nature portfolio

Peer Review File

Cell-free DNA from germline TP53 mutation carriers reflect cancer-like fragmentation patterns

Corresponding Author: Professor Trevor Pugh

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In the paper entitled "Cell-free DNA from germline TP53 mutation carriers reflect cancer-like fragmentation patterns", Wong et al. used shallow whole genome sequencing data from nearly 200 cases to build up a diagnostic tool for LFS patients. This manuscript can be divided into 3 essential parts. In the first part, the authors showed us the difference in fragmentation patterns in TP53 mutation carriers, including increased shorter fragment proportion, intra-nucleosome fragmentation, and increased A/T frequency. The authors then discussed the genomic context, functional classes, cancer-associated copy number alterations, and nucleosome positioning to study where the differences in fragmentation are from. In the last section, they successfully build a model with multiple fragmentomic features. In general, they generated pretty rich data, and the logic is convincing. However, some concerns need to be carefully addressed before accepting.

Q1: It's very rigorous to make sure other variables such as patient age, cancer history, time of sample collection, sample collecting institutions, collection tubes, sequencing run, and cluster density would not affect the conclusion. I do notice that patients might have different cancer types in your supplementary table 1. Would this affect the result?

Q2: In Figure 2A&B, we can also observe an obviously increased proportion of relatively long fragments (250-500 bp). Fragments in this size range can be regarded as multi-nucleosomal related. Do you have any explanation for this increase?

Q3: In Figure 4B, what does HBC stand for? TP53-wildtype? this shall be denoted at least in the figure legend. Does the tetranucleotide you noted as the X-axis stand for all 256 combinations of 4-mer end motifs? Does this figure suggest a decreased motif frequency in TP53m-carriers versus wildtype comes along with the higher GC content in the 4-mer motif?

Q4: In lines 117-115, we know genomic instability associated with telomere attrition has always been linked with LFS let'ss call it A), and we know short fragments are associated with ctDNA let'ss call it B). It's still weird to try to combine those two factors and see their connections. Is this a well-known phenomenon that deserves a test, and is the result of no relationship between A&B meaningful?

Q5: In lines 127-145, similar to the last question, why would you group mutations in those categories? Take the splice site, for instance. Some of those sites might contribute positively to generating shorter fragments, while some of those sites might not, or even contribute negatively to generating shorter fragments. If so, this method of grouping would blur the real difference. Maybe you can try different angles of grouping those mutations. I understand that the conclusion is that broad TP53 mutation classes may not be useful for explaining the generation of short fragments, but due to the actual utility in future clinical use, I do think you can figure out a "not that broad classes" and take that as real cancer biomarkers.

Q6: In line 176, how do you calculate sliding frequency? The genomic context? Detailed definitions shall be included.

Q7: In line 177, why 167+-50bp and 167bp +-83bp in line 151? How do you pick the flanking size range?

Q8: In line 198, what do you mean by "disproportionate increase/ decrease"?

Reviewer #2

(Remarks to the Author)
In this paper, Wong et al. investigate cell-free DNA from patients with Li-Fraumeni syndrome. Circulating cell-free DNA in the

plasma of cancer patients contains cell-free tumour DNA (cfDNA) derived from tumour cells. Sequencing of cell-free DNA in the blood of cancer patients (liquid biopsy) provides attractive opportunities for early diagnosis. In this paper, Wong et al. comprehensively investigated the cfDNA of 82 Li-Fraumeni and 30 healthy TP53-wild-type patients using shallow whole genome sequencing. However, the paper is confusing, and lacks biological insights.

