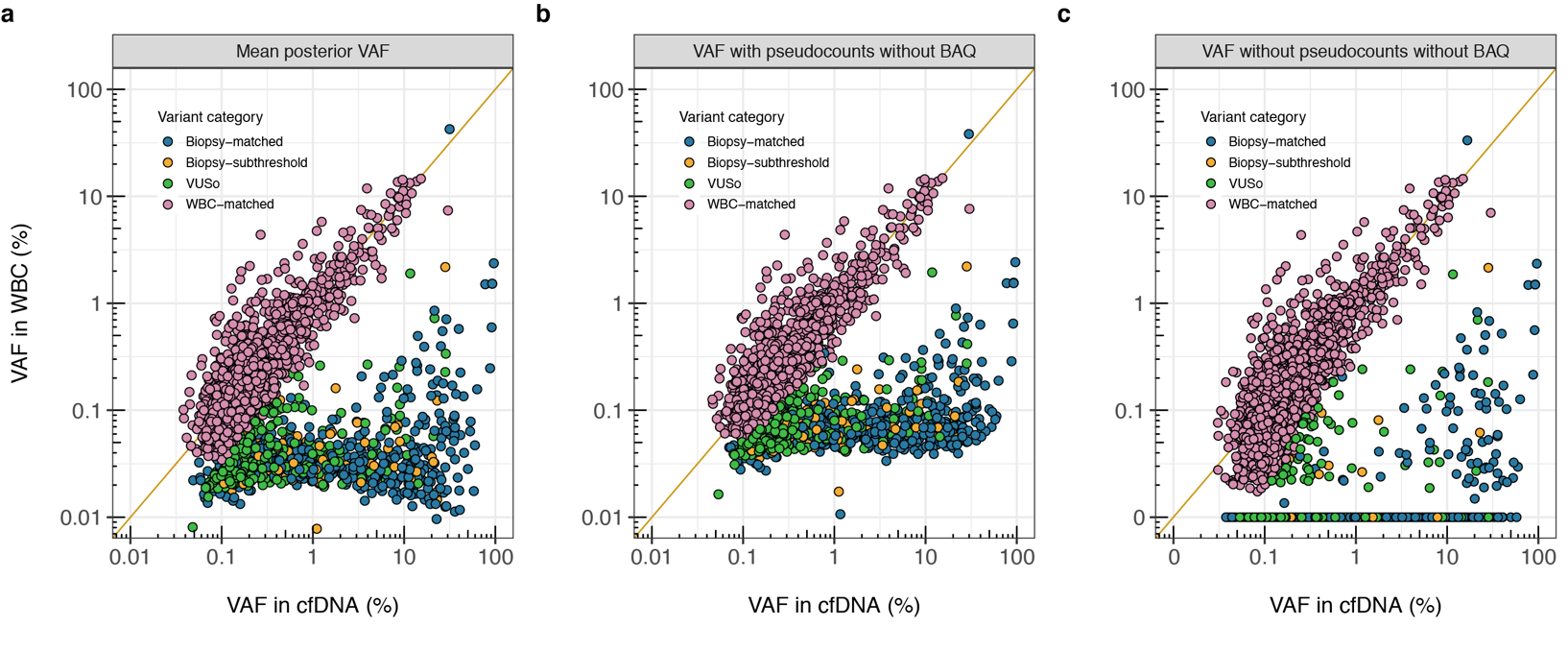
6. Figure 4 depicts CH results analogous to those by Liu et al 2018 (Ref 26), who reported very similar results in observing essentially the same recurrent mutations and made the same observation of similar AF in WBC and cfDNA. Separately, the scatterplot depiction of the VAFs in leukocytes vs cfDNA appears to be from modeled VAFs, not actual VAFs. The presentation of modeled VAFs would seem reasonable only in the context of a performance characterization of the model against empirically observed/orthogonally validated methods, and this seems lacking.

Authors: We thank the Reviewer for pointing this out. We would like, however, to highlight a few important aspects distinguishing our study from that of Liu *et al.* (PMID: 30475948):

1. The study by Liu *et al.* (PMID: 30475948) included only healthy individuals. The study showed, as expected, that the majority of the somatic mutations in plasma originated from CH. Our study, however, is the first to show the large contribution of CH mutations to the cfDNA mutational burden in cancer patients. In the 114 non-hypermutated patients, more than 50% of the mutations identified in cfDNA likely originated from CH.

1. Our study utilized the same sequencing methodology as the cfDNA to sequence the cfDNA and WBC (i.e. we utilized UMI barcoding and targeted raw sequencing depth of >60,000X for both cfDNA and WBCs). This allowed us to identify CH mutations at much lower limits of detection both in cfDNA and WBCs compared to Liu *et al.* (PMID: 30475948) where the WBCs were sequenced the to a mean target depth of 406X resulting in a VAF limit of detection threshold at least 10 times higher than our study.

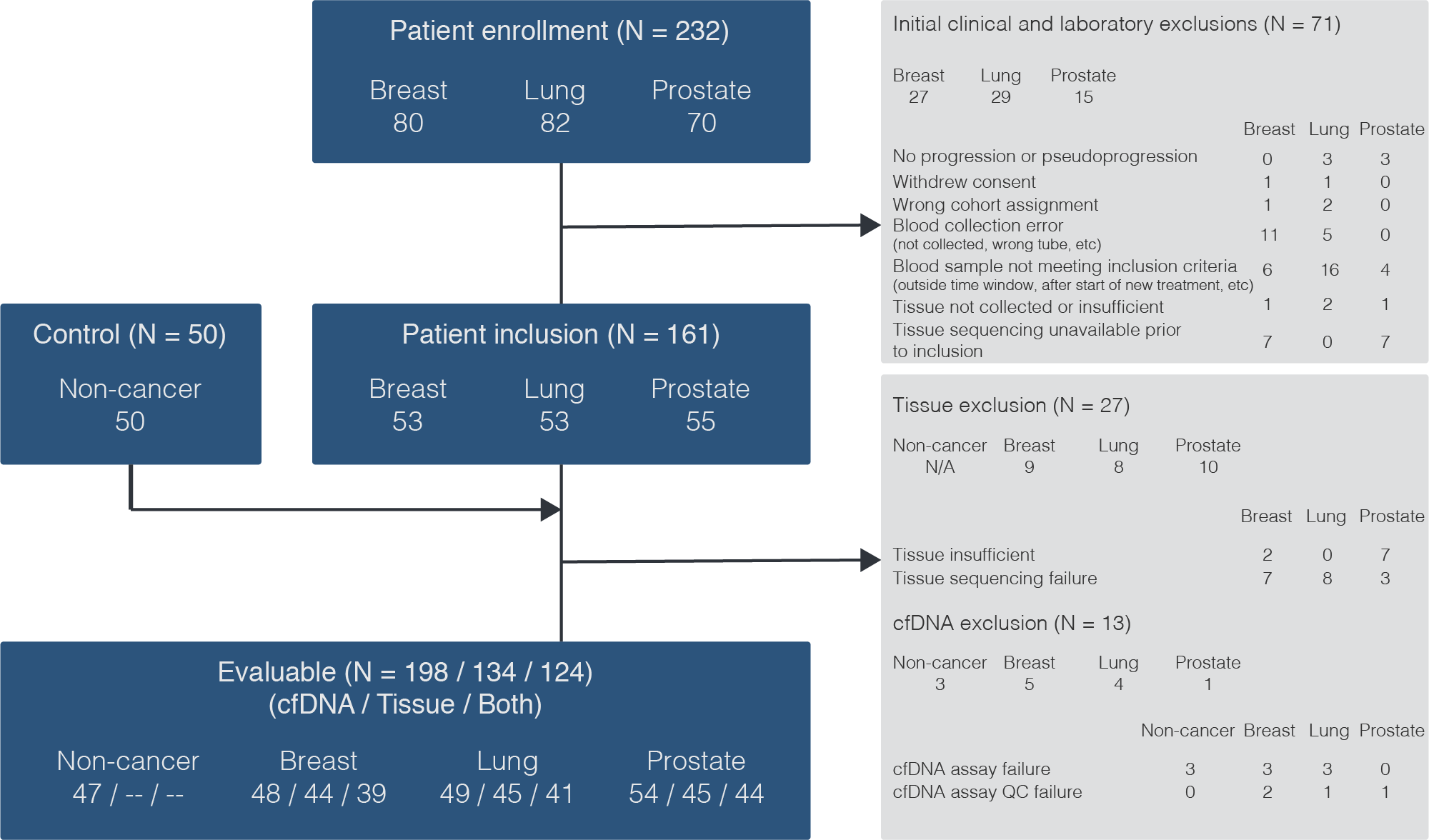
We appreciate the Reviewer’s comments regarding the modeled cfDNA and WBC VAFs reproduced in **Response to Reviewers Figure 5a** below. Modeled VAFs were obtained as the mean posterior estimate from the Bayesian hierarchical model after Markov chain Monte-Carlo resampling. We now additionally present as **Supplementary Figure SX** and **Response to Reviewers Figures 5b** and **5c** below the two alternative representations of this scatterplot where the VAFs are estimated from the raw pileup without BAQ filtering and with or without the pseudocounts such that and where and are the alternate allele and total base counts, respectively. Whilst being more factual, raw VAFs without pseudocounts cannot be displayed accurately since variants detected in cfDNA with zero alternate allele count in WBC cannot be represented in a typical scatterplot with logarithmic axes. Similarly, whilst the raw VAFs with added pseudocounts is a possible alternative, the equivalent depiction using posterior estimates of VAF does not affect the interpretation of the underlying data and has the advantage of being part of the model rather than being transformed for display purposes.



**Response to Reviewers Figure 5: Comparison of VAF of somatic mutations in cfDNA and WBC.** Comparison of VAF in cfDNA (*x*-axis) and genomic DNA isolated from WBC (*y*-axis) using three alternative definitions to estimate VAF. Shown is (a) posterior estimate of VAF obtained from the Bayesian hierarchical model after Markov chain Monte-Carlo resampling, (b) estimate of VAF based on reference and alternate allele counts in cfDNA and WBC obtained from raw pileup without base alignment quality filtering (BAQ) and added pseudocounts and (c) same as in (b) without pseudocounts. In all panels, the variant category i.e. Biopsy-matched, Biopsy-subthreshold, VUSo and WBC-matched are color coded and described in the corresponding legends. The diagonal line represents the line *y = x*. Panel (a) represents modeled VAF as displayed in Figure 4(e) of the manuscript. VAF based on pseudocounts displayed in panel (b) are obtained as where AD and DP are the variant level alternate and total collapsed depths. In (c), variants detected with zero alternate depth in WBC were leveled at 0.01% VAF and labeled as 0 due to the logarithmic scaled axes.

7. The patient enrolment flow suggests an unusually high rate of failure / rejection of subjects and/or specimens for a variety of reasons, which makes the current assay from GRAIL to seem an outlier when considering the rate of failures of similar assays from Guardant, PGDx, Foundation ACT, etc. This major discrepancy seems concerning and is poorly addressed.

Authors: We thank the Reviewer for bringing this issue to our attention. We now present as **Supplementary Figure S1** and **Response to Reviewers Figure 6** below the updated CONSORT diagram providing a detailed breakdown of the patient enrollment and the reasons for exclusion. As indicated below, the number of cfDNA assay failures appear to be relatively low and comparable to the previous reports (5.6%, 9/161). Additionally, 3 more cases (1.9%) had to be excluded due to cfDNA assay QC failure, including one case of mislabeled blood tube identified after performing the assay. The diagram has now been updated in the revised version of the manuscript.



**Response to Reviewers Figure 6: Study overview.** Patient enrollment, inclusion, and evaluable group are defined in the blue boxes. Clinical, tissue and cfDNA exclusions are shown in the gray boxes.

8. The raw mutation data provided in Supplemental tables 6-8 are largely un-interpretable due to the fact that it is not clear what mutations in the Supplemental tables were included in the analysis. For example, there are 3,354 cfDNA mutations reported in Supp Table 7, but only 740 cfDNA mutations listed in Fig. 1a, and not clear why this might be the case. The Supplemental tables are very hard to interpret as there is no legend.

Authors: We apologize for the lack of clarity in describing these results. **Supplementary Table S6** of the manuscript now provides details on the patients selected to evaluate the reproducibility of the cfDNA assay. **Supplementary Tables S7-S9** were provided following the standard Mutation Annotation Format (MAF) with aggregated information for each variant. A description of the file format together with the headers can be found on the NIH’s website (<https://docs.gdc.cancer.gov/Data/File_Formats/MAF_Format/>). Additionally, the PHENO column (i.e. column 99/CT) provides the category of the particular variant i.e. Biopsy-matched, Biopsy-subthreshold, VUSo, and WBC-matched for variants detected in cfDNA and compiled in **Supplementary Table S7**, CH-derived for variants found in genomic DNA extracted from buffy coat and shown in **Supplementary Table S8** and finally, Biopsy-matched, Biopsy-subthreshold, and Biopsy-only for variants detected in the matched tumor and provided in **Supplementary Table S9**. The captions of **Supplementary Table S7-S9** have

been updated accordingly in the revised version of the manuscript. Supplementary tables of this format are often submitted in a very similar manner in sequencing publications to allow other investigators to easily input the data into their preferred software for additional analyses. To ensure that the readers can scrutinize these supplementary tables and utilize them in subsequent analyses of their own, we have now provided captions for each of these supplementary tables to render their interpretation easier (page XX, paragraph XX).

9. The supplemental Excel tables 6-8 provide variant level estimates, but not sample level estimates of depth etc, and the associated Supplemental Table describing the cohort poorly captures these indices. Information such as plasma cfDNA concentration, DNA input into library preparation, median deduped depth for each sample run, amount of sequencing given to each sample, error statistics…etc. would seem critically important for such a paper.

Authors: We apologize for the omission of these important pieces of data. We have now provided the sample level metrics as **Supplementary Table S10**. A description of the column headers is as follows:

1) Patient\_ID: Patient identifier

2) Sample\_Type: The type of DNA i.e. cfDNA or gDNA

3) Tissue: The type of tissue i.e. Breast, Lung, Prostate, or Healthy

4) Volume\_of\_blood\_mL: Volume of blood used for DNA extraction

5) Volume\_of\_DNA\_source\_mL: Volume of plasma or buffy coat used for cfDNA extraction

6) DNA\_extraction\_yield\_ng: Yield of DNA after extraction

7) DNA\_input\_concentration\_ng\_uL: Concentration of input DNA

8) Library\_praparation\_input\_ng: Amount of DNA used for library preparation

9) Uncollapsed\_Mean\_Coverage: The mean coverage of all baits in the experiment obtained from deduplicated and uncollapsed reads

10) Collapsed\_Mean\_Coverage: The mean coverage of all baits in the experiment from deduplicated and collapsed reads

11) Collapsed\_Fragment\_Mean\_Coverage: Mean unique coverage across target regions obtained from deduplicated and collapsed reads

12) Indel\_and\_Substitution\_Error\_Rate: Percentage of collapsed bases with SNV or indels

13) Substitution\_Error\_Rate: Percentage of collapsed bases with SNVs

14) Assay\_Version: Version of the assay used i.e. V1 or V2

In addition, we have also included a data dictionary in **Supplementary Table S10**, explaining each of the fields included.

10. This reviewer considered the median depth levels across variants found in cases versus controls and between leukocytes and cfDNA as tabulated in Excel Supplemental tables 6-8. Unfortunately, there are major differences in the depth of sequencing of cases and controls when considering cfDNA specimens, and also between cfDNA and leukocytes (again with significant differences between cases:controls for this ratio). It is unclear whether this reflects pre-analytic variation or analytic variation, but this could lead to major technical artifacts that could provide for trivial explanations to some of the other differences between cases:controls (eg, plasma cfDNA concentration).

Authors: Sample level metrics are compiled in **Supplementary Table S10**. **Response to Reviewers Figure 7a** demonstrates that the distribution of mean target coverage of uncollapsed reads of the cfDNA assay were largely overlapping comparing the three cancer cohorts (two-sided Mann-Whitney *U-*test, all p>0.05). The healthy control individuals had a higher cfDNA mean target uncollapsed sequencing depth. This was deliberate to accommodate the lower sequencing depth after collapsing due to lower input DNA for library preparation and reduced library complexity. The distribution of mean target uncollapsed depth of the WBC assay showed no statistically significant differences between cancer and control or between the three cancer cohorts (two-sided Mann-Whitney *U-*test, all p>0.05). Similarly, except for healthy controls who had a higher target cfDNA mean depth for reasons described above, the comparisons involving the mean target uncollapsed depth of cfDNA and WBC showed no statistically significant differences (two-sided Mann-Whitney *U-*test, all p>0.05).

The distribution of cfDNA mean target coverage from collapsed reads (**Response to Reviewers Figure 7b** below) was also comparable within the different cohorts (two-sided Mann-Whitney *U-*test, all p>0.05). However, there was a statistically significant difference of mean target collapsed sequencing depth between cfDNA and WBC DNA for each cohort with cfDNA samples having on average higher collapsed mean target coverage than WBC (two-sided paired Mann-Whitney *U*-test, p=1.03e-13). This is expected given that a maximum input of 75 ng cfDNA with no lower limit was deemed acceptable for library preparation. On the other hand, a fixed 50 ng of sheared size selected genomic DNA (gDNA) extracted from buffy coat was used for the WBC assay. **Response to Reviewers Figure 7b** illustrates that this effect is present across all cancer types and healthy control individuals. As shown in **Response to Reviewers Figure 7c** and **7d**, the input DNA for library preparation is the main pre-analytical condition which explains the variability of cfDNA collapsed mean target coverage (*R2* =0.62, p = 4.7e-37).