Major comments
1. Were the healthy and Li-Fraumeni patient plasma samples indeed collected in the exact same way? This will be important to note for any changes seen. We need to be sure this is not due to collection biases and is really due to real differences.
2. The observations should be put into context. What are the other cfDNA studies? What have they found?
3. There is very little biological insight into the data. The authors make a throwaway statement “Other factors such as germline genetics, epigenetic, or lifestyle may also contribute to the variability observed between patients” – but they do not investigate this at all. The authors say “we observed a decrease in A/T motifs and an increase in G/C motifs in LFS-AC samples, which may be due to differential expression and activity of nucleases in cancer cells (Supplemental Figure 9D)” – the authors make such statements like these, but do not go on to explore anything further about what nucleases might be the reason.
4. The “chromatin architecture” part is especially confusing. I cannot understand how genome-wide fragmentation profiles can give insights into chromatin architecture. I do not understand even what chromatin architecture is being referred to here – nucleosomes? Open chromatin? 3D genome organization? It would help if there are perhaps some citations of previous work done in this aspect.
5. The authors say “our observations suggest that TP53m-carriers exhibit differences in cfDNA fragmentation that may be driven by abnormal chromatin organization in specific tissues” but given they do not look at the original cancers that the cfDNA came from, I do not see how this association or correlation can be made
6. Same comment for the nucleosome positioning – I am also very confused as to how this work was done. How did the authors find their nucleosome positioning? The p53 ChIP-Seq in the database may not reflect the actual p53 binding sites in the cancer or cfDNA in these specific patients. Also, nucleosome is not the same as p53 transcription factor binding sites.
7. Same comment for the section on “increased chromatin accessibility at developmental and cancer-associated loci” – I do not understand how the authors “ inferred the nucleosome positioning at the top 10,000 binding sites (meta peaks) for 335 transcription factors (TFs)” from their cfDNA WGS data. The authors say “we explored nucleosome positioning at cancer-associated sites using ATAC-seq peaks from 23 TCGA cancer types” – these may not be relevant to the specific patient samples examined, as ATAC-Seq is known to be different in different cell types and different people. To do this analysis, it should be done in the same patient (i.e. matched cfDNA and patient cancer ATAC-Seq).
8. For the prediction of phenoconverters, this is very exciting results, but it would be best if an independent cohort could be obtained and the predictor be applied to test, to check if there the original predictor was over-fitted or not. If the predictor is good, then that means it will be able to predict with an independent cohort.

Minor comments
1. Define cfDNA in the abstract. Also, explain what is cell-free DNA more clearly in the introduction, with more citation of the relevant literature.
2. In the introduction, it is important to state they also looked at 30 healthy TP53-wild-type controls.

Reviewer #3

(Remarks to the Author)
In this manuscript, Wong and colleagues aim to assess if mutations in the tumor suppressor TP53, which are prevalent in cancer and when occurring in the germline cause the cancer-predisposing Li-Fraumeni Syndrome (LFS), is associated with specific patterns of circulating cell-free DNA (cfDNA) fragmentation, with potential diagnostic/prognostic value. To this aim, the authors got liquid biopsies from a cohort of cancer-free/past-cancer/active-cancer LFS patients and healthy wtTP53 individuals and performed cfDNA fragmentomics to detect nucleosome-free/degradation-prone and nucleosome-protected/degradation-resistant DNA fragments.
Data analysis showed that TP53 missense/splice/LOF LFS mutations were associated, independent of cancer status, with reduced cfDNA fragment size. Analyses of fragments' end and bp composition indicated increased cleavage of nucleosome-associated sequences, resulting in fragments with specific sequence patterns. Interestingly, mapping LFS fragment profiles with annotated TFBS identified a set of BS for developmental TFs whose accessibility correlated with cancer development. Furthermore, fragment profiling could also detect, in active-cancer LFS, increased accessibility in chromatin regions that were found more accessible also in advanced cancer from TCGA ATAC-seq datasets. The authors also developed ML algorithms validating the diagnostic/prognostic value of cfDNA profiles in LFS.
These results led to the identification of cfDNA markers of cancer-free LFS, and also of possible nucleosome alterations, associated with TP53 mutations, correlating with cancer development, with possible diagnostic/prognostic value in LFS and potentially also in somatic cancers driven by TP53 mutations. However, I have some concerns that should be addressed to improve the MS.

1) Data are very dense and not always clearly described. In some cases the use of different terms to describe the same readout is confusing. Overall, the text should be revised to improve readability.
As one example, it should be briefly introduced what the parameters “midpoint coverage”, “central coverage”, ”amplitude” “variability” and “meta peaks” represent. The explanation that decreased “central coverage = increased accessibility/binding”, “indicative of binding activity”, should be anticipated from line 309 to the section describing nucleosome positioning analyses (lines 277-281).
Also, whenever possible, biologically meaningful terms should be used instead of technical parameters (e.g. replacing

“decreased central/midpoint coverage” with “increased accessibility”).

2) To support the relevance of the identification of TFBS whose accessibility correlates with cancer development in LFS, the authors should assess whether the TFs targets are upregulated in mutant p53 cancers, by analyzing available LFS and TCGA RNAseq datasets.