Although we would be delighted to be in a position of extensively validating the mutations detected using orthogonal methods as the Reviewer suggested, the reality is that for the vast majority of the samples included in this study, no biological material, other than tumor samples, is still available. In addition, the development of assays to validate orthogonally our findings would require several months, even in partnership with commercial vendors, given that the vast majority of the variants identified are not hotspots. In the original version of manuscript, we included some experiments to demonstrate the accuracy of the ‘high intensity’ assay; we have now expanded this validation utilizing the samples available and the reagents we had access to from commercial vendors that could be employed as a means to define a `ground truth`. The authors would like, therefore, to bring the following results to the attention of the Reviewer:

* + - 1. The raw error profiles of cfDNA and WBC assays are displayed in **Response to Reviewers Figure 7e** and **7f** and provided in **Supplementary Table S10**. The substitution error rate represents the percentage of collapsed bases with single nucleotide variants and the indel and substitution error rate represents the percentage of collapsed bases with single nucleotide variants or indels per sample. Although WBC samples have, on average, lower error rates compared to cfDNA, the distribution is comparable for cfDNA samples across the three cancer cohorts (two-sided Mann-Whitney *U*-test; all p>0.05) and higher than in healthy control individuals. The error profiles of white blood cells show no statistically significant differences within the three cancer cohorts (two-sided Mann-Whitney *U*-test; all p>0.05).
      2. The manuscript included technical replicates carried out to test the reproducibility of the targeted sequencing assay for cfDNA and WBC using two different versions (V1 and V2) of the protocol. The main differences were the UMI sequences in the library adapters and the reaction volumes in the hybridization enrichment process. Six patient samples, detailed in **Supplementary Table S6** of the manuscript, were selected for reprocessing with both assay protocols. The results were shown in **Figure 1** and **Supplementary Figure S3** of the original version of the manuscript. The measured VAFs between the two technical replicates for samples from the five non-hypermutated patients showed a strong agreement (*R2* = 0.9997, **Figure 1C**) as well as for one sample from a hypermutated case (*R2* = 0.9972, **Supplementary Figure S3**). In response to the Reviewer’s comment, we have now performed additional replicate cfDNA sequencing of three of the above patients using version V2 of the sequencing assay. The pairwise comparison of VAFs between versions V1 vs V2 and V2 vs V2 for all the samples that have been retested are shown in the **Response to the Reviewers Figure** 8 and the revised **Supplementary Figure S3** of the manuscript. The number of variants from version V1 broken down per variant source category which have been validated across the replicates are displayed in **Response to Reviewers Tables 7 and 8**. Overall, across the five non-hypermutated patients, 170 of variants detected in version V1 have been tested and 152 (89.4%) of these were also detected in version V2. In the hypermutated breast cancer patient, 630 of 659 (95.6%) of variants detected in version V1 were also detected in version V2. Comparing version V1 and V2 for the three patients who have been retested for the purpose of this response, 45 of 51 (88.2%) variants were found present in two non-hypermutated patients whilst the corresponding number in the hypermutated breast cancer patient was 640 of 659 (97.1%).

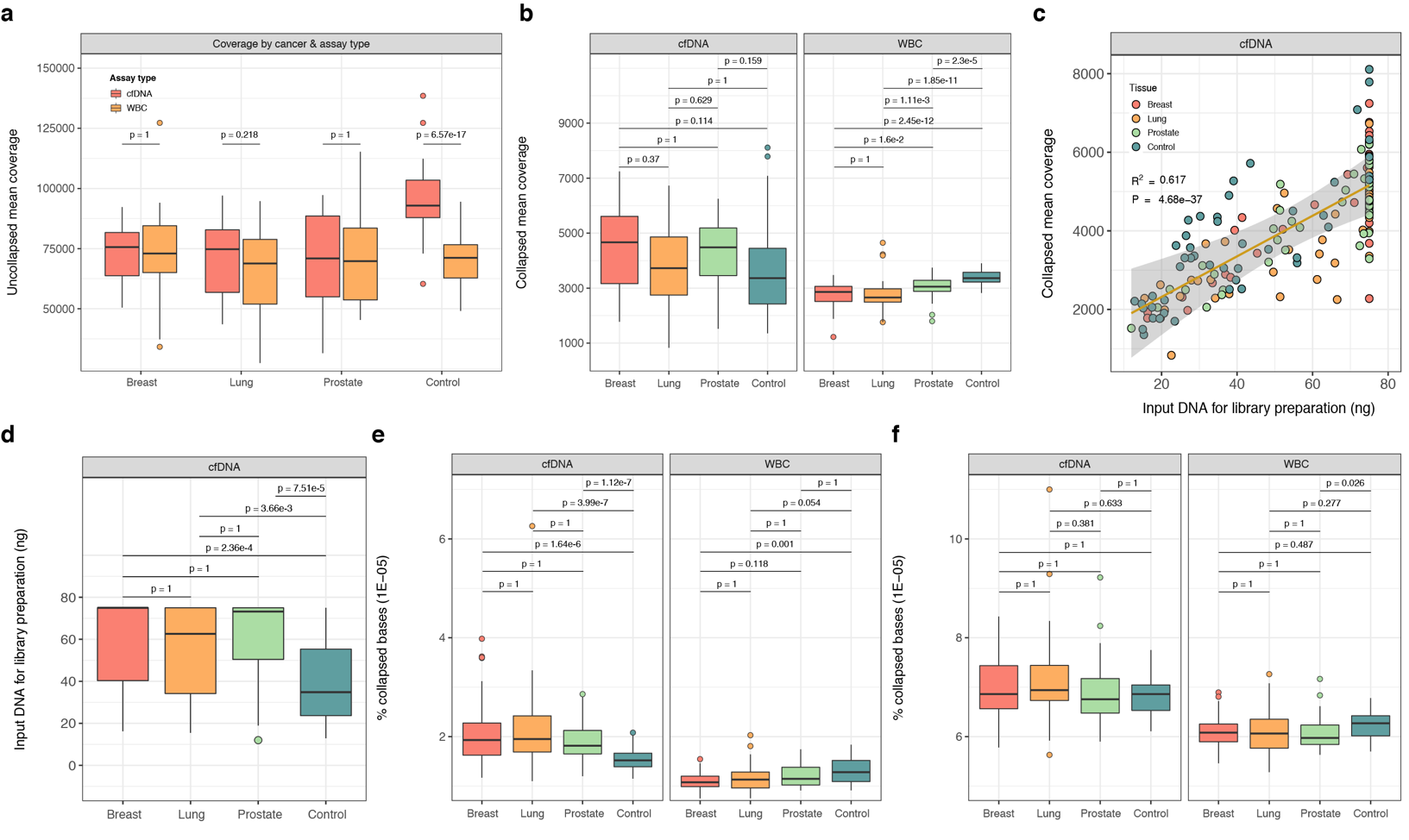
1. Although the technical repeats of the samples using V1 and V2 of the high-intensity assay demonstrate the robustness of the method, one could claim that the measurements were robustly incorrect. Hence, to test the accuracy of the high-intensity sequencing approach to determine the presence and measure the VAF of the mutations, we performed an orthogonal validation of our results utilizing droplet digital PCR (ddPCR), which is considered in many contexts to constitute a ‘gold standard’ for the assessment of variants at low VAFs. In the original version of the manuscript, we had performed ddPCR experiments targeting five canonical hotspot tumor-matched somatic mutations (*PIK3CA* E542K, *PIK3CA* H1047R, *KRAS* G12C, *KRAS* G12A, *EGFR* L861Q) in five patients. **Figure 1b** of the manuscript and **Response to Reviewers Figure 9a** demonstrate a 100% positive percent agreement (PPA) and 100% negative percent agreement (NPA) considering ddPCR as the benchmark whilst the VAF measured in cfDNA using either version of the assay protocol closely mirrors that obtained with ddPCR. In response to the Reviewer’s comment, we have now performed additional ddPCR assays targeting VUSo (i.e. somatic cfDNA variants that were not matched to the paired tumors or WBCs) to evaluate the specificity of the cfDNA sequencing assay for low VAF non tumor-matched somatic variants. For this analysis, we first identified the subset of VUSo for which a validated ddPCR assay was available. We then identified all the patients with residual extracted cfDNA and/or leftover pre-enrichment sequencing libraries harboring any of these mutations. Our final analysis included four ddPCR assays (*PIK3CA* E545K, *NRAS* Q61K, *CTNNB1* T41I, *CTNNB1* S33C) targeting VUSo in seven patients. For one patient, only cfDNA extracted from plasma was available. For two patients, both cfDNA and pre-enrichment sequencing libraries were available and tested, whereas for the remaining four patients, only pre-enrichment libraries were available. As a negative control, the pre-enrichment sequencing libraries from 12 patients where the target VUSo were not detected in cfDNA were used. All experiments were performed in triplicate. The results are shown in **Response to Reviewers Figure 9b** below and **Figure 4** of the revised manuscript whereby the cfDNA sequencing assay had a 100% PPA and 100% NPA considering ddPCR as the benchmark.
2. In a previous publication of this group (Li *et al.* PMID: 30891595)…

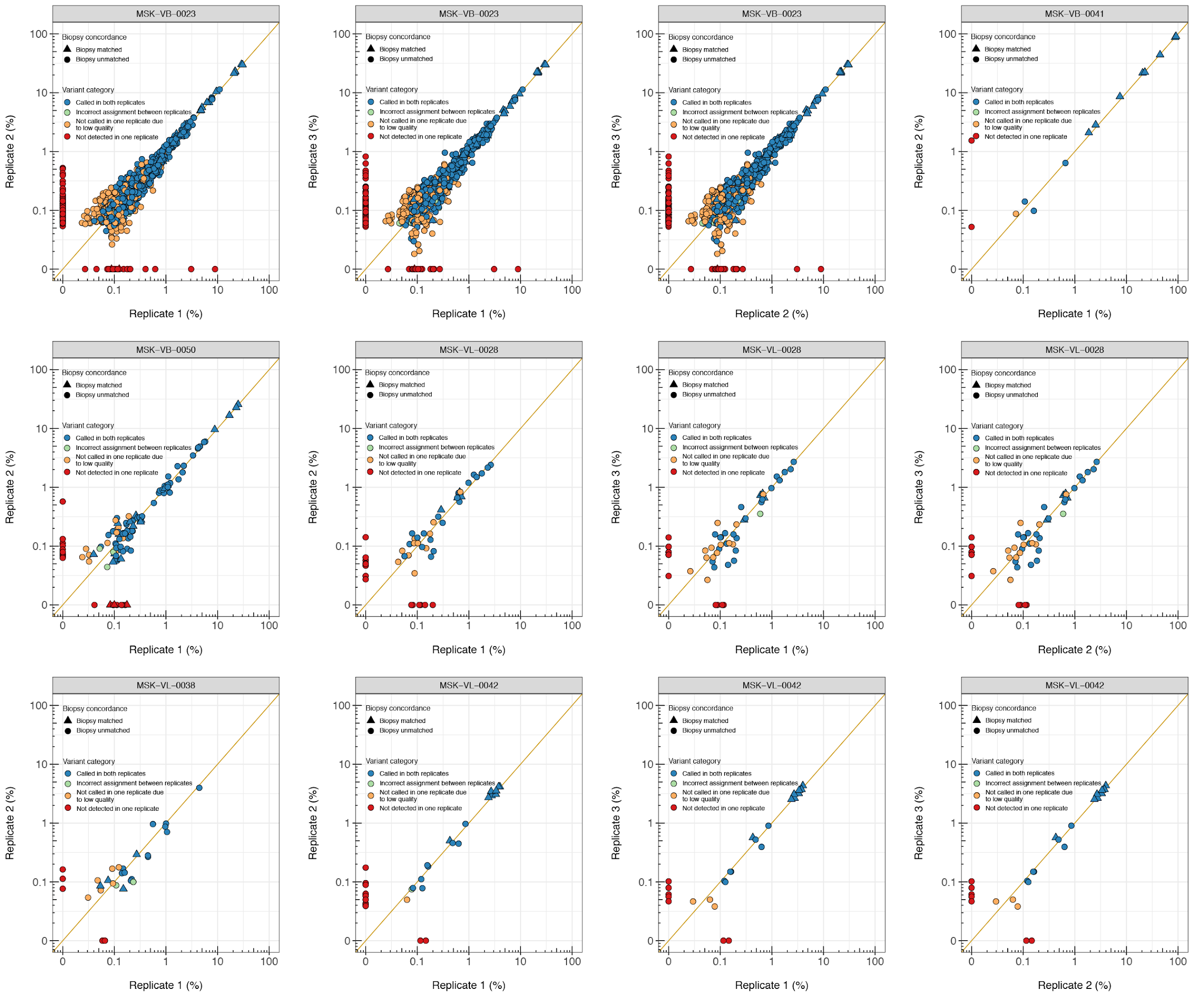
**Response to Reviewers Table 7**: Number of somatic mutations per patient detected in version V1 and confirmed present in six initial technical replicates using version V2 of assay protocol

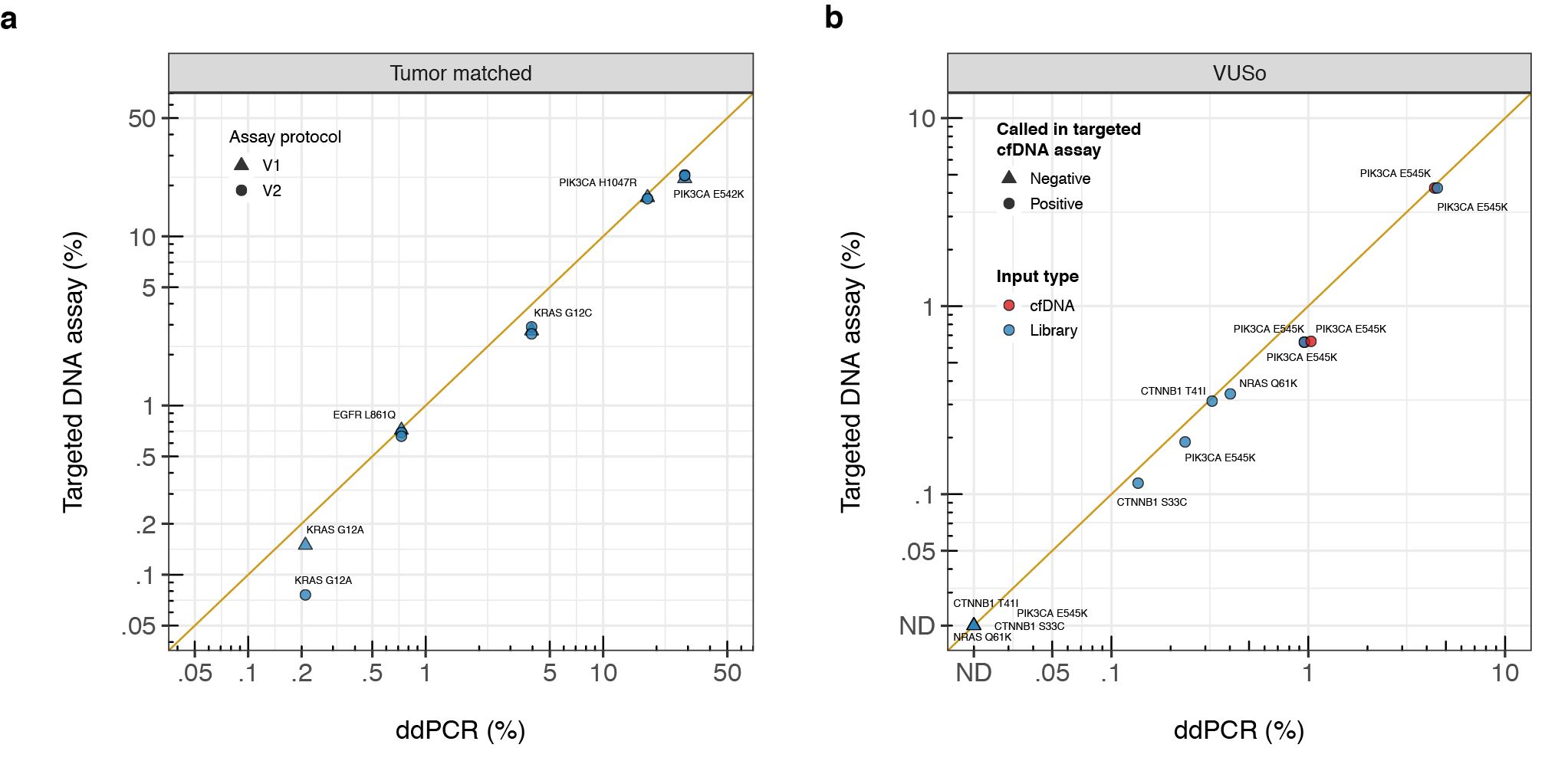
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Patient ID | No. of Biopsy -matched variants | No. of Biopsy -subthreshold variants | No. of VUSo | No. of WBC -matched variants | Total no. of variants |
| MSK-VB-0050 | 13 / 16 (81.3%) | 5 / 5 (100%) | 53 / 59 (89.8%) | 5 / 6 (83.3%) | 76 / 86 (88.4%) |
| MSK-VB-0041 | 8 / 8 (100%) | 0 (N/A) | 2 / 2 (100%) | 1 / 1 (100%) | 11 / 11 (100%) |
| MSK-VL-0028 | 4 / 4 (100%) | 0 (N/A) | 1 / 3 (33.3%) | 22 / 26 (84.6%) | 27 / 33 (81.8%) |
| MSK-VL-0042 | 7 / 7 (100%) | 1 / 1 (100%) | 4 / 4 (100%) | 6 / 6 (100%) | 18 / 18 (100%) |
| MSK-VB-0023 | 16 / 16 (100%) | 51 / 53 (96.2%) | 553 / 576 (96.0%) | 10 / 14 (71.4%) | 630 / 659 (95.6%) |
| MSK-VL-0038 | 4 / 4 (100%) | 0 (N/A) | 2 / 3 (66.7%) | 14 / 15 (93.3%) | 20 / 22 (90.9%) |
| Total | 52 / 55 (94.5%) | 57 / 59 (96.6%) | 615 / 647 (95.1%) | 58 / 68 (85.3%) | 782 / 829 (94.4%) |

**Response to Reviewers Table 8**: Number of somatic mutations per patient detected in version V1 and confirmed present in three additional technical replicates using version V2 of assay protocol

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Patient ID | No. of Biopsy -matched variants | No. of biopsy -subthreshold variants | No. of VUSo | No. of WBC -matched variants | Total no. of variants |
| MSK-VL-0028 | 4 / 4 (100%) | 0 (N/A) | 1 / 3 (33.3%) | 24 / 26 (92.3%) | 29 / 33 (87.9%) |
| MSK-VL-0042 | 7 / 7 (100%) | 1 / 1 (100%) | 2 / 4 (50.0%) | 6 / 6 (100%) | 16 / 18 (88.9%) |
| MSK-VB-0023 | 16 / 16 (100%) | 52 / 53 (98.1%) | 561 / 576 (97.4%) | 11 / 14 (78.6%) | 640 / 659 (97.1%) |
| Total | 27 / 27 (100%) | 53 / 54 (98.1%) | 564 / 583 (96.7%) | 41 / 46 (89.1%) | 685 / 710 (96.5%) |

**Response to Reviewers Figure 7: Comparison of mean target sequencing depth and error rate distribution across the cancer cohorts and healthy controls.** Shown are the distributions of (a) deduplicated and uncollapsed mean target sequence depth and (b) deduplicated and collapsed mean target sequence depth. Panel (c) shows the association between the amount of cfDNA used for library preparation and the mean target deduplicated and collapsed sequencing depth whilst (d) shows the distribution of mean target deduplicated and collapsed sequencing depth across the different cohorts. Panels (e) and (f) show the distribution of the substitution error rate and substitution and indel error rate, respectively across the different cohorts. In (a), the p-values were obtained using paired two-sided Mann-Whitney *U*-tests. In (b) and (d)-(f), the p-values were obtained from pairwise comparisons using a two-sided Mann-Whitney *U*-test and adjusted for multiple testing using the Bonferroni method. In (c), the diagonal line represents a linear regression with 99% confidence intervals. The p-value was obtained using an *F*-test. In (e), the substitution error rate represents the percentage of collapsed bases with non-reference base. Similarly, in (f) the combined error rate represents the percentage of collapsed bases with non-reference base or indels.

**Response to Reviewers** **Figure 8: Reproducibility of cfDNA targeted sequencing assay**. Six patient samples were selected for processing using two versions of the assay protocol (V1 and V2). These are labelled Replicate 1 and Replicate 2. A subset of three samples were further retested using version V2 and these are labelled Replicate 3. The panels show the pairwise comparisons of measured VAF between all available replicates for each patient. In all panels, the variants are shape coded based on their origin i.e. whether they were also detected in the matched tumor biopsy and color coded according to their category i.e. whether they were called in both replicates and assigned to similar source categories i.e. VUSo, WBC-matched or noise. In all panels, the samples are labelled on top.

**Response to Reviewers Figure 9: Orthogonal validation of tumor-matched and VUSo detected in cfDNA using ddPCR.** Comparison of VAF measured using ddPCR (*x*-axis) and the cfDNA targeted assay (*y*-axis) for (a) tumor-matched canonical hotspot mutations and (b) low VAF (<10%) non tumor-matched somatic mutations detected in cfDNA i.e. VUSo. In (a), plasma cfDNA samples from five patients with hotspot tumor-matched mutations were subjected to ddPCR. An aliquot of the same cfDNA isolate was used for the targeted DNA assay using two versions of the assay protocol (V1 and V2). In (b), plasma cfDNA samples and pre-enrichment libraries of seven cancer patients with hotspot mutations not detected in the matched tumor sequencing were subjected to four ddPCR assays. For one patient, only cfDNA isolated from plasma was available. For two patients, both cfDNA and pre-enrichment sequencing libraries were available whilst for the remaining four patients, only libraries were assayed. Sequencing libraries from 12 patients where the ddPCR target variants were not detected by sequencing were used as negative controls. Four negative libraries were used for each ddPCR assay and all experiments were performed in triplicate.

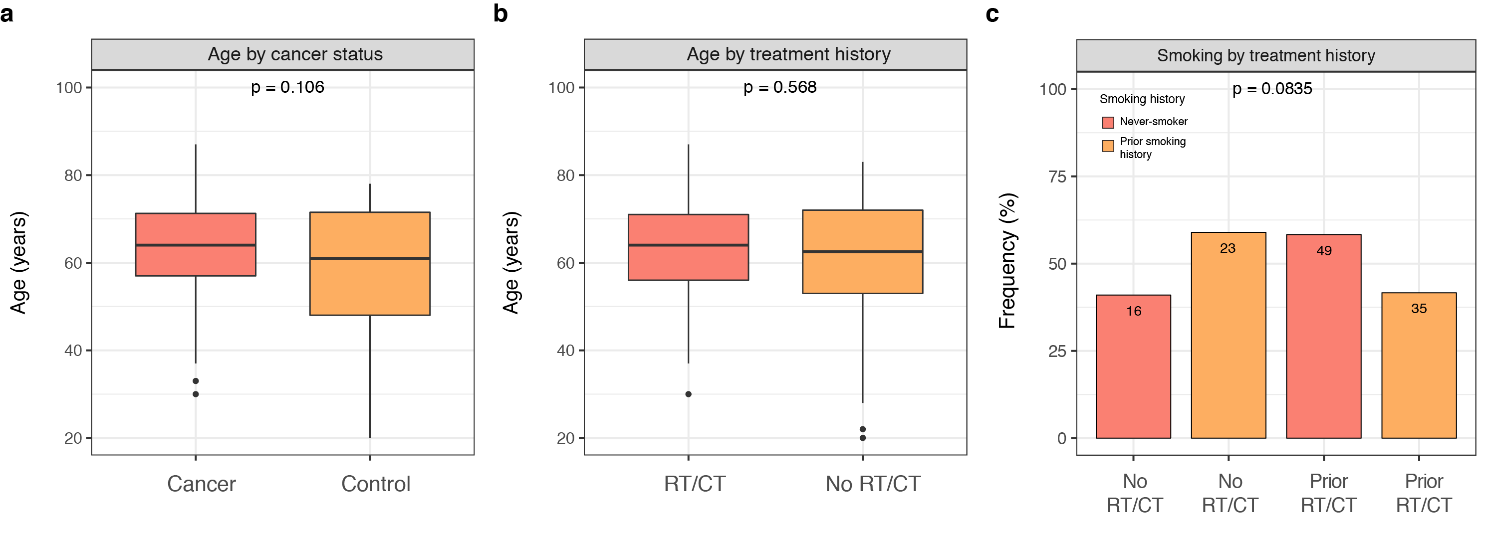
11. Figure 5 depicts the gene-wise distribution of CH variants in cases and controls, and when considering cases with or without prior history of chemotherapy and radiation. While the latter observation about PPP1MD has been made as a unique CH variant by other groups, in the current analysis presented by the authors, there are too many potential confounders and potential other variables (including mismatching for depths in cfDNA and in ratio of cfDNA:leukocyte profiling described above) that could affect the presented results. Other potential confounders include age, smoking, or other unmeasured factors that limit the analysis.

Authors: We thank the Reviewer for this comment and apologize for the lack of clarity. First, we would like to highlight that **Figure 5a** of the original manuscript does not depict results from cfDNA but instead, was generated from CH-related mutations detected in WBC. The cfDNA results are only used to estimate the overlap between CH-derived mutations in WBC and WBC-matched variants in cfDNA. The comparisons between cancer and control or RT/CT and no prior history of RT/CT presented in Figures 5b and 5c of the original manuscript were drawn utilizing solely the CH-derived mutations detected in WBC.

We acknowledge the differences in the mean collapsed sequencing depths of cfDNA between cancer patients and healthy controls. As shown in response to the previous comment (see **Response to Reviewers Figure 6**), however, the mean uncollapsed sequencing depth of WBC as well as the error rate distributions computed from collapsed reads are similar across the different cancer cohorts and comparable to that of healthy controls.

The Reviewer is absolutely correct in that other potential confounders such as age and smoking history could have affected the results of this analysis. Hence, following the Reviewer’s recommendation, prior smoking history was retrieved from the records of patients included in this study. The healthy controls accrued through the San Diego Blood Bank did not have a recorded smoking history and this information, unfortunately, could not be obtained retrospectively. In the **Response to Reviewers Figures 10a** and **10b**, comparisons of the age distribution in cancer patients and healthy controls and between the different treatment arms (i.e. RT/CT versus no RT/CT) are presented. There was no statistically significant difference between the different groups under consideration. Similarly, excluding healthy control individuals, the frequency of prior smoking history by treatment arms was not statistically different. Nevertheless, given the strong association of CH with age and previously reported association with tobacco use (PMID: 28803919), we further adjusted the analyses presented in **Figure 5b** of the original manuscript using a permutation based Likelihood ratio test to compute the p-value of a logistic regression model with age or history of smoking as covariates. For the age adjusted models, the results remained unchanged with cancer patients overall having a higher rate of CH mutations in *PPM1D* than controls (age adjusted p = 1.15e-2) and prior RT/CT being associated with increased rate of CH mutations in *TET2* and *PPM1D* (age adjusted p = 2.9e-3 and p = 6.61e-2, respectively). Due to missing smoking history for the healthy controls, the age and smoking history could only be used as covariates comparing prior RT/CT after discarding observations with missing information. The corresponding p-values were p = 8e-4 and p = 0.669 for *TET2* and *PPM1D*, respectively. We have updated the manuscript accordingly to reflect this updated analysis (page XX, paragraph XX).

Additionally, we further adjusted the analysis presented in **Figure 4c** of the original manuscript showing the association of age and number of somatic mutations in cfDNA categorized into WBC-matched, VUSo, Biopsy-matched and Biopsy-subthreshold for smoking history and found all the results to remain unchanged. Please see the updated **Figure 4c** of the manuscript presented below as **Response to Reviewer Figure 11**. We have updated the manuscript accordingly to reflect this updated analysis (page XX, paragraph XX).



**Response to Reviewers Figure 10: Age and smoking as confounding factors.** Shown is the distribution of age by (a) cancer status and (b) prior history of RT/CT versus no prior history of RT/CT. Panel (c) shows the frequency of patients who were never smokers and who had a prior smoking history by treatment history. In (a) and (b), the p-values were obtained using a two-sided Mann-Whitney *U*-test. In (c), the p-value was computed using Fisher exact test.

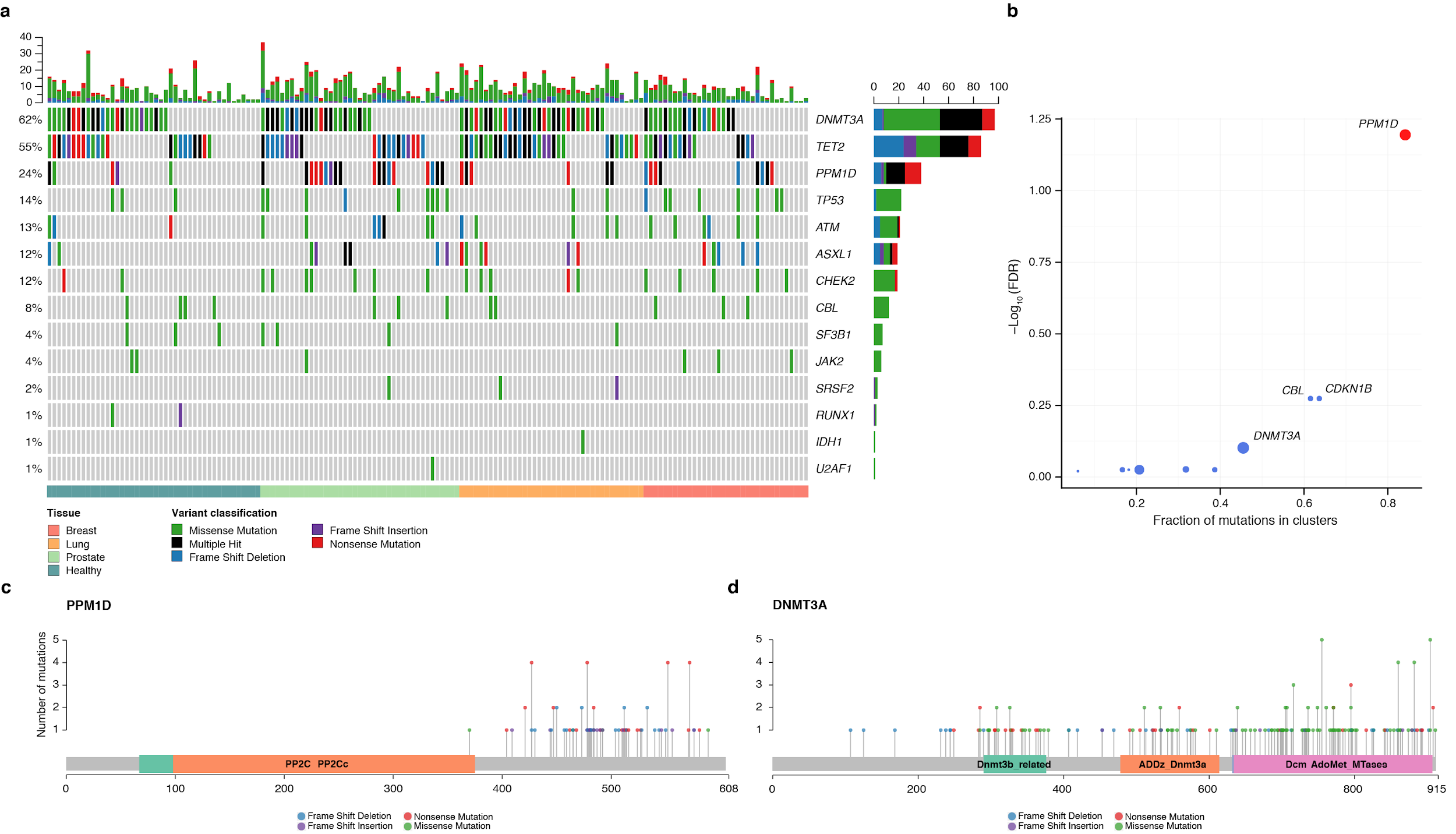
**Response to Reviewers Figure 11: Association of age and mutational burden in cfDNA.** Shown are the association of age and number of somatic cfDNA variants categorized as (a) WBC-matched, (b) VUSo, (c) tumor biopsy-matched and (d) biopsy-subthreshold. In all panels, the p-values were obtained using Wald’s test based on the coefficients of the age variable in a Poisson zero-inflated regression model with age and cancer status as explanatory variables. p\* indicates the p-values obtained using a similar model with prior smoking history as covariate. In (a) and (b), the smoking history of the healthy control individuals were treated as not available in the regression.

12. Fig 5b depicts the comparison of RT/Chemo vs treatment-naïve patients seems too small to make any firm conclusions about. The use of a ?Fisher exact or ?Chi-square test in this setting seems problematic, and this analysis would seem best done using a permutation framework for scrambling the case vs control and treatment naïve vs chemo-RT labels and seeing whether the surprisingly small p-values are valid. This is especially because the independence assumption seems violated as depicted in Fig 5a, where the observation of a given CH variant is frequently associated with CH variants in other CH genes. Also, even if assuming this observation to be real, it is unclear what the significance of the finding is. If the difference can be proven to be real, what causes the varied distribution of CH genes affected and the mutation spectrum observed in them? Are any of these of any clinical significance? Finally, the last panel in Fig 5 seems not only seems highly redundant with the preceding panel, but it seems unintelligible as currently depicted, since the legend for the circle sizes does not match the figure.

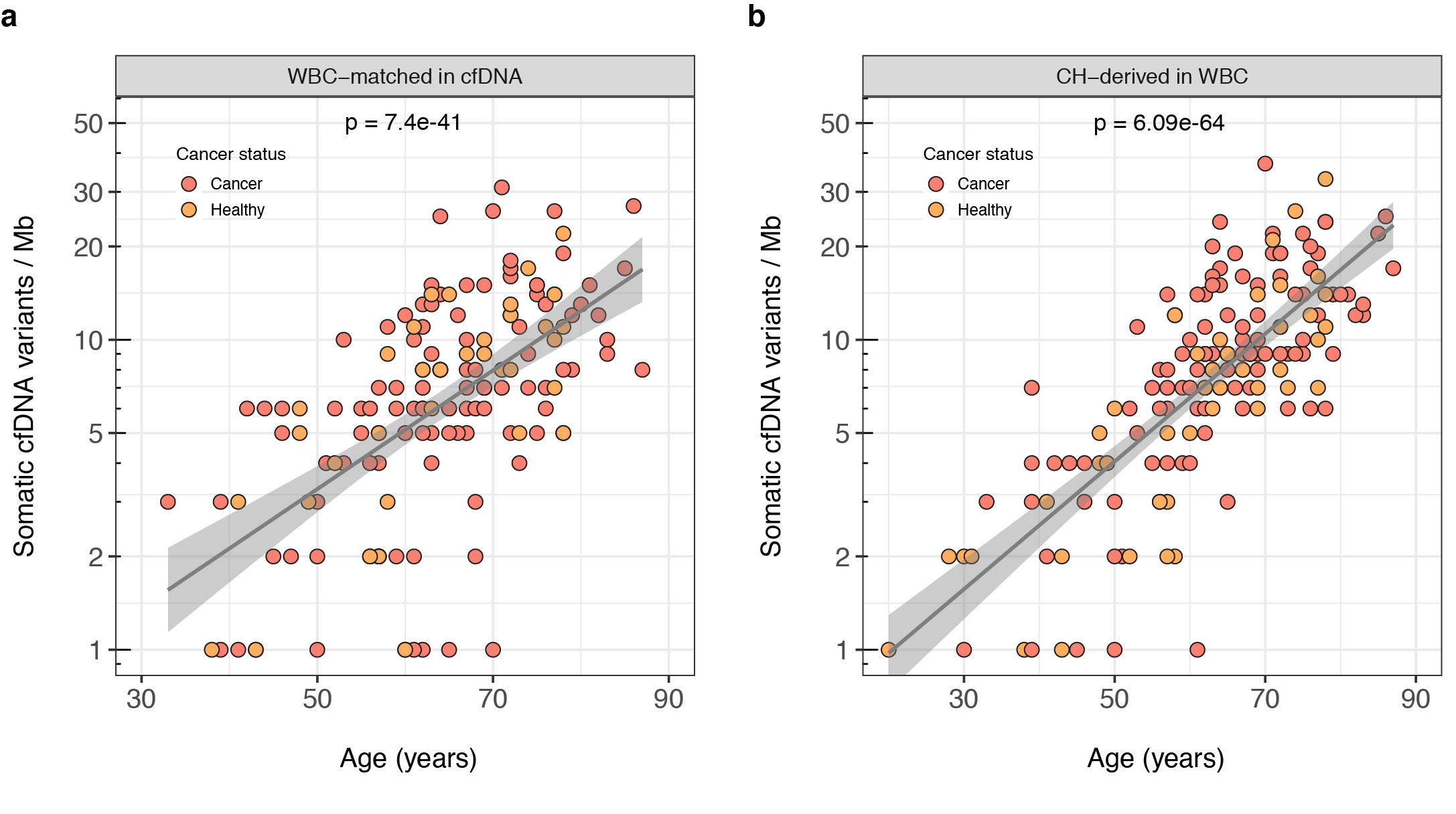
Authors: We apologize for the lack of clarity in the original version of the manuscript and thank the Reviewer for this opportunity to further elaborate. The assumption of independence in the contingency table requires only that one subject be assigned uniquely to one category when drawn from the relevant population and that this is not influenced by other subjects. Each subject in this case is a patient or healthy control, and each patient or healthy control is assigned to a unique category based on the mutation with highest VAF in CH genes within that patient. No patient or healthy control appears twice in the table based on having more than one mutation, so the cell counts in the contingency table satisfy independence. Fisher's exact test is a standard choice in this situation for comparing differences in marginal frequencies. Nonetheless, following the Reviewer’s suggestion, we redid these analyses using a permutation based Likelihood ratio test to determine the statistical significance of a logistic regression. The updated results are presented in our response to point #11 of this Reviewer’s comments.