3) In Fig.1 the authors included a schematic of the whole study. To improve readability of results, graphical abstract(s) of conclusions should also be included.

Minor concerns:

- In lines 288-298, please briefly describe the rationale for analyzing CTCF.

- In lines 281-284, the authors describe decreased midpoint coverage and increased amplitude at TSS of p53 targets but not housekeeping genes (Figure 6B, Supplemental Figure 12B). However, in lines 288-290, they state that TP53m-carriers displayed decreased central coverage at the TSS of housekeeping genes (Supplemental Figure 12A-B). Please clarify this apparent contradictory conclusion.

- Lines 295-295: a steep drop-off of fragments <90bp in healthy TP53-wildtype at open-associated chromatin sites is reported, but no statistically significant difference is highlighted in Fig.S12H. Please clarify. Also, it should be better described how the authors reach the conclusion that, in TP53m-carriers, a decreased fraction of total reads suggests increased degradation.

- Nucleosome positioning (midpoint and amplitude) was not found to be significantly different across different TP53 mutation classes (Supplemental Figure 12C-F). Please clarify what groups were compared in statistical analyses. It seems that some mutations behave differently (e.g. LOF impacts amplitude only in p53 TSS, Supplemental Figure 12E-F). Please discuss these differences.

- Some Figure panels are not referenced in the text (e.g. 3C-D). Please check them all.

- In some Figure panels, some annotation is lacking/cropped (e.g. asterisks indicating statistically significant differences in Fig. S12). Please review all figure panels.

- It would be helpful to indicate, in Fig.7A and S13A plots, the TFs described in lines 310-314.

Author Rebuttal letter:

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author): expertise in DNA fragmentomics bioinformatics

In the paper entitled "Cell-free DNA from germline TP53 mutation carriers reflect cancer-like fragmentation patterns", Wong et al. used shallow whole genome sequencing data from nearly 200 cases to build up a diagnostic tool for LFS patients. This manuscript can be divided into 3 essential parts. In the first part, the authors showed us the difference in fragmentation patterns in TP53 mutation carriers, including increased shorter fragment proportion, intra-nucleosome fragmentation, and increased A/T frequency. The authors then discussed the genomic context, functional classes, cancer-associated copy number alterations, and nucleosome positioning to study where the differences in fragmentation are from. In the last section, they successfully build a model with multiple fragmentomic features. In general, they generated pretty rich data, and the logic is convincing. However, some concerns need to be carefully addressed before accepting.

We thank this Reviewer for their comments and feedback. We have performed additional analyses to address this Reviewerâs concerns and to further validate our findings. Further analysis assessing the contribution of technical variables have been performed and included as additional figure panels. We hope that these analyses have addressed this Reviewerâs concerns.

Q1: It's very rigorous to make sure other variables such as patient age, cancer history, time of sample collection, sample collecting institutions, collection tubes, sequencing run, and cluster density would not affect the conclusion. I do notice that patients might have different cancer types in your supplementary table 1. Would this affect the result?
â We appreciate this important point that has been brought up by this Reviewer. We have investigated this in the context of fragment lengths and found there to be no influence of these factors on the fragmentation, and have added the following to the manuscript:
â Page 4: âWe have previously described an increased proportion of short (<150bp) cfDNA fragments in TP53m-carriers which was independent of cancer status (ANOVA p-value < 0.001; Figure 2A/B), cancer history (Studentâs t-test p-value = 0.085; Supplemental Figure 2A-D) and patient age at the time of sample collection (Pearsonâs R2 = 0.01; Supplemental Figure 2E)(13). Technical variables were also accounted for, with no