In **Figure 5c** of the original manuscript, the frequency of mutations in the 15 canonical CH-related genes is expressed as a function of the number of cancer patients and healthy controls in the given arm. **Response to Reviewers Figure 12** demonstrates that many of the genes e.g. *DNMT3A*, *TET2*, *PPM1D*, and *ASXL1* had multiple variants per patient. The size of the circles can therefore exceed 100%. Separately, the color scale provides an indication of the ratio of truncating to missense mutations in those genes displayed. This information is not conveyed by other panels in **Figure 5** of the manuscript. We agree that this was insufficiently described and have revised the caption of **Figure 5** of the manuscript to reflect the above discussion. In the revised version of the manuscript, the caption for the figure now reads: “XXXXX”. Additionally, **Response to Reviewers Figure 12** has been added to the manuscript as **Supplementary Figure SX**.

**Response to Reviewers Figures 12b** and **12c** of this response further illustrate that most of the mutations in *PPM1D* were clustered towards the C-terminus of the protein. For comparison, **Response to Reviewers Figure 12d** shows the distribution of mutations in *DNMT3A*. Truncating mutations clustered in the terminal exon 6 of *PPM1D* are known to be associated with clonal hematopoiesis and myeloid neoplasms where they are enriched in patients who have previously been exposed to chemotherapy (PMID: 28803919). Truncated forms of PPM1D usually display increased activity due to loss of the C-terminal degradation domain and contribute to chemotherapy resistance through selective expansion of mutants in the presence of several cytotoxic DNA damaging agents (PMID: 30388424, PMID: 29954749). Although the current cohort comprises advanced stage cancer patients with follow-up shorter than the latency period of therapy related myeloid neoplasms, technological advances such as the one in the present study are beginning to shape our understanding of the etiology and pathogenesis of secondary hematological malignancies following cytotoxic treatment (PMID: 28835720).

**Response to Reviewers Figure 12: CH-derived somatic mutations detected through direct analysis of WBC.** The majority of variants detected in WBC are related to clonal hematopoiesis (CH). Shown is (a) CH-related somatic mutations in the top 14 mutated genes across the 124 cancer patients and 47 healthy controls together with the marginal frequencies by patient (top) and by gene (right), (b) clustering of CH-derived mutations in WBC, (c) distribution of mutations in *PPM1D* according to genomic coordinates and (d) same as in (c) for *DNMT3A*. In (a), (c) and (d), the variants are color coded according to their variant classification and described in the corresponding legend. In (a), *DNMT3A*, *TET2* and *PPM1D* are the top mutated genes in WBC and harbor multiple hits i.e. two or more mutations per patient. In (b), the fraction of mutations per gene occurring in clusters is shown on the *x-*axis against the log-transformed FDR-corrected p-value on the *y-*axis. The clusters and associated p-values were computed using OncodriveCLUST (PMID: 23884480). In (c), mutations detected in *PPM1D* are clustered in the C-terminus of the protein. For comparison, there are more mutations detected in *DNMT3A*. As shown in panels (b) and (d), these are distributed throughout the length of the protein.

The Reviewer highlights a rather important point, that in the samples analyzed, if a CH variant was detected in a patient, there was a high likelihood of other CH mutations being detected in the same patient. Although we do not have a definitive biological explanation for this finding, we have performed additional analyses that demonstrate that the number of CH mutations in cfDNA and in WBCs from in metastatic cancer patients and healthy who have at least one CH mutation correlates strongly with age (**Response to Reviewers Figure 13**), providing strong circumstantial evidence that this observation may a biological phenomenon (e.g. CH).



**Response to Reviewers Figure 13: Association of age and CH burden.** The number of CH-derived mutations per patient or healthy control (y-axis) as measured in (a) cfDNA and (b) WBC is plotted against age (x-axis). In both panels, the p-values were obtained from a zero-inflated Poisson regression model.

13. One of the major contributions of this study could be making the mutation and sequencing data available as a resource, and it is not immediately clear if this is the case.

Authors: We thank the Reviewer for this insightful suggestion and entirely agree that releasing the sequencing data to the public domain will make this manuscript a resource paper for the research community to replicate our methodology and also to develop novel analytical approaches utilizing our high-intensity sequencing approach. We now deposited all the sequencing data reported in the manuscript including the raw BAM, collapsed BAM and VCF files of cfDNA and WBC for the full cohort (>45 TB of data) at the European Genome-phenome Archive, which is hosted by the EBI and the CRG, under accession number EGAS00001003755.

14. The Bayesian model depicted graphically in the supplemental figures seems to be missing key analysis of its performance in terms of AUC, Sn, Sp in the context of held out data.

Authors: We thank the Reviewer for this opportunity to clarify this point. The cfDNA samples of the 43 healthy control individuals used to train the hierarchical Bayesian model and learn the site-specific error rates have no ground truth by which the specificity can be evaluated. Furthermore, in the present context, the AUC is likely not a useful metric for evaluating the performance of variant calling since the underlying ROC would evaluate the test over regions of the ROC space in which one would not operate. The low specificity high sensitivity region would be uninteresting and inadequate for uncovering biological truth. Instead, the key characteristics of the hierarchical Bayesian model are detailed in the response to point #3 of this Reviewer’s comments, where we supplied false positive-recall curves directly measuring the effective false positive versus recall rates of biopsy-matched variants and determine reasonable cancer-specific thresholds on continuous scores returned by the hierarchical Bayesian model. The procedures used for doing so are detailed in the Methods of the manuscript under sections “Machine learning error model” and “Joint variant analysis using the machine learning error model” (on pages 30 and 33 of the original manuscript).

As for the validation cohorts reported in the manuscript, the 47 healthy control individuals have no orthogonal validation which would allow one to assess the performance in a supervised machine learning approach. Similarly, the 124 cancer patients reported in the manuscript were recruited for the purpose of evaluating this prototype high-intensity sequencing assay relative to their matched tumor sequencing. Although the variants detected in the matched tumor biopsy using the MSK-IMPACT assay are what comes closest to ground truth, due to intra-tumor heterogeneity and the lower sequence depth of the tissue assay which affords a lower sensitivity/power to detect low allele frequency subclonal variants, biopsy-matched variants by no means constitute the gold standard with which one could train a supervised model and assess the performance. Instead, the authors would like to refer the Reviewer to replicate experiments and the ddPCR experiments detailed at point #10 this Reviewer’s comments and the section “Sensitivity and specificity of the targeted DNA assay” of the manuscript involving titrations of gDNA from cell lines where ground truth is known (on page XX of the revised manuscript).

15. There are 2 technical experiments in the manuscript, (1) an experiment intended to assess the reproducibility of their assay after implementing a change in protocol (Fig 1b-c and Supp Fig 3) and (2) in a spike experiment to determine probability of detection of mutations at different allele frequencies (supp fig 2).

a. In experiment 1, the authors validate five mutations with ddPCR in one patient. While this is a good approach to validate the mutation calls of their assay, it is not applied broadly enough to be meaningful. The results based on validating these 5 mutations are very much overstated in that only 1/5 of the mutations validated has an AF < 0.5% (a KRAS mutation at 0.2% by ddPCR). Thus, the authors are only validating one of the many hundred low AF mutations called by their cfDNA assay in the study. However, in reference to this experiment at line 354 the authors state “This cfDNA sequencing approach allowed for robust de novo detection of somatic mutations (Fig. 1c) with a sensitivity similar to that of ddPCR (Fig. 1b).” This statement seems problematic for obvious reasons. Further, given the fact that they report 454 mutations < 0.1% in Supp Table 7, it would be useful to validate a more substantial portion of these low AF mutations (the VUSo mutations in particular) using ddPCR or some other method to empirically determine what fraction (if any) of these mutations are potential noise in their data.

Authors: We thank the Reviewer for the opportunity of addressing this comment through the provision of additional experiments. As detailed at point #10 of this Reviewer’s comments, in addition to the technical replicates provided in the manuscript, three patients have now been retested using version V2 of the protocol. The pairwise comparison of measured VAFs between versions V1 vs V2 and V2 vs V2 for all the samples that have been retested are shown **Response to Reviewers Figure 8**, whereas the number of variants from version V1 broken down per variant source category which have been validated across the replicates are displayed in **Response to Reviewers Tables** **7** and **8**. Overall, across the five non-hypermutated patients, 170 of variants detected in version V1 have been tested and 152 (89.4%) of these were also detected in version V2. In the hypermutated breast cancer patient, 630 of 659 (95.6%) of variants detected in version V1 were also detected in version V2. Comparing version V1 and V2 for the three patients who have been retested for the purpose of this response, 45 of 51 (88.2%) variants were found present in two non-hypermutated patients whilst the corresponding number in the hypermutated breast cancer patient was 640 of 659 (97.1%).

Additionally, as the VAF of these variants range between 0.3% and 91.2%, we limited the list to those which were initially detected in version V1 with VAF<1%. The results are summarized in **Response to Reviewers Tables 9** and **10**. Finally, following the Reviewer’s suggestion, additional ddPCR experiments for VUSo occuring at VAF<1% were performed for the revised version of the manuscript. This is detailed in our response at Comment #10. In summary, ddPCR measurements of selected VUSo mutations from 7 patients (6 of 7 mutations have VAF<1%) were performed. The experiments are limited by residual sample availability i.e. cfDNA or pre-enrichment library, VAF, and availability of validated ddPCR assay. The ddPCR and cfDNA targeted assay were found to have 100% PPA.

**Response to Reviewers Table 9**: Number of somatic mutations per patient detected in version V1 with VAF<1% and confirmed present in six initial technical replicates using version V2 of assay protocol

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Patient ID | No. of Biopsy -matched variants | No. of Biopsy -subthreshold variants | No. of VUSo | No. of WBC -matched variants | Total no. of variants |
| MSK-VB-0050 | 10 / 13 (76.9%) | 4 / 4 (100%) | 35 / 41 (85.4%) | 5 / 6 (83.3%) | 54 / 64 (84.4%) |
| MSK-VB-0041 | 0 (N/A) | 0 (N/A) | 0 (N/A) | 1 / 1 (100%) | 1 / 1 (100%) |
| MSK-VL-0028 | 4 / 4 (100%) | 0 (N/A) | 1 / 3 (33.3%) | 17 / 21 (81.0%) | 22 / 28 (78.6%) |
| MSK-VL-0042 | 0 (N/A) | 1 / 1 (100%) | 2 / 4 (50%) | 6 / 6 (100%) | 9 / 11 (81.8%) |
| MSK-VB-0023 | 1 / 1 (100%) | 44 / 46 (95.7%) | 479 / 502 (95.4%) | 10 / 12 (83.3%) | 534 / 561 (95.2%) |
| MSK-VL-0038 | 4 / 4 (100%) | 0 (N/A) | 2 / 3 (66.7%) | 11 / 12 (91.7%) | 17 / 19 (89.5%) |
| Total | 19 / 22 (86.4%) | 49 / 51 (96.1%) | 519 / 553 (93.9%) | 50 / 58 (86.2%) | 637 / 684 (93.1%) |

**Response to Reviewers Table 10**: Number of somatic mutations per patient detected in version V1 with VAF<1% and confirmed present in three additional technical replicates using version V2 of assay protocol

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Patient ID | No. of Biopsy -matched variants | No. of biopsy -subthreshold variants | No. of VUSo | No. of WBC -matched variants | Total no. of variants |
| MSK-VL-0028 | 4 / 4 (100%) | 0 (N/A) | 1 / 3 (33.3%) | 19 / 21 (90.5%) | 24 / 28 (85.7%) |
| MSK-VL-0042 | 0 (N/A) | 1 / 1 (100%) | 2 / 4 (50%) | 6 / 6 (100%) | 9 / 11 (81.8%) |
| MSK-VB-0023 | 1 / 1 (100%) | 45 / 46 (97.8%) | 487 / 502 (97.0%) | 11 / 12 (91.7%) | 544 / 561 (97.0%) |
| Total | 5 / 5 (100%) | 46 / 47 (97.9%) | 490 / 509 (96.3%) | 36 / 39 (92.3%) | 577 / 600 (96.2%) |

b. In the spiking experiment in Supp fig 2, a reference standard DNA was spiked into genomic DNA extracted from a cell line and the following mixtures were made in triplicate 0%, 0.1%, 0.25%, 0.5% and 1%. The mean depth of a replicate is stated to be 2,430X, and it is stated that the 3 replicates were pooled at the fastq level to simulate a higher input mass, resulting in a depth of ~4,577X. It is entirely unclear why the pooled depth us roughly 2X the depth of an individual replicate, rather than 3X. Shouldn’t the combined depth be closer to 7,200X? In the manuscript there are many mutations at positions with > 10,000 deduped depth (215 mutations in supp table 7) and there are numerous mutations at low AFs reported (as mentioned above) and I am insufficiently convinced that all these low AF mutations are somatic mutations and not noise. Also, a much better spike would utilize a high AF ctDNA sample that is spiked into healthy cfDNA as the error profile of cfDNA may be different from that of sonicated reference standard DNA and gDNA from a cell line as used here.

Authors: The spike-in experiments are detailed in section “Sensitivity and specificity of the targeted DNA assay” of the original manuscript. Five DNA titrations using the HD753 standard and NA12878 genomic DNA were prepared in triplicate i.e. fifteen titrations overall to have nominal expected variant allele fractions of 0, 0.1, 0.25, 0.5, and 1% for a majority of variants. The mean deduplicated and collapsed target coverage of all fifteen experiments was 2,430X. In addition, *in silico* spike-in experiments were performed to simulate higher input mass where at each given titration, half of each FASTQ was subsampled and combined to create three additional FASTQ i.e. if the triplicates are labelled A, B and C, the three simulated samples are AB = 0.5A+0.5B, AC = 0.5A+0.5C and BC = 0.5B+0.5C. Therefore, the mean target collapsed depth of these three simulated samples is approximately twice rather than three times that of a single replicate at any given titration.

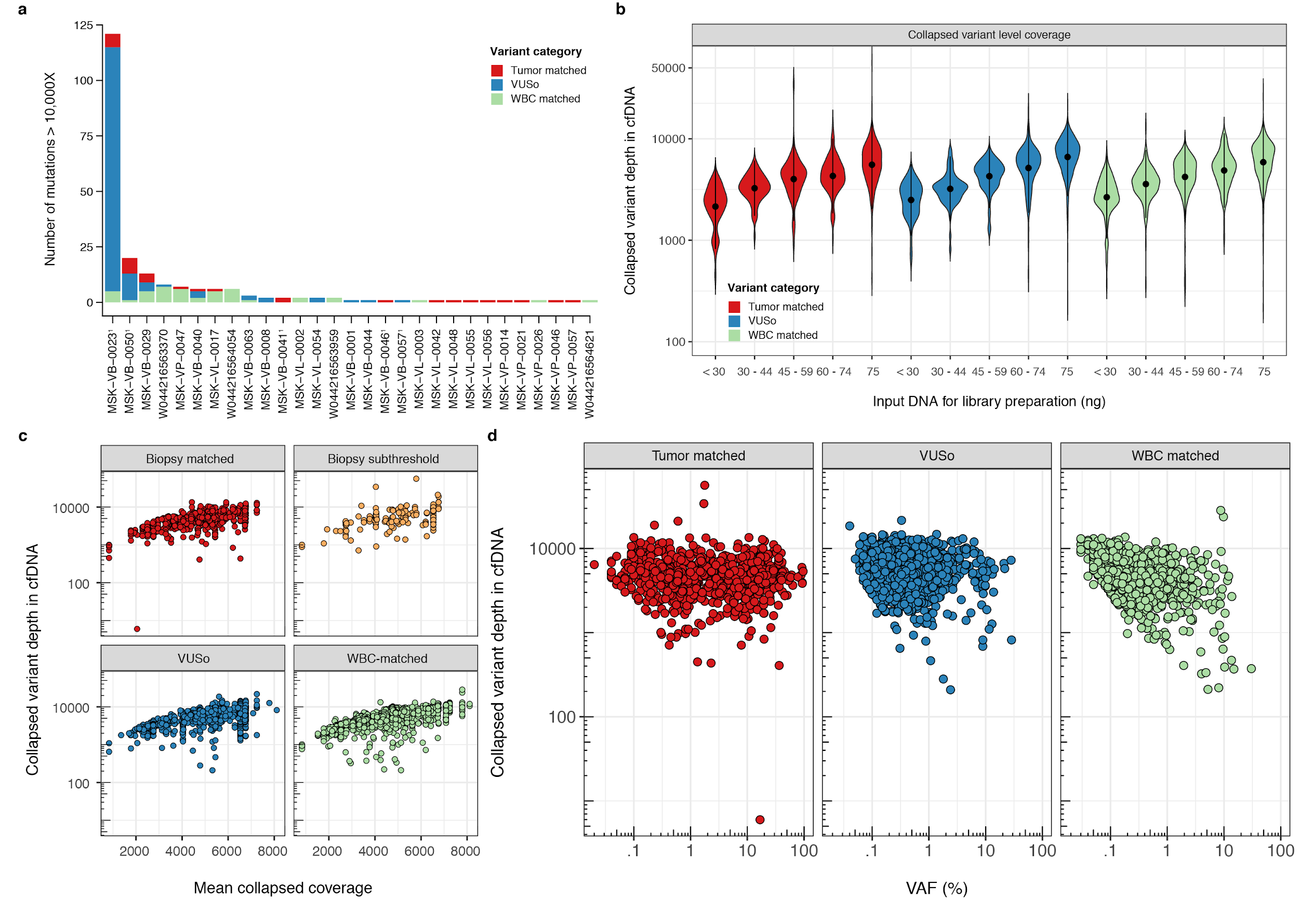
**Supplementary Table S7** of the original manuscript reports 215 somatic mutations labelled as either one of Biopsy-matched, Biopsy-subthreshold, WBC-matched or VUSo at collapsed depth >10,000X. **Response to Reviewers Table 11** and **Response to Reviewers Figure 14a** show that 121 (56.3%) and 20 (9.3%) of these mutations were detected in samples from patients MSK-VB-0023 and MSK-VB-0050, respectively, both of whom were hypermutated patients. Besides these two patients, only 74 mutations occurred with depth >10,000X of which 57 (77%) were also detected in the matched tumor biopsy or WBC. **Response to Reviewers Figures 14b** and **14c** demonstrate that the total depth of somatic mutations detected in cfDNA is a function of the mean collapsed target coverage in cfDNA which is itself a function of the amount of input DNA used for library preparation. Furthermore, there is no association between the VAF and the sequencing depth of variants irrespective of source of origin.

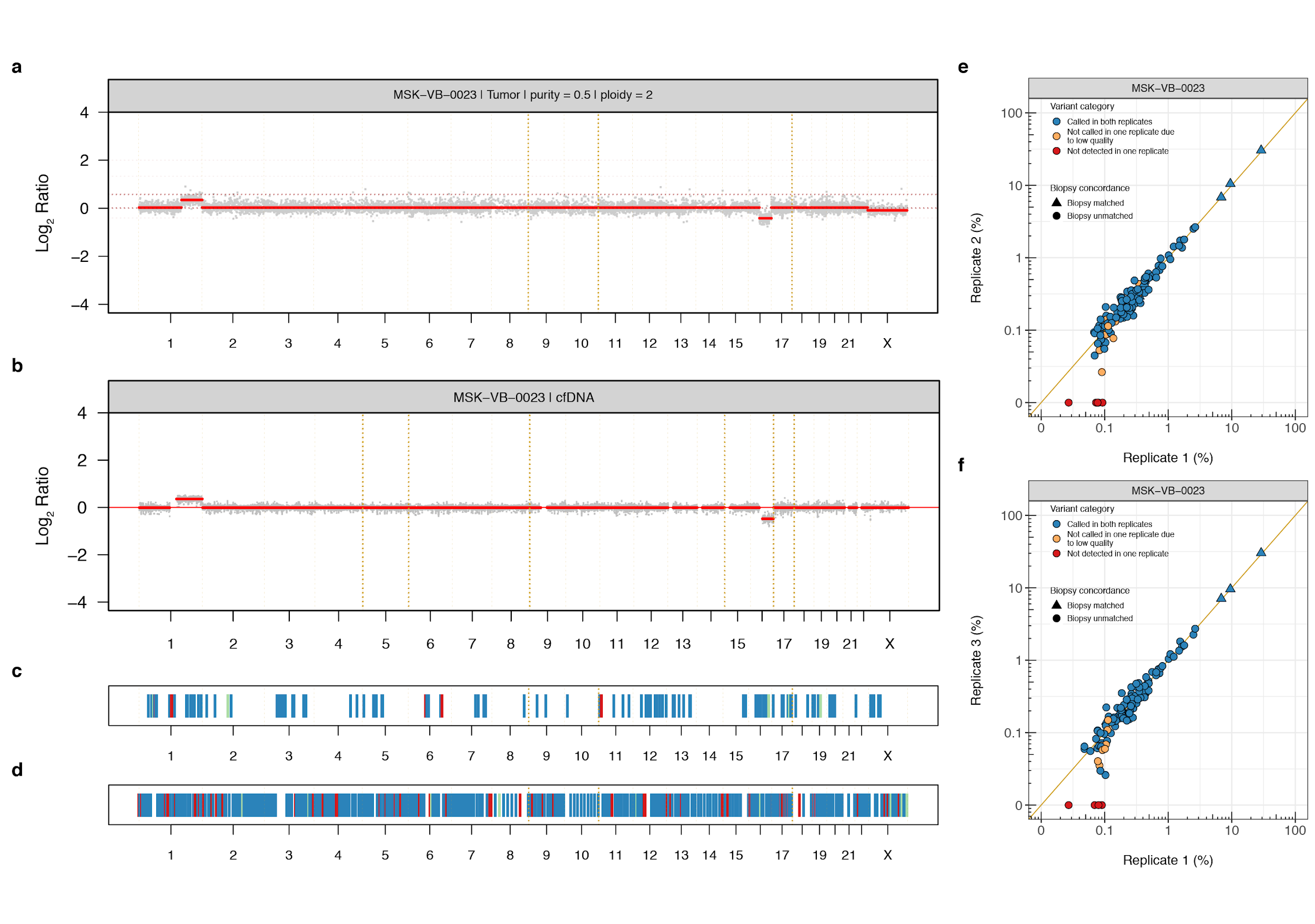
As the majority of variants detected at >10,000X occurred in patient MSK-VB-0023, we investigated whether these co-occurred with copy number amplifications. **Response to Reviewers Figures 15a** and **15b** show that other than a gain of chromosome arm 1q and a hemizygous loss of chromosome arm 16q, both of which were detected in cfDNA and the matched tumor biopsy, the profile of patient MSK-VB-0023 was devoid of any copy number alterations. The mutations detected at >10,000X depth, therefore, cannot be attributed to amplifications of the corresponding loci. As the pair cfDNA and WBC samples from patient MSK-VB-0023 was sequenced in triplicate, the pairwise comparisons of VAF of all 121 mutations occurring >10,000X across the three replicates are shown in **Response to Reviewers Figures 15e** and **15f**. We have now incorporated these data to **Supplementary Figure SX** of the revised manuscript.

**Response to Reviewers Table 11**: Somatic mutations detected at high sequencing depth in cfDNA

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Patient ID | Total no. of somatic mutations | Total no. of somatic mutations > 10,000X | No. of biopsy matched mutations > 10,000X | No. of biopsy subthreshold mutations > 10,000X | No. of WBC matched mutations > 10,000X | No. of VUSo mutations > 10,000X |
| MSK-VB-0023† | 659 | 121(18.4%) | 2 | 4 | 5 | 110 |
| MSK-VB-0050† | 86 | 20 (23.3%) | 4 | 3 | 1 | 12 |
| MSK-VB-0029 | 30 | 13 (23.3%) | 4 | 0 | 5 | 4 |
| W044216563370 | 15 | 8 (43.3%) | 0 | N/A | 7 | 1 |
| MSK-VP-0047 | 33 | 7 (21.2%) | 1 | 0 | 6 | 0 |
| MSK-VB-0040 | 24 | 6 (25%) | 0 | 1 | 2 | 3 |
| MSK-VL-0017 | 35 | 6 (17.1%) | 0 | 1 | 5 | 0 |
| W044216564054 | 23 | 6 (26.1%) | N/A | N/A | 6 | 0 |
| MSK-VB-0063 | 22 | 3 (13.6%) | 0 | 0 | 1 | 2 |
| MSK-VB-0008 | 17 | 2 (11.8%) | 0 | 0 | 0 | 2 |
| MSK-VB-0041 | 11 | 2 (18.2%) | 2 | 0 | 0 | 0 |
| MSK-VL-0002 | 15 | 2 (13.3%) | 0 | 0 | 2 | 0 |
| MSK-VL-0054 | 40 | 2 (5.0%) | 0 | 0 | 0 | 2 |
| W044216563959 | 3 | 2 (66.7%) | N/A | N/A | 2 | 0 |
| MSK-VB-0001 | 9 | 1 (11.1%) | 0 | 0 | 0 | 1 |
| MSK-VB-0044† | 110 | 1 (0.9%) | 0 | 0 | 0 | 1 |
| MSK-VB-0046† | 118 | 1 (0.8%) | 0 | 1 | 0 | 0 |
| MSK-VB-0057† | 55 | 1 (1.8%) | 0 | 0 | 0 | 1 |
| MSK-VL-0003 | 25 | 1 (4%) | 0 | 0 | 1 | 0 |
| MSK-VL-0042 | 18 | 1 (5.6%) | 1 | 0 | 0 | 0 |
| MSK-VL-0048 | 25 | 1 (4%) | 1 | 0 | 0 | 0 |
| MSK-VL-0055 | 46 | 1 (2.2%) | 0 | 1 | 0 | 0 |
| MSK-VL-0056 | 17 | 1 (5.9%) | 1 | 0 | 0 | 0 |
| MSK-VP-0014 | 22 | 1 (4.5%) | 0 | 1 | 0 | 0 |
| MSK-VP-0021 | 13 | 1 (7.7%) | 0 | 1 | 0 | 0 |
| MSK-VP-0026 | 21 | 1 (4.8%) | 0 | 0 | 1 | 0 |
| MSK-VP-0046 | 7 | 1 (14.3%) | 0 | 1 | 0 | 0 |
| MSK-VP-0057 | 14 | 1 (7.1%) | 1 | 0 | 0 | 0 |
| W044216564621 | 16 | 1 (6.3%) | N/A | N/A | 1 | 0 |
| Total | 1529 | 215 (14.1%) | 17 (7.9%) | 14 (6.5%) | 45 (20.9%) | 139 (64.7%) |

† denotes hypermutated samples

**Response to Reviewers Figure 14: Somatic mutations occurring at high sequencing depth in cfDNA.** Somatic mutations detected at sequencing depth >10,000X in cfDNA occur mostly in hypermutated samples and are related to sample level mean target collapsed depth which is itself a function of the amount of input DNA used for library preparation. Shown in (a), the number of somatic mutations occurring at >10,000X per patient and categorized into WBC-matched, VUSo or Tumor-matched where the latter category is composed of Biopsy-matched and Biopsy-subthreshold mutations, (b) variant level collapsed depth for all somatic mutations detected in cfDNA categorized in Tumor-matched, VUSo or WBC-matched and grouped according to the amount of input DNA used for library preparation, (c) variant level collapsed depth for all somatic mutations detected in cfDNA against sample level mean collapsed target depth and (d) variant level collapsed depth for all somatic mutations against the modeled VAF in cfDNA. In (a), ‘1’ denotes hypermutated samples.

**Response to Reviewers Figure 15: Somatic mutations occurring in the cfDNA of hypermutated patient MSK-VB-0023.** 121 of 215 (56.3%) somatic mutations detected at sequencing depth >10,000X in cfDNA occurred in the hypermutated patient MSK-VB-0023. Panels (a) and (b) show the Log2 Ratio profiles of the tumor biopsy and cfDNA, respectively. Panels (c) and (d) show the genomic coordinates of somatic mutations detected in cfDNA at >10,000X and ≤10,000X, respectively. The tumor biopsy and cfDNA showed similar copy number alterations i.e. 1q+/16q- and somatic mutations did not cluster at specific genomic loci nor did they co-occur with copy number amplifications. Three replicate sequencing of cfDNA and WBC were available for patient MSK-VB-0023. Panels (e) and (f) show the pairwise comparisons of VAF for the 121 mutations detected at >10,000X in the index cfDNA sequencing reported in the manuscript. In (c) and (d), the variants are color coded according to their origin i.e. Tumor biopsy-matched, VUSo and WBC-matched. In (e) and (f), the variants are shape coded based on their origin (i.e. whether they were also detected in the matched tumor biopsy and color coded according to their category; whether they were called in both replicates and assigned to similar source categories i.e. VUSo, WBC-matched or noise). The methods used to infer the copy number alterations are described in response to Reviewer 2’s Comment #9.

Minor comments:

1. It is unclear what the authors mean by “high intensity sequencing”, which reads a bit like advertising. If considering intensity to mean the depth of sequencing allotted for a given input of DNA, then several prior papers have exceeded the depths targeted by the authors, which is only relevant in the context of the achieved error profile from UMI-informed suppression of PCR artifacts. If considering the breadth of sequencing as intensity, then prior studies that have targeted broader portions of the genome, including the whole genome or whole exome would seem to be higher intensity. If considering multiple specimens/analytes from the same subject as the basis for their intensity, then multi-regional sequencing studies of tumours, premalignant lesions, plasma, etc have far exceeded this current study in intensity. If the authors consider their seemingly unique combination of depth/breadth/samples profiled as “high intensity”, then it would seem important to much better capture how this unique combination informs a useful tool when considering biological discovery, clinical utility, etc

Authors: The authors thank the Reviewer for this insightful suggestion. “high intensity sequencing” in fact refers to the unique combination of depth/breadth/sample profiled in this study. We have now revised the text to better emphasize this point (page XX, paragraph XX).

**Reviewer #2:**

Remarks to the Author:

The manuscript by Razavi and colleagues reports the development and validation of a technology to assess mutational characterization in plasma samples from patients with solid tumors. The approach robustly detects somatic mutations related to the tumor but also identifies clonal hematopoiesis as a source of additional mutations present in plasma.

This work addresses critical concepts in the era of molecular characterization of cancer patients by using liquid biopsy/plasma samples with direct clinical implications. The development of a broad gene-sequencing assay with high sensitivity and the analysis of the origin of these variants detected are of utmost importance. Indeed, the identification and analysis of the role of clonal hematopoiesis as a source of gene variants that may not be related to the tumor under scope has direct clinical implications.

There are a few suggestions to clarify data presented in the manuscript and improve on the relevance of the clinical findings:

1. The authors have analyzed some of the most prevalent tumor types. However, no group of colorectal cancer patients has been included. What was the reason for skipping this segment of patients? Were there some technical issues? Would it be possible to include a similar group of colorectal cancer patients?

Authors: We thank the Reviewer for these most pertinent questions. The patients were prospectively recruited for the purpose of this study. The selection of breast, lung and prostate cancers was based on the fact that these three cancer types cover the most common cancers both genders, and it is, regrettably, not possible to perform additional analysis of colorectal cancer patients as they have not been included in this prospective study. We appreciate the Reviewer’s suggestion and have added to the paragraph outlining the limitations of the study in the discussion of the manuscript (page XX, paragraph XX), the following passage: “Colorectal carcinomas, another common form of cancer, were not included in this study; hence, further studies are warranted to define the sources of cfDNA mutations in patients in this disease”.

2. The authors describe the use of two versions of the protocol V1 and V2 (Figure 1b). It would be of value to detail what was the distribution of the protocol used among samples and controls?

Authors: We apologize for the lack of clarity. We have now provided this information in **Supplementary Table S10** of the revised manuscript. The samples used for training the noise model were sequenced using both versions V1 and V2 of the assay and internal cross-validation showed no apparent differences in error control between the two assays. As shown in the replicate analyses (**Response to Reviewers Figure 8**), the results from versions V1 and V2 of the assay were highly concordant on samples harboring biological variants. Owing to the timing of assay development and availability of the samples, all the cancer patients were sequenced using version V1 as were approximately half of the training samples. The V2 assay was used for further experiments: replicates as a bridging study between the V1 and V2 assays, a set of control individuals run on V2 allocated to additional noise model training, and an independent set of non-cancer individuals run on V2 reported in this manuscript to demonstrate error control. Given the equivalence of the results of the assays observed during the technical development of the assay and in the direct comparison of actual cases (**Response to Reviewers Figure 8**), we have considered the assays equivalent in their performance.

3. A good correlation is found between ddPCR and the targeted DNA assay in the 5 cases with high amount of cfDNA, although protocol V1 provides lower MAF in the variant KRAS G12A that has a ddPCR MAF of around 0.2%. It would be important to confirm the reproducibility of mutations with low MAF from some additional samples

Authors: We thank the Reviewer for bringing this important point to our attention. We have now performed additional technical repeats using V1 and V2 of the high-intensity sequencing assay and additional ddPCR testing VUSo with low VAFs. Please see our response to Reviewer #1’s comment #10, and this Reviewer’s comment #5. The authors kindly refer the Reviewer to the corresponding sections for a more elaborate discussion of the results. Briefly here, in the technical replicates utilizing V1 and V2 of the high-intensity sequencing assay, the pairwise comparison of VAFs between versions V1 vs V2 and V2 vs V2 for all the samples that have been retested are shown in the **Response to the Reviewers Figure 8** and the revised **Supplementary Figure S3** of the manuscript. The number of variants from version V1 broken down per variant source category which have been validated across the replicates are displayed in **Response to Reviewers Tables 7 and 8**. Overall, across the five non-hypermutated patients, 170 of variants detected in version V1 have been tested and 152 (89.4%) of these were also detected in version V2. In the hypermutated breast cancer patient, 630 of 659 (95.6%) of variants detected in version V1 were also detected in version V2. Comparing version V1 and V2 for the three patients who have been retested for the purpose of this response, 45 of 51 (88.2%) variants were found present in two non-hypermutated patients whilst the corresponding number in the hypermutated breast cancer patient was 640 of 659 (97.1%). As for the ddPCR assays, in the revised version of the manuscript, we have now assessed the performance of the high-intensity sequencing assay for detecting VUSo and accurately measuring VAF. For this analysis, we first identified the subset of cfDNA VUSo for which a validated ddPCR assay was available. We then identified all the patients with residual extracted cfDNA and/or leftover pre-enrichment sequencing libraries harboring any of these mutations. Our final analysis included four ddPCR assays (*PIK3CA* E545K, *NRAS* Q61K, *CTNNB1* T41I, *CTNNB1* S33C) targeting VUSo in seven patients. For one patient, only cfDNA extracted from plasma was available. For two patients, both cfDNA and pre-enrichment sequencing libraries were available and tested, whereas for the remaining four patients, only pre-enrichment libraries were available. As a negative control, the pre-enrichment sequencing libraries from 12 patients where the target VUSo were not detected in cfDNA were used. Four negative libraries were used for each ddPCR assay and all experiments were performed in triplicate. The results are shown in **Response to Reviewers Figure 9b** and **Figure 4** of the revised manuscript whereby the cfDNA sequencing assay had a 100% PPA and 100% NPA considering ddPCR as the benchmark.

4. In the technique validation some variants are not called in one of the replicates due to WBC filtering (green dots, figures 1c and suppl fig3). How can this observation be interpreted? In general variants present in WBC have MAFs higher than 1% that should be detected with good reproducibility

Authors: The category “Not called in one replicate due to WBC-filtered” in **Figure 1** and **Supplementary Figure S3** of the manuscript is composed of low allele fraction variants (<1%) which, although they are definite present, are classified in one of the replicates as VUSo and WBC-matched in the other, rather than being filtered out as germline variants. This occurs as a result of varying alternate read support in the matched WBC sequencing and to ensure the highest confidence in those mutations which are detected as VUSo, even a low number of supporting alternate reads in the matched WBC would be deemed sufficient to classify a given mutation as WBC-matched. Across the six patients who were chosen to assess the reproducibility of the assay, only 75 of the 1,726 (4.3%) variants were found to be misclassified in one or the other replicate. **Response to Reviewers Table 12** summarizes those variants that were discordant in the five non-hypermutated patients.

We agree with the Reviewer that the labelling of **Figure 1** and **Supplementary Figure S3** of the manuscript may be misleading. This important issue raised by this Reviewer and Reviewer 3, however, strengthens our critical argument in response to Reviewer 1, namely that cfDNA sequencing without matched WBC is likely to identify CH-derived somatic mutations which may be misinterpreted as cancer derived. As this is most likely to occur close to or at the limit of detection, it stresses the need for ultra-deep sequenced WBC and further emphasizes the timeliness and importance of the findings reported in our study. The legend of **Figure 1** and **Supplementary Figure S3** of the manuscript have been changed to “Incorrect assignment between replicates” to reflect the above discussion and the captions have been revised accordingly.