differences detected between institutions (Studentâs t-test = 0.637; SickKids = SK, Princess Margaret = PM), collection tubes (SK = EDTA, PM = Streck; Supplemental Figure 2F), sequencing run (ANOVA p-value = 0.0601; Supplemental Figure 2G), or flow cell cluster density (Pearsonâs R2 = 0.004 â 0.049; Supplemental Figure 2H).â
â We have additionally sequenced a subset of our samples using an alternate sequencing method (Ultima Genomics) and found that this platform recapitulates the fragment frequency distributions observed in Illumina. We have added these new data to the text as follows as well as to the Supplemental material:
â Page 4: âUsing an alternate sequencing method (Ultima Genomics), in a subset of samples (n = 58), we observed a similar distribution of fragment lengths in TP53-wildtype (n = 18) and TP53m-carriers (cancer negative n = 26, cancer positive n = 14) with the canonical peak at 167bp and 10bp periodicities (Supplemental Figure 2J-K). Comparing the frequency of fragments in 5bp bins, we saw high correlation between Illumina and Ultima sequencing suggesting good concordance between the two sequencing technologies (R2 = 0.973 â 0.988; Supplemental Figure 2L) (18).â
â To further address this, we have included a comparison of our LFS patients to external healthy control cohorts from published studies (1â3). Increased frequency of short fragments was consistent when comparing to all other healthy cohorts.
â Page 4: âThis increase in short fragments also persisted when compared to external cohorts of healthy controls (Supplemental Figure 2I)(15â17).â
â To similarly address this point, we have investigated the influence of each variable on the probability outputs for our TP53-wildtype vs TP53m-carrier models and observed similar results to our fragment length analysis:
â Page 11 â 12: âFurther analysis of the prediction scores generated by the integrated model (LFS fragmentation score [LFS-FS]) showed no differences between TP53m-carriers independent of cancer history or germline TP53 class (Supplemental Figure 15A-B). Although, we did observe higher LFS-FS in pediatric compared to adult TP53m-carriers ($p < 0.001$; Supplemental Figure 15C), LFS-FS were not correlated to age in both TP53-wildtype (R2 = 0.06) and TP53m-carriers (R2 = 0.12; Supplemental Figure 15D). To account for technical variation, we compared LFS-FS across different sequencing runs (Supplemental Figure 15E) and did not observe correlation with cluster density (R2 = 0.006; Supplemental Figure 15E).â
â We did not perform this analysis for CFD and CFS scores as these scores are normalized against a patient specific baseline. Longitudinal samples were sequenced across multiple sequencing runs thus patient-specific normalization would account for technical differences across sequencing runs.

Q2: In Figure 2A&B, we can also observe an obviously increased proportion of relatively long fragments (250-500 bp). Fragments in this size range can be regarded as multi-nucleosomal related. Do you have any explanation for this increase?
â In Figure 2A/B we do also observe an increase in di-nucleosome fragments. The release of di-nucleosome fragments is currently less well understood compared to mono-nucleosome fragments. Studies have shown that the mechanism is not merely a result of undigested mono-nucleosome fragments and that it may be due to chromatin architecture as di-nucleosome fragments are consistently released by normal cells into the circulation (4,5). We suspect that the increase in di-nucleosome fragments may also be related to the differences in the underlying biology of LFS. However, in this manuscript we focused on mono-nucleosome fragments as the mechanism of their release is much better understood.
â To address this comment, we have also included additional figures to highlight that di-nucleosome fragments are also increased in LFS individuals and added the following text to the manuscript.
â Page 4: âWe also observed an increase in longer fragments (250-500bp), representing di-nucleosome fragments, in TP53m-carriers compared to TP53-wildtype (Supplemental Figure 2I-M).â
Q3: In Figure 4B, what does HBC stand for? TP53-wildtype? this shall be denoted at least in the figure legend. Does the tetranucleotide you noted as the X-axis stand for all 256 combinations of 4-mer end motifs? Does this figure suggest a decreased motif frequency in TP53m-carriers versus wildtype comes along with the higher GC content in the 4-mer motif?
â HBC stands for TP53-wildtype. This has been made consistent in all figures and figure legends.
â The x-axis denotes all 256 combinations of 4-mer end-motifs. This has also been clarified in the figure and figure legends.
â We observed a decrease in the frequency of GC rich end-motifs in LFS individuals compared to controls which is highlighted in the Results text:
â Page 7: âTP53m-carriers (cancer negative) were characterized by an increase in A/T rich motifs and a decrease in G/C rich motifs relative to TP53-wildtype (Figure 4B).â

Q4: In lines 117-115, we know genomic instability associated with telomere attrition has always been linked with LFS let's call it A), and we know short fragments are associated with ctDNA let's call it B). It's still weird to try to combine those two factors and see their connections. Is this a well-known

phenomenon that deserves a test, and is the result of no relationship between A&B meaningful?
â The rationale for this investigation is that increased genomic instability may result in increased