**Response to Reviewers Table 12**: Somatic mutations with incorrect assignment of variant category between replicates

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Patient ID | Gene Symbol | HGVSp | Variant Type | cfDNA VAF (%) | Alternate allele count in cfDNA | WBC VAF (%) | Alternate allele count in WBC |
| MSK-VB-0050 | *NOTCH2* | Q2367E | SNV | 0.077 | 0.154 | 5 | 11 | 0 | 0 | 0 | 0 |
| MSK-VB-0050 | *TET1* | S919\* | SNV | 0.055 | 0.098 | 7 | 13 | 0 | 0 | 0 | 0 |
| MSK-VB-0050 | *ARID2* | L1452V | SNV | 0.073 | 0.044 | 10 | 6 | 0 | 0.018 | 0 | 1 |
| MSK-VB-0050 | *KMT2D* | E1136\* | SNV | 0.103 | 0.059 | 14 | 8 | 0 | 0 | 0 | 0 |
| MSK-VB-0050 | *MAPK3* | F346L | SNV | 0.095 | 0.080 | 13 | 11 | 0 | 0.026 | 0 | 1 |
| MSK-VB-0050 | *MAPK3* | S159Y | SNV | 0.134 | 0.061 | 12 | 6 | 0 | 0 | 0 | 0 |
| MSK-VB-0050 | *FANCA* | L910V | SNV | 0.109 | 0.105 | 9 | 10 | 0 | 0 | 0 | 0 |
| MSK-VB-0050 | *INSR* | S54C | SNV | 0.094 | 0.178 | 10 | 20 | 0 | 0 | 0 | 0 |
| MSK-VB-0050 | *PIK3CG* | Q432H | SNV | 0.052 | 0.090 | 5 | 9 | 0.023 | 0 | 1 | 0 |
| MSK-VL-0038 | *TET1* | K1162M | SNV | 0.109 | 0.087 | 6 | 7 | 0.030 | 0 | 1 | 0 |
| MSK-VL-0038 | *STAG2* | Y1155\* | SNV | 0.232 | 0.100 | 11 | 7 | 0.035 | 0.053 | 1 | 2 |
| MSK-VL-0042 | *TET2* | Q1534\* | SNV | 0.078 | 0.075 | 7 | 10 | 0 | 0.080 | 0 | 4 |
| MSK-VL-0042 | *AMER1* | G432A | SNV | 0.083 | 0.079 | 6 | 10 | 0 | 0 | 0 | 0 |

5. The identification of VUSo in the different cohorts analyzed provides challenging information. Were these VUSo somehow confirmed? Replicates? ddPCR of specific variants? The low MAFs detected suggest the possibility of artifacts and it would be important to confirm these results. Were VUSo analyzed in sequential samples? Are they stable in time?

Authors: We thank the Reviewer for highlighting this issue and the opportunity to strengthen our findings. The manuscript included technical replicates carried out to test reproducibility using two different versions (V1 and V2) of the assay for cfDNA and WBC. Samples from six patients, detailed in **Supplementary Table S6** of the manuscript, were selected for reprocessing with both assay protocols. The results are shown in **Figure 1** and **Supplementary Figure S3** of the manuscript where the measured VAFs between the two technical replicates for samples from the five non-hypermutated patients showed a strong agreement (*R2*= 0.9997, **Figure 1C**) as well as for one sample from a hypermutated case (*R2*= 0.9972, **Supplementary Figure S3**). We now performed additional sequencing of replicates for three of the above patients utilizing version V2 of the assay. The pairwise comparison of VAFs between versions V1 vs V2 and V2 vs V2 for all the samples that have been retested are shown **Response to Reviewers Figure 8**.

**Response to Reviewers Table 7** (in response to Reviewer 1’s Comment #15) further summarizes the number of Biopsy-matched, Biopsy-subthreshold, VUSo and WBC-matched variants detected in the index cfDNA sequencing of those six patients i.e. using version V1 of the assay and reported in Figures 2 to 5 of the manuscript together with the percentages of those variants which were confirmed present (i.e. have non-zero alternate read support irrespective of variant source category using version V2). Additionally, **Response to Reviewers Table 8** (in response to Reviewer 1’s Comment #15) shows the same for the three patients who were selected for retesting.

Overall, 647 of 829 (78.0%) mutations were classified as VUSo. Comparing versions V1 and V2 as reported in **Figure 1** and **Supplementary Figure S3** of the original manuscript, 615 of 647 (95.1%) mutations classified as VUSo were positively validated. Additionally, comparing versions V1 and V2 for the three patients who were selected for retesting, 564 of 583 (96.7%) of VUSo mutations were positively validated.

Additionally, following the Reviewer’s comment, we performed additional ddPCR assays targeting VUSo in seven patients with leftover cfDNA or pre-enrichment libraries or both. We kindly refer the Reviewer to our response to Reviewer 1’s Comment #10 and **Response to Reviewers** **Figure 9b**. Briefly, our additional analyses demonstrated 100% PPA and 100% NPA, considering ddPCR as the ‘gold standard’.

6. It would be important to incorporate the distribution of VUSo among genes in controls as part of figure 2b to infer the possible confounding effect when assessing patient samples

Authors: We agree with the Reviewer’s comment and have now provided the genes harboring VUSo in healthy control individuals as part of the revised **Figure 2b** and **Supplementary Figure S8** (see also **Response to Reviewers** **Figure XX**) together with the full list of these mutations in **Supplementary Table S7** of the revised manuscript.

7. In page 7, line 174, the authors say “somatic mutations with the highest VAF”, it would be informative to include the mean/median value of this VAF

Authors: **Figure 2** of the original manuscript shows the VAF in cfDNA of all variants classified as Biopsy-matched, Biopsy-subthreshold and VUSo across all patients. The mean and median of the highest VAF mutations are 15.10% and 9.18%, respectively. Following the Reviewer’s suggestion, the manuscript has been revised to provide the mean/median value of the VAF (page XX, paragraph XX)

8. The number of patients with high TMB is unfortunately very low (only 6), to confirm the capacity of the system in robustly identifying this marker, although the results are promising. Interestingly 10 samples showed high TMB in CfDNA. The variants defining these high TMB were probably obtained after filtering for WBC, but this should be indicated. The authors justify that the TMB is not detected in breast and prostate cancers because in this patients the hypermutator phenotype is acquired later in evolution. And in NSCLC cases? How could the differences between CfDNA and tumor biopsy be explained? Lapse in sample collection? Additional therapy?

Authors: The inclusion criteria of this study are outlined in section “Patient enrolment” of the manuscript. Briefly, patients with MBC, NSCLC, or CRPC with disease progression as assessed by the investigator were eligible. Disease progression was based on objective radiographic and/or physical exam and/or biomarker results. Patients diagnosed with *de novo* or recurrent stage IV NSCLC or breast cancer were allowed to be included if enrolled prior to initiation of the first line of treatment for metastatic disease. Patients who initiated new therapies between tissue biopsy and blood draw were not eligible. Patients with progressive disease on stable doses of treatment (e.g. hormone therapy) were eligible. Blood was drawn within 6 weeks of tissue biopsy for MSK-IMPACT analysis either prior to or after tissue biopsy. Whole blood samples received outside of the stability timeframe for Streck DNA BCT (5 days) were excluded.

The tumor mutation burden (TMB) was defined as the number of mutations per megabase (Mb) in the matched tumor biopsies and was computed from somatic non-synonymous mutations reported as part of the clinically validated MSK-IMPACT pipeline (PMID: 2580182, PMID: 28481359), namely Biopsy-matched and Biopsy-only categories using the definitions in the manuscript. The cfDNA mutation burden was computed from somatic non-synonymous mutations categorized as VUSo, Biopsy-matched and Biopsy-subthreshold. Both estimates were restricted to mutations occurring within 1Mb of the human genome common to both assay designs. The threshold of mutation burden used to define hypermutated patients was 13.8 mutations/Mb for the tumor biopsy (PMID: 28481359) whereas the corresponding value for cfDNA was computed *de novo* as the median (cfDNA mutation burden) + 2 × IQR (cfDNA mutation burden), where IQR is the interquartile range. A threshold of 22.7 mutation/Mb was used for the cfDNA assay.

The mutation burden is a compound measure which depends on the region of genome sequenced, the clonal structure of the sample being assessed, the sensitivity of the assay to detect subclonal mutations and the purity of the sample which, in the case of cfDNA, is the ctDNA fraction estimate. As illustrated in the response to this Reviewer’s Comment #9 and **Response to Reviewers** **Figure 15**, there is a larger number of breast and prostate cancer patients displaying a high ctDNA fraction compared to lung cancers. The tumor purities estimated using FACETS (PMID: 27270079) from copy number alterations and ctDNA fraction estimates of the patients defined as hypermutated using either the tumor or cfDNA assays are displayed in **Response to Reviewers** **Table 13**. For the two lung cancer patients found to be hypermutated based on the tumor biopsy sequencing only, the ctDNA fraction estimates were below 10%, providing the basis for the discrepancy between the biopsy and cfDNA results. Conversely, for the five cases found to be hypermutated through the cfDNA assay only, based on the known intra-tumor genetic heterogeneity elegantly documented in the numerous studies published by Charles Swanton based on the analysis of the samples from the TRACERx (PMID: 28445469), it is plausible that despite having a high tumor purity, the VUSo mutations contributing to the high cfDNA mutation burden are either (1) late evolving at the same tumor site, (2) present in parts of the tumors which were not sampled by the biopsy procedure or (3) found in tumor metastases at different anatomical locations. Thus, despite the limitations imposed by low purities, this heterogeneity is more readily captured through cfDNA sequencing, which emphasizes the importance of this study. The authors thank the Reviewer this most pertinent of observations. The manuscript has been revised accordingly to reflect the above discussion.

**Response to Reviewers Table 13**: Mutation burden and purity estimates for hypermutated cases defined through tumor biopsy or cfDNA sequencing. Discordant cases are highlighted.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Patient ID | TMB | cfDNA mutation burden | Tumor purity (%) | ctDNA fraction (%) |
| MSK-VB-0023 | 14.5 | 585.6 | 50.0 | 59.7 |
| MSK-VB-0044 | 5.45 | 94.4 | 80.3 | 30.6 |
| MSK-VB-0046 | 1.82 | 106.2 | 58.9 | 11.4 |
| MSK-VB-0050 | 14.5 | 72.6 | 90.2 | 25.3 |
| MSK-VB-0057 | 20.0 | 48.1 | 77.3 | 19.0 |
| MSK-VL-0035 | 30.0 | 55.4 | 37.1 | 19.7 |
| MSK-VL-0054 | 5.45 | 32.7 | 40.4 | 22.2 |
| MSK-VP-0031 | 2.72 | 33.6 | 30.0 | 14.9 |
| MSK-VP-0054 | 4.54 | 30.9 | 74.3 | 31.0 |
| MSK-VP-0041† | 15.4 | 39.0 | 80.0 | 39.1 |
| MSK-VL-0060 | 16.3 | 20.0 | 40.1 | 8.56 |
| MSK-VL-0065 | 18.2 | 18.2 | 30.0 | 0.19 |

† denotes the MSI-high prostate cancer

9. The approach used to characterize the molecular profile should allow for detection of CNVs and fusion genes. What results have been obtained in this area in the three groups of patients? Is the system reliable to detect these alterations which may be relevant for treatment decision making in the subset of patients analyzed? And in controls? Were any CNVs and/or gene fusions detected that may condition the interpretation of these findings in patient samples?

Authors: We greatly appreciate the Reviewer’s comment and acknowledge the importance and clinical relevance of accurate detection of copy number alterations and fusion genes utilizing cfDNA. We would like to first point out that the current cfDNA assay had not been optimized to detect CNVs and/or fusion genes, as this was not one the aims that motivated the development of the assay or this study. Given the importance of the Reviewer’s insightful comment, although we could not perform an analysis of the detection of fusion genes, as the assay did not include the intronic regions involved in fusion that would potentially allow for their detection, we still performed an exploratory, hypothesis generating analysis of CNVs.

For the CNV exploratory analysis, we first computed the Log2 Ratios, absolute copy numbers, purity and ploidy of the tumor biopsies using FACETS (PMID: 27270079). The raw read counts of the cfDNA and WBC samples were extracted from deduplicated uncollapsed BAM files and processed using CNVkit (PMID: 27100738) to obtain an estimate of Log2 Ratios for both on- and off-target regions. The raw read counts were corrected for library size, GC content and target length after which the WBC samples were averaged and used as baseline to normalize each cfDNA sample. The resulting Log2 Ratios were smoothed using the median absolute deviation and segmented by penalized least squares regression using the R/Bioconductor library copy number (PMID: 23442169) with default parameter values, except for the minimum number of probes in each segment and the penalty parameter for allowing discontinuities in Log2 Ratios. Absolute copy numbers were inferred from the segmented Log2 Ratios based on the ctDNA fraction estimates using the following equation:

where is the absolute copy number of a given segment, is the ctDNA fraction estimate of the sample, is the ploidy of the cfDNA sample, is the Log2 Ratio of the given segment and is a compression ratio. Since is unknown, we use an iterative approach to minimize the weighted sum of squared errors (SSE) over a range of ploidy values according to the following equation:

where is the length of segment and is the number of segments in a given sample. The optimal ploidy is given by:

Amplifications and homozygous deletions were inferred for the 410 genes included in the MSK-IMPACT assay based on the amplitude of and as follows:

where is the ordinal copy number call of a given gene, and are ploidy specific thresholds such that:

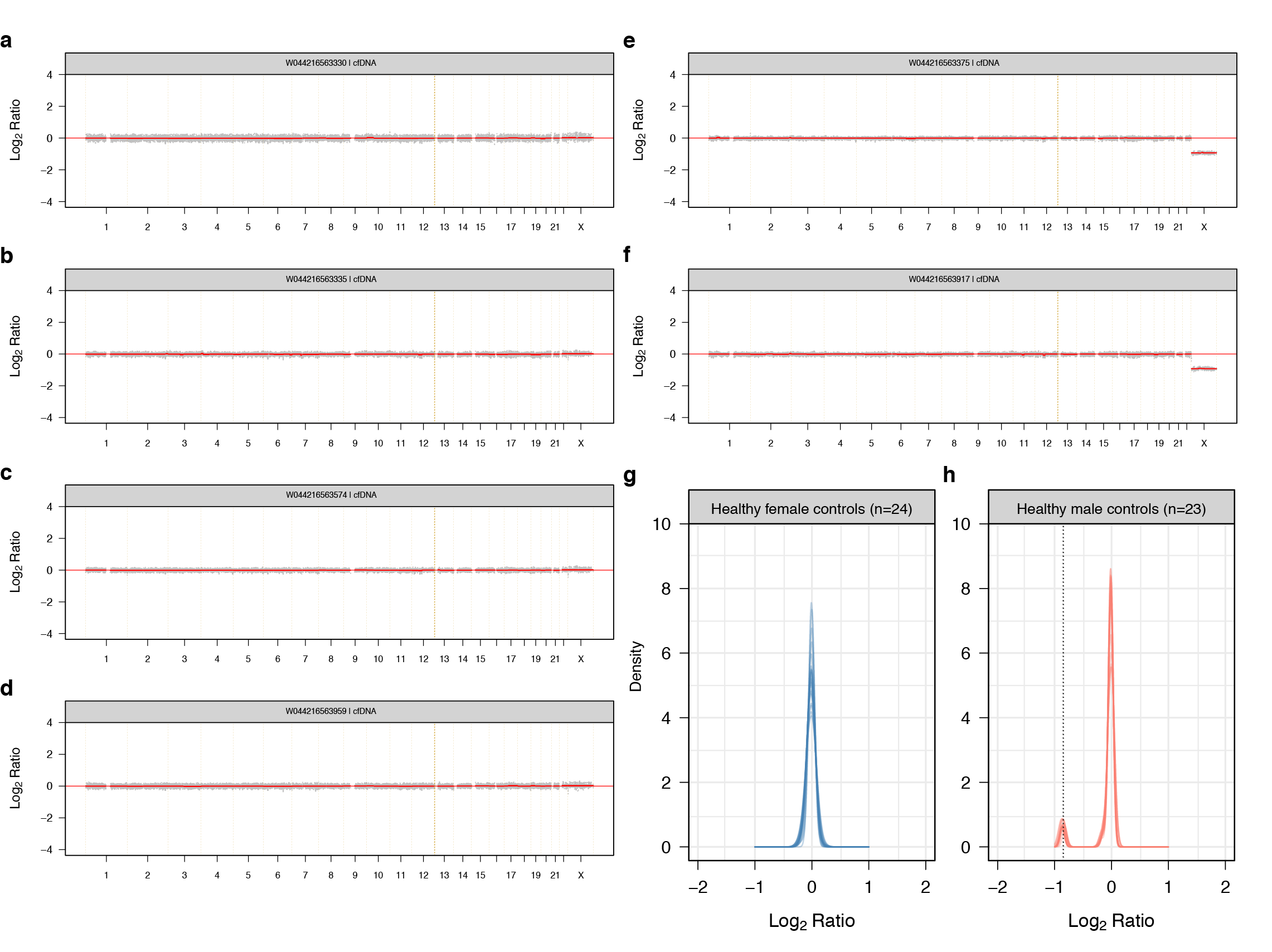
The Log2 Ratios of representative healthy male and female controls are shown in **Response to Reviewers Figures 16a-16f**. The low variance of raw Log2 Ratios observed is expected due to the ultra-high depth sequence coverage achieved for the cfDNA and WBC samples. Since no systematic technical artefact was observed in the healthy controls, these data can be used to derive parameter estimates for the analysis of cancer samples. There is, usually, a platform specific compression of Log2 Ratio (PMID: 20837533) which describes the ratio of observed to expected Log2 Ratio for a clonal loss of one copy in a diploid background in a sample with 100% purity. The compression ratio was estimated from the mean segmented Log2 Ratio of chromosome X across healthy male controls (**Response to Reviewers Figures 16g and 16h** below).

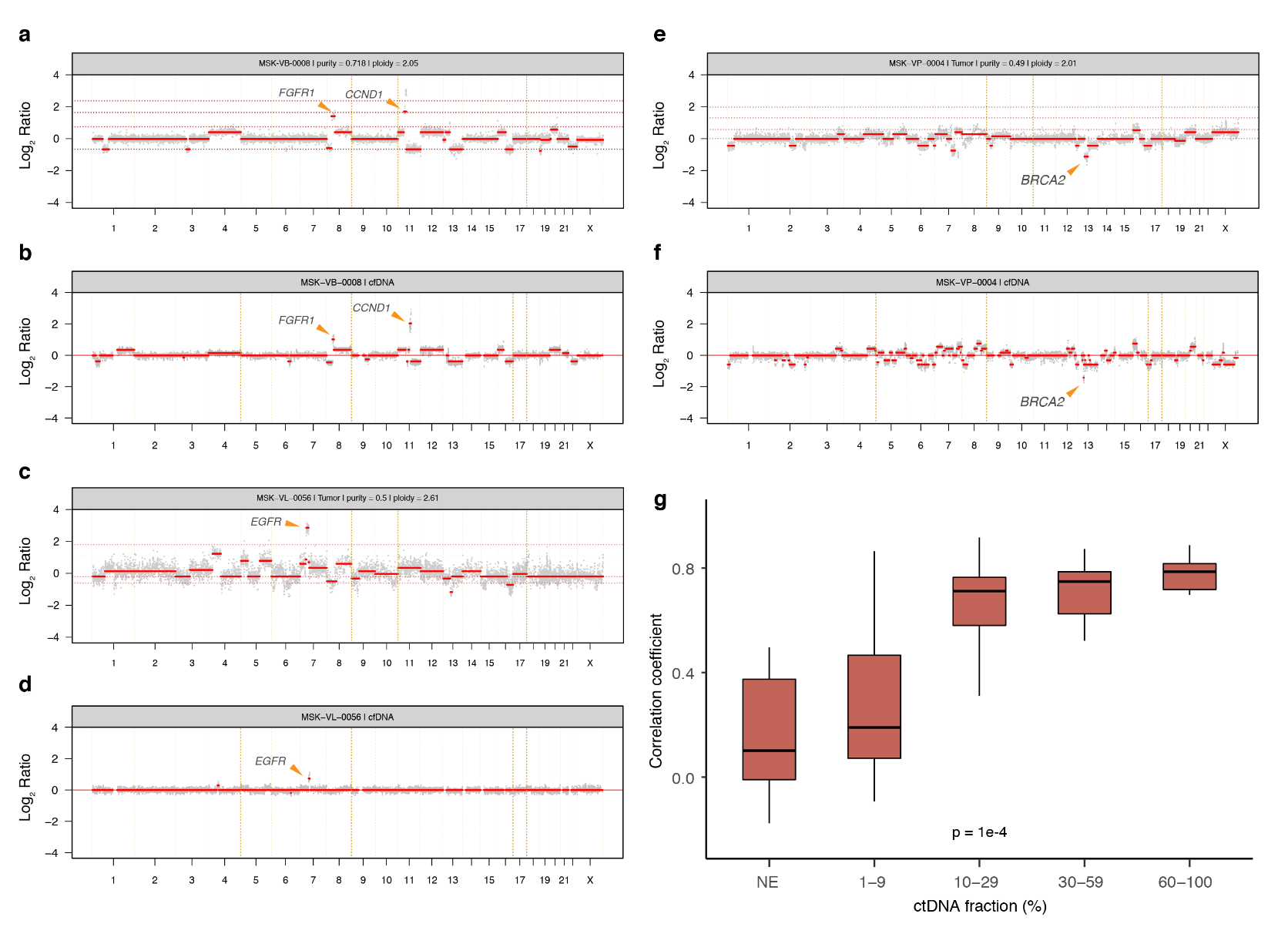
A comparison of Log2 Ratios of three tumor biopsy and matched cfDNA sample pairs, one of each cancer types, where amplifications (*FGFR1* and *CCND1* in the breast cancer case, *EGFR* in the lung cancer case) and a homozygous deletion (*BRCA2* in the prostate cancer case) were detected in the tumor biopsies is shown in **Response to Reviewers Figures 17a-17f** of this response. Based on these initial observations, the concordance between the tumor biopsies and cfDNA samples was evaluated using two complementary approaches:

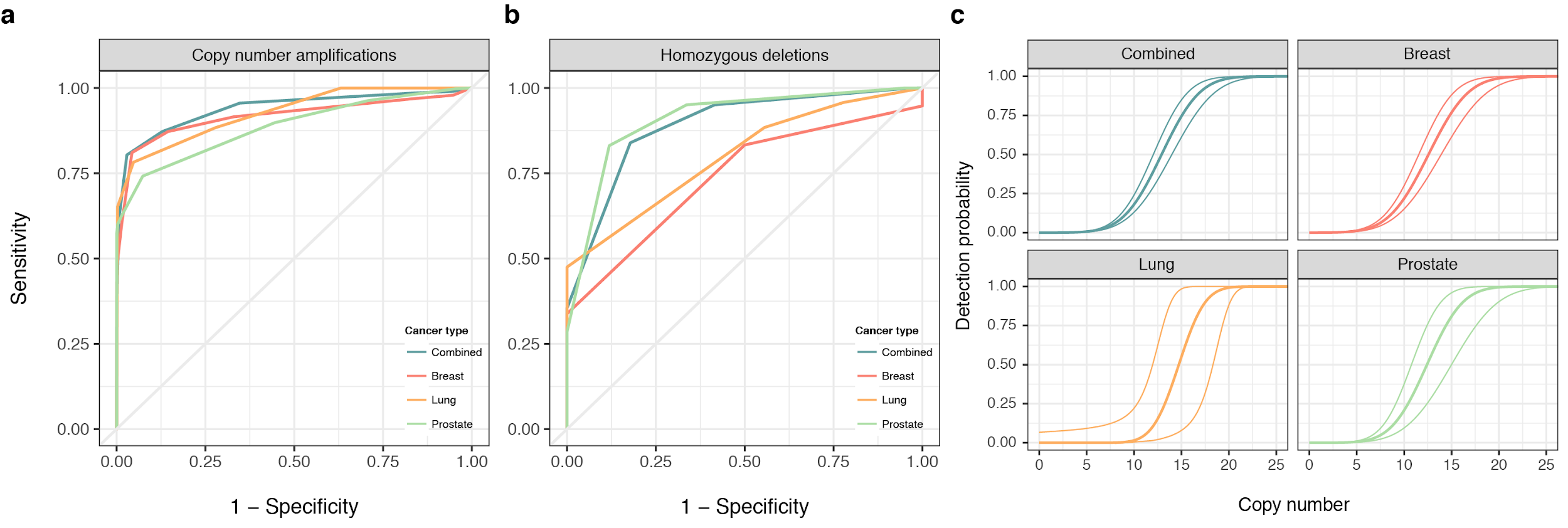
1. As the correlation of segmented Log2 Ratios for segments overlapping ≥75%. The distribution of Pearson’s *r* is shown in **Response to Reviewers Figure 17g-17h**. There was a statistically significant association between the Pearson’s *r* and the ctDNA fraction (one-sided Jonckheere-Terpstra test p = 1e-4). A larger number of the breast and prostate cancer cases had high ctDNA fractions and Pearson’s *r* > 0.5. There was no association between the Pearson’s *r* and the purity of the matched tumor biopsy.
2. Based on the previous observation that the Log2 Ratios estimated from the tumor biopsy and cfDNA are comparable only for samples with high ctDNA fraction estimates, we compared amplifications and homozygous deletions in any of the 410 genes included in the MSK-IMPACT assay for only N = 49 tumor biopsy and cfDNA sample pairs with ctDNA fraction estimate ≥10%. The ROC curves for predicting amplifications or homozygous deletions from the absolute copy numbers in cfDNA are shown in **Response to Reviewers Figures 18a** and **18b**. Pairwise comparisons between the cancer cohorts showed that none of the AUC for amplifications were significantly different (DeLong two-sided test for AUC; p > 0.05) except comparing breast and prostate cancers (DeLong two-sided test for AUC; p = 0.042). This was expected given that breast cancers had the highest number of amplifications whilst prostate cancers harbored the fewest. Similarly, comparing homozygous deletions across the different cancer cohorts showed that prostate cancers had a statistically higher AUC than breast (p = 3.95e-3) and lung (p = 8.56e-2) cancers possibly explained by the higher number of such alterations in prostate cancers. We also estimated the probability of detecting an amplification in cfDNA as a function of the amplitude of the corresponding copy number in the matched tumor biopsy for sample pairs with concordant diploid or near diploid genome mass. The probit regressions are shown in **Response to Reviewers Figure 18c**.

Overall, 34 of 49 (69%) patients had at least one amplification or homozygous deletion reported in their matched tumor biopsies. Among these, 26 (76.5%) patients displayed at least one concordant. **Response to Reviewers Table 14** provides a detailed breakdown of detection rate by cancer type. In summary, 154 amplifications in 100 unique genes were detected in the 34 tumor biopsies and 106 (68.8%) of these were also detected in cfDNA. Only 23 of 51 (45.1%) homozygous deletions detected in the same tumor biopsies, mostly in prostate cancer patients, were also detected in cfDNA. This was expected given the sparse number of events as noted above and low sensitivity at high specificity observed for the breast and lung cohorts. **Response to Reviewers Figure 19** illustrates all the amplifications and homozygous deletions detected in cfDNA and the paired tumor tissue. Whilst in many cases, concordant alterations were found, CNVs present in the cfDNA but not in tumor biopsies were detected, consistent with the tumor CNV heterogeneity extensively documented in cancers of advanced stage (PMID:).

Finally, we evaluated the concordance between the tumor biopsies and cfDNA samples of clinically actionable alterations (OncoKB Level 1; PMID: 28890946) for all patients irrespective of the ctDNA fraction. Four breast cancer patients harbored an amplification of *ERBB2* on 17q and one lung cancer patient had an amplification of *MET* on 7q. The Log2 Ratio of the tumor biopsies and matched cfDNA samples are shown in **Response to Reviewers** **Figure 20**. All but two of the actionable CNV alterations could be detected. One of the four metastatic *HER2* amplified breast cancer cases did not have detectable *HER2* amplification in the cfDNA, and the *MET* amplified metastatic lung cancer did not have detectable *MET* amplification in the cfDNA. In these cases, the ctDNA fractions estimated were 1.3% and 1.9%, respectively, providing an explanation as to why these could not be detected in cfDNA using the high-intensity cfDNA assay.

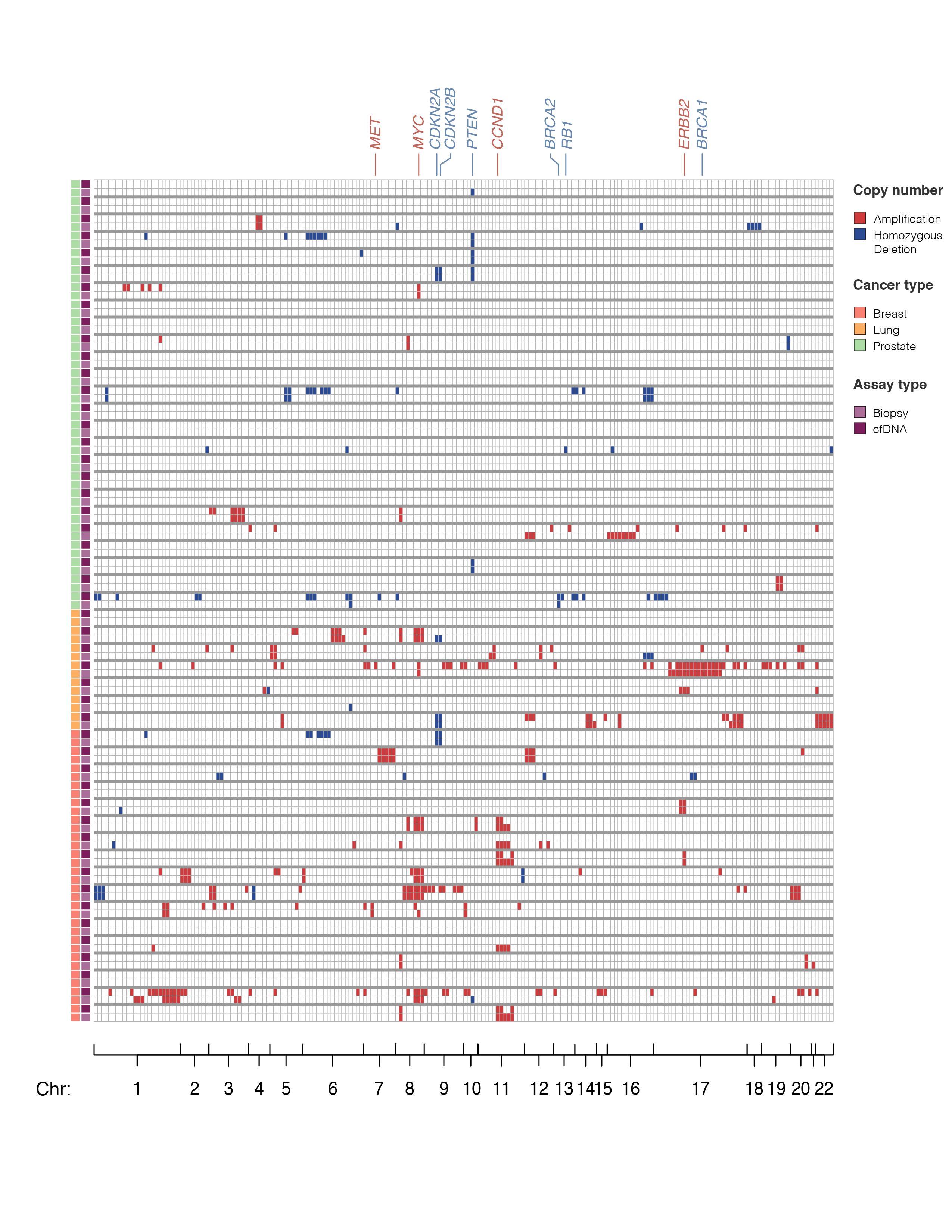
**Response to Reviewers Figure 16: Log2 Ratios derived from cfDNA of healthy control individuals.** Example Log2 Ratios estimated from the cfDNA of four healthy (a)-(d) female and (e)-(f) male control individuals. (g) and (h) show the density of segmented Log2 Ratios for the female and male controls, respectively. In (a) to (f), the Log2 ratios are displayed according to their genomic coordinates. The grey dots show the raw estimates while the red lines represent segmented values.

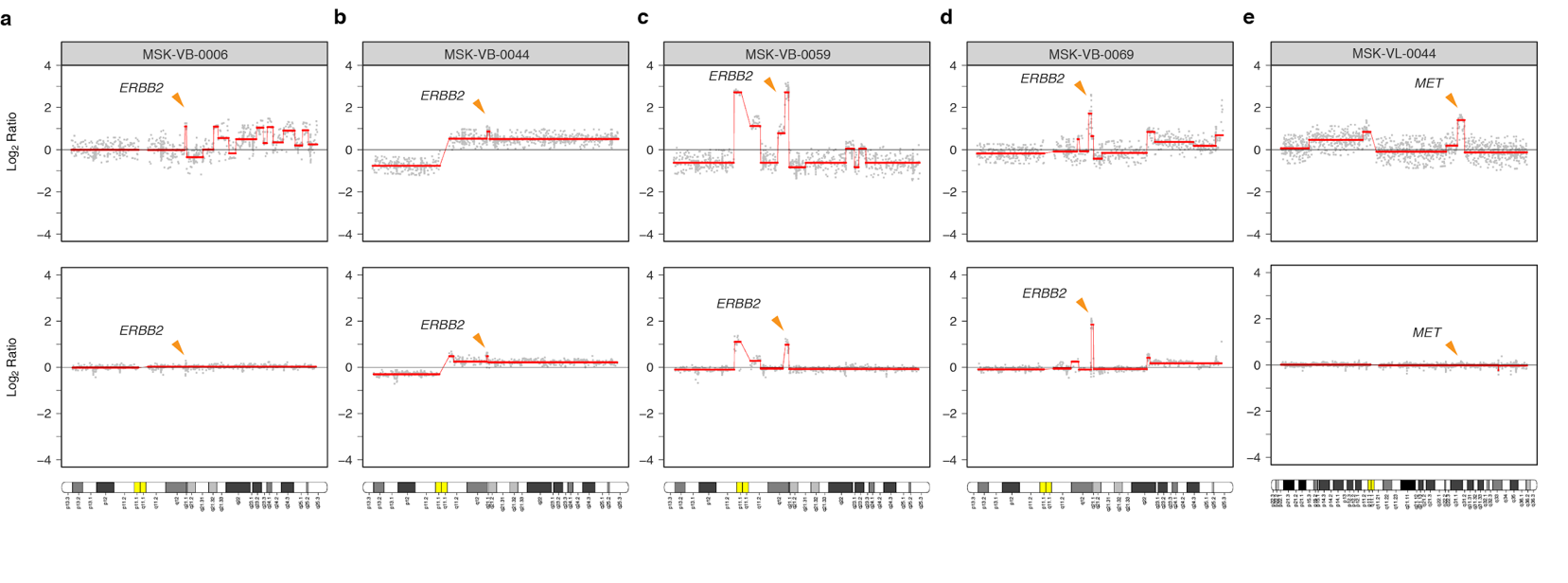
**Response to Reviewers Figure 17: Comparison of copy number alterations in tumor biopsy and matched cfDNA.** Log2 Ratios of tumor biopsies for three cases (a) MSK-VB-0008, (c) MSK-VL-0056 and (e) MSK-VP-0004 where amplifications of *CCND1, FGFR1*, *EGFR* and a homozygous deletion of *BRCA2* were reported. The corresponding Log2 Ratios of the matched cfDNA are shown in (b), (d) and (f), respectively where the arrows point to the reported amplifications or deletions. The segmented Log2 Ratios were used to compute the Pearson’s correlation coefficient comparing segments overlapping >75% in the tumor biopsies and cfDNA samples. The association of the Pearson’s *r* is shown in (g) against the ctDNA fraction. In (a) to (f), the Log2 Ratios are displayed according to their genomic coordinates. The grey dots show the raw estimates while the red lines represent the segmented values. In (g), the p-value was obtained using a permutation based one-sided Jonckheere-Terpstra test for increasing Pearson’s *r* with ctDNA fraction. NE; not evaluable.

**Response to Reviewers Figure 18: Performance assessment of cfDNA for detecting amplifications and homozygous deletions.** The ROC curves compare (a) copy number amplifications and (b) homozygous deletions detected in the tumor biopsy with the absolute copy numbers inferred in cfDNA. The probability of detecting a copy number amplification as a function of the absolute copy number in the tumor biopsy is shown by cancer type in panel (c). In (a) and (b), each tumor-cfDNA sample pair was used to construct individual curves. These were averaged after fitting a local polynomial regression and estimating the sensitivities over fixed intervals of specificities. In (c), only tumor-cfDNA sample pairs with concordant diploid or near diploid genome mass were used. In all panels, only sample pairs with ≥10% ctDNA fraction were used. NE; not evaluable.

**Response to Reviewers Table 14**: Summary statistics of concordance of amplifications and homozygous deletions between cfDNA and tumor biopsy assays

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Patients for whom ≥ 1 amplification or homozygous deletion detected in tumor was also detected in plasma cfDNA | | Detection rate of all aberrations | |
| No. of patients cfDNA / tissue | Detection rate (%) | Amplifications (%) | Homozygous deletions (%) |
| All | 26 / 34 | 76.5% | 106 / 154 (68.8) | 23 / 51 (45.1) |
| Breast | 11 / 14 | 78.6% | 57 / 84 (67.9) | 7 / 16 (43.8) |
| Lung | 4 / 6 | 66.7% | 38 / 48 (79.2) | 2 / 9 (22.2) |
| Prostate | 11 / 14 | 78.6% | 11 / 22 (50.0) | 14 / 26 (53.8) |

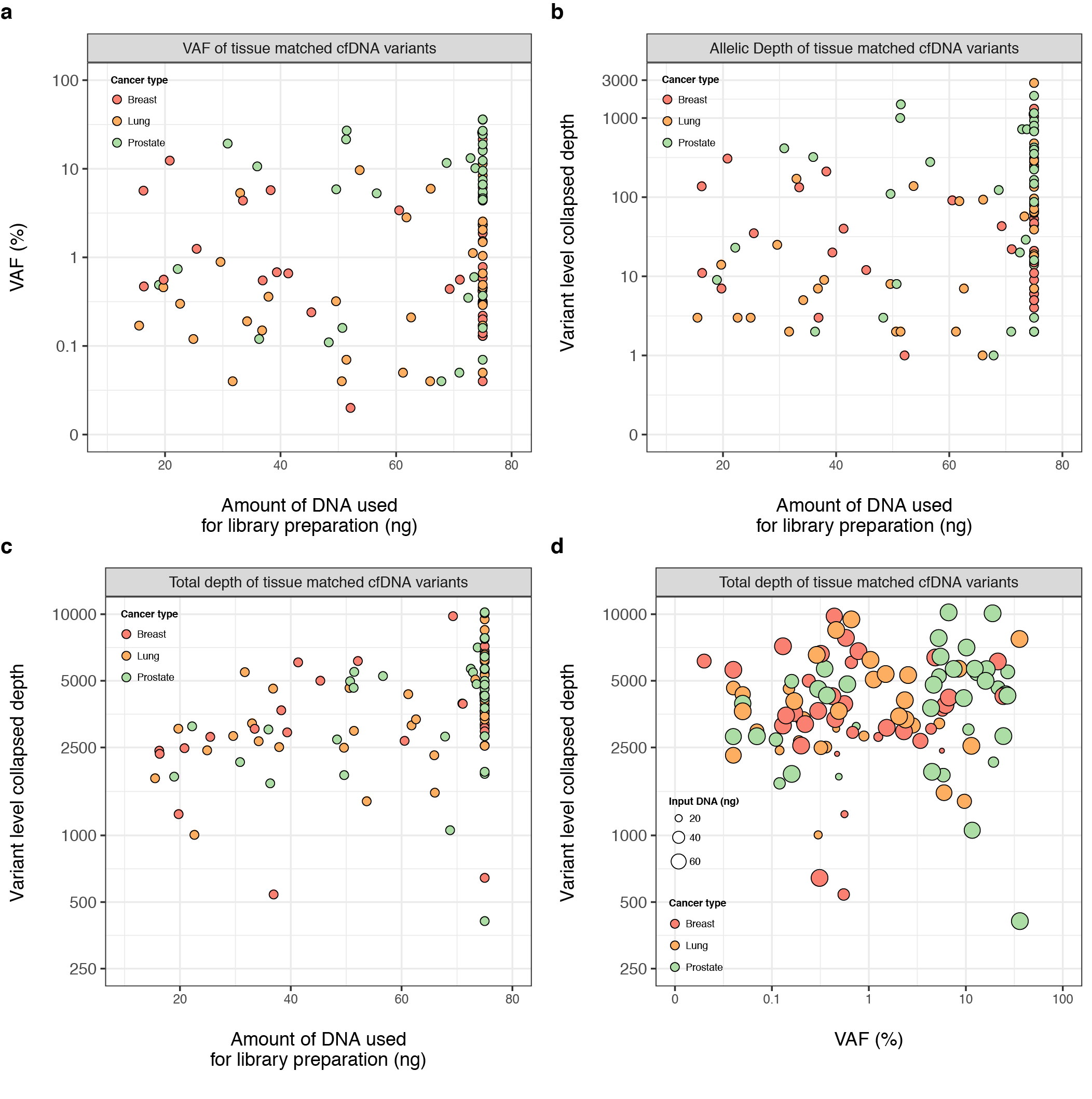
**Response to Reviewers Figure 19: Comparison of copy number alterations in tumor biopsy and matched cfDNA.** Heatmap of all genes where an amplification or a homozygous deletion was found in either the tumor biopsy or cfDNA. The alterations are color coded and indicated in the accompanying legend. The samples are interleaved i.e. tumor biopsy and cfDNA and represented along the rows whilst genes are ordered in columns relative to their genomic coordinates. Chromosome numbers are indicated along the bottom axis.

**Response to Reviewers Figure 20: Comparison of copy number alterations in tumor biopsy and matched cfDNA.** Four breast cancer patients (a) MSK-VB-0006, (b) MSK-VB-0044, (c) MSK-VB-0059 and (d) MSK-VB-0069 with a reported amplification of *ERBB2* on chromosome 17q are shown together with one lung cancer patient (e) MSK-VL-0044 with a reported *MET* amplification on chromosome 7q. In panels (a) to (e), the tumor biopsies are displayed on top and the cfDNA is shown below together with a chromosome ideogram. The genomic coordinates of *ERBB2* and *MET* are displayed by orange arrows and labelled accordingly.

Taken together, our exploratory, hypothesis generating analysis of CNVs utilizing the high-intensity cfDNA assay demonstrated that for samples with ≥10% ctDNA fraction, amplifications and homozygous deletions can be detected with relatively good accuracy. The assay, however, in the way it was originally designed, is not optimal for the detection of complex structural rearrangements and gene fusions. We have now included a brief passage in the manuscript, on page XX, section “High-intensity cfDNA assay analysis of copy number variants” of the revised manuscript, and included **Supplementary Figures XX** and **XX**, illustrating the above findings.

10. Library preparation involves an amount of 75 ng … which were the minimal requirements for performing library preparation? Which was the LODs for known-tissue mutations in cases in the low range of initial cfDNA?

Authors: A maximum input of 75 ng cfDNA with no lower limit was deemed acceptable for library preparation. For the WBC assay, a fixed 50 ng of sheared size selected genomic DNA extracted from buffy coat was used. **Response to Reviewers Figure 7** shows that the input DNA for library preparation explains the variability of cfDNA mean target coverage from collapsed reads. **Figure 2** of the revised manuscript shows the distribution of VAF in cfDNA of all Biopsy-matched mutations while **Response to Reviewers Figure 21** displays the input DNA used for library preparation against the VAF, allelic, and total depth of the lowest VAF Biopsy-matched mutation observed in cfDNA for every case with ≥1 concordant mutation.

**Response to Reviewers Figure 21: Limits of detection in cfDNA of tumor-matched variants.** For each patient, tumor-matched variants in cfDNA occurring at lowest (a) VAF, (b) allelic depth and (c) total depth are shown against the amount of input DNA used for library preparation. Panel (d) further shows the variant level collapsed depth of the tumor-matched variant at lowest depth against the corresponding VAF where the size of the individual dots is proportional to the amount of input DNA for library preparation. In all panels, the variants are color coded according to cancer type and indicated in the corresponding legends.

11. It would be advisable that the authors present the commands used and packages developed for the identification of the variants, as this will allow to obtain reproducible and comparable results with other series

Authors: As recommended by the Reviewer, we have now packaged all the computer codes on GitHub. These can be accessed at <https://github.com/ndbrown6/MSK-GRAIL-TECHVAL>

12. PPM1D and TP53 gene mutations were associated with chemotherapy and/or radiotherapy. These results confirm previous observations. However it has also been described that mutations in these genes in these settings associate with the development of therapy related myeloid neoplasms. It would be very interesting if the authors could provide the experience in this regard in their series.

Authors: The Reviewer raises a very important aspect of paired cfDNA-/WBC analysis as frequent detection of CH in routine cancer care may have important clinical significance. This analysis is not possible utilizing our current cohort as all the patients included in our study had advanced progressive metastatic disease at the time of enrollment, and unfortunately, many of them have succumbed to their disease before developing myeloid neoplasms. In fact, of the metastatic breast, lung and prostate cancer patients included in this study, 23, 27 and 27 have expired since the enrollment. Additionally, our sample size and the relatively short follow-up period make our study underpowered to further assess this association. We believe that cfDNA studies of patients with early-stage disease who receive adjuvant or neoadjuvant chemotherapy and/or radiation therapy as part of their definitive therapy would provide a better clinical setting to further elaborate on the excellent suggestion the Reviewer has offered. We have already embarked on such a study and are hopeful to be able to present our results within the next two to three years.

**Reviewer #3:**

Remarks to the Author:

Summary: Razavi and colleagues describe a study of 124 patients with metastatic cancer that had plasma cell-free DNA and WBC DNA sequenced to very high depth (60,000x) using a ~2Mb capture panel at GRAIL, with matched tumor biopsy sequencing using the MSK-IMPACT panel at standard depth (900x). In addition, they evaluate plasma and matched WBC in 47 healthy controls. They find that most of the somatic mutations in cfDNA are due to clonal hematopoiesis (CH), with important implications for clinical cfDNA testing in cancer. The paper is well-written and the findings are important and compelling. The manuscript might be improved by addressing the following comments.

Major Comments:

1. Authors might give more detail and discussion on other sources of variants of unknown source (“VUSo”) beyond tumor heterogeneity/biopsy sampling issues. This is an important area of exploration and in scope for the stated purpose of defining the sources of cell-free DNA. Particularly of interest would be a discussion of the VUSo in the healthy control subjects as that opens a window into possible sources of somatic variation beyond CHIP (CH) and known cancers. This seems to be at the heart of GRAIL needs to understand to do early detection with cfDNA. Benign neoplasms? Occult cancer(s)? Other cell types beyond hematopoietic cells with evidence of age-related clonal somatic mosaicism? The specific 67 VUSo in the healthy controls should be provided as a supplement with as much detail as possible. Reviewer’s apologies if this was provided and missed it. I looked in Supplemental Table 7 and had trouble figuring out which cases were these, if any.

Authors: We thank the Reviewer for highlighting this important aspect of our study. We fully agree with the Reviewer that other sources of somatic mutations such as benign tumors (e.g. nevi, polyps, etc.), somatic mosaicism beyond CH and occult malignancies could explain a portion of the VUSo. We have now more clearly discussed these possibilities in the revised version of the manuscript text (page XX, paragraph XX). We believe that definitive clinical validation of these other possible sources of VUSo is extremely challenging clinically and would be outside the of scope of the current study. Perhaps, studies involving longitudinal monitoring of cfDNA in hitherto healthy individuals who would later be diagnosed with cancer combined with autopsies of cancer patients and healthy individuals could help determine the full genomic landscape of the tumors and other possible sources of somatic mutations.

The full list of VUSo along with all other cfDNA variants are provided in **Supplementary Table 7** of the manuscript. Following the Reviewer’s excellent recommendation, we updated **Supplementary Figure S8** to separate healthy controls from the cancer patients to highlight the genes that are most frequently altered as VUSo in healthy controls (**Supplementary Figure S8** of the manuscript and **Response to Reviewers Figure 21** below).

2. There is some good discussion of the problem of CHIP (CH) interfering clones in cfDNA assays used for cancer care, but this could be emphasized more, including in the abstract. This is a major problem with the current analysis approach and interpretation of these assays. Harm is very likely happening to cancer patients as a result of hematopoietic clones being often misinterpreted as being cancer-derived in clinical cfDNA testing. Do authors think standard of care that all cfDNA assays used for cancer require matched WBC sequencing, for example? Could one or two specific case study examples be given that highlight this issue? The clinical authors are likely to have knowledge of such cases.

Authors: We thank the Reviewer for appreciating this important aspect of our study as nearly all commercially available cfDNA assays and many of those used in academia do not incorporate paired WBC sequencing and could erroneously report cfDNA variants originating from CH as tumor derived. As we highlighted in our response to Reviewer 1’s Comment #2, many of such mutations are pathogenic (e.g. *NF1, ATM, PIK3CA, TP53*) and may adversely affect clinical decision making. Additionally, with the increasing use of cfDNA assays for the detection of minimal residual disease (MRD) to direct clinical care, any false positive results regardless of their actionability can potentially result in misclassification of the MRD status of patients. As per the Reviewer’s suggestion, we have modified the discussion and highlighted this issue raised by the Reviewer (page XX, paragraph XX).

3. Cost of the ultra-deep sequencing approach should be discussed as a potential limitation of the broad applicability of this method. Sequencing a ~2Mb capture to >60,000x depth is likely to be cost-prohibitive for most labs.

Authors: We agree with the Reviewer and acknowledge that the high-intensity cfDNA sequencing assay reported here may be cost-prohibitive and the financial burden of such comprehensive assay could potentially be a major hurdle for its adoption in routine clinical practice. Given the rapid advances in sequencing technology, however, as of the writing of this manuscript, the initial costs of the raw sequencing at the time of sequencing are outdated. Taken together with the overhead of storing, processing and analyzing the sequence data, the global costs are difficult to quantify. It is possible to evaluate approximately what it would cost by today’s standards based on the sample raw mean sequence depth which required four libraries per HiSeq X Ten flow cell or 0.5 flow cell per sample for cfDNA and WBC sequencing. The prototype assay in this study was, however, designed to guide the development of future assays using ultra-deep sequencing with UMI barcoding in order to optimize sequence depth and narrow down informative targets. Following the Reviewer’s suggestion, we have revised the manuscript to highlight the cost as a potential deterrent to a broader adoption of this technology (page XX, paragraph XX).

4. Since this is presented as a technical report focused on assay performance, between-run and within run reproducibility of the assay might be discussed. This is especially important given the very low limit of detection.

Authors: We kindly refer the Reviewer to the response to Reviewer 1’s Comment #10, which addresses a similar concern, for a more elaborate discussion. In summary, the manuscript included technical replicates carried out to test reproducibility using two different versions (V1 and V2) of the assay for cfDNA and WBC. Samples from six patients, detailed in **Supplementary Table S6** of the manuscript, were selected for processing with both assay protocols. The results are shown in **Figure 1** and **Supplementary Figure S3** of the manuscript. The measured VAFs between the two technical replicates for samples from the five non-hypermutated patients showed a strong agreement (*R2* = 0.9997, **Figure 1C**) as well as for one sample from a hypermutated case (*R2* = 0.9972, **Supplementary Figure S3**). Additionally, three of the above patients have now been retested using version V2 of the protocol. The pairwise comparison of VAFs between versions V1 vs V2 and V2 vs V2 for all the samples that have been retested are shown in **Response to Reviewers Figure 8**. In addition, we have now also performed ddPCR of low VAF VUSo, which confirmed both the technical sensitivity and specificity of the high-intensity cfDNA sequencing assay as well as the accuracy for measuring VAFs utilizing ddPCR as the ‘gold standard’.

Minor Comments:

1. At first I was confused and thought MSK-IMPACT had been used for both plasma and tumor; run at ~60,000x with UMI and error correction added on. Instead, a different assay from GRAIL was used for plasma at ultra high-depth with UMI and error correction, and MSK-IMPACT for tumor at the regular depth (~900x). This could be made clearer to the reader early on the abstract/introduction. Fig 1a schematic is good showing this, but could be better outlined in text too.

Authors: We apologize for the lack of clarity in the original version of our manuscript. Tumor biopsies and matched normal tissue were sequenced using the MSK-IMPACT assay whilst plasma cfDNA and genomic WBC DNA extracted from buffy coat were sequenced by GRAIL, Inc using an ultra-high depth assay with UMI barcoding to allow error correction and a bespoke bioinformatic pipeline. The two assays have approximately 1Mb of human genome in common. There is, at present, no equivalent of MSK-IMPACT using UMI barcoding in routine clinical use for plasma cfDNA sequencing although a research version exists (PMID: 30675060) which is still under development. Following the Reviewer’s suggestion, the introduction and caption of **Figure 1** have been updated in the revised version of the manuscript (page XX, paragraph XX).

2. Underlying mismatch repair gene mutation(s) identified in the MSI-high prostate cancer case described? If not, why not?

Authors: The case illustrated in Figure 3(e) of the manuscript is a 55-year old patient with castration- and enzalutamide-resistant prostate cancer. The tumor biopsy harbored 27 somatic mutations including 6 frameshifting indels, 10 of which were labelled Biopsy-subthreshold and only detected by parallel analysis of the matched cfDNA (**Supplementary Table S9**). None of these mutations were related to the DNA mismatch repair pathway (MMR). Previous analyses had shown that the mutational profile of these tumor-matched mutations was consistent with the MMR signature (PMID: 28481359). **Figure 3** of the original manuscript shows that sequencing of the matched cfDNA captured the MMR signature and MSI-high phenotype. Analysis of the germline DNA revealed a missense *MUTYH* Y176C variant which is considered pathogenic in ClinVar (PMID: 29165669) and associated with MUTYH-associated polyposis (MAP). It is possible that the germline *MUTYH* mutation resulted in the MMR signature or that, equally likely, epigenetic factors lead to MMR deficiency in this tumor. Although there is no definite evidence of MMR pathway lesion, the patient was enrolled on a clinical trial of an anti-PD-L1 immunotherapy regimen, has exhibited a marked and sustained response to treatment for 35 months after initiating the immunotherapy regimen.

3. Why were synonymous variants separated out of some of the analyses? Certainly many of these will be passenger mutations, but the same is true for non-synonymous mutations, and does pathogenicity matter here?

Authors: All somatic mutations reported as part of the clinically validated MSK-IMPACT pipeline are non-synonymous variants. All comparisons in the manuscript were, therefore, limited to somatic non-synonymous variants which were detected in overlapping regions interrogated by both MSK-IMPACT and GRAIL’s cfDNA assay. Although the submitted version of the manuscript does not involve an elaborate discussion of the pathogenicity of these mutations, for the purpose of this response, all non-synonymous variants were annotated using OncoKB (PMID: 28890946).

We hope to have satisfactorily addressed the Reviewers’ insightful comments and constructive criticisms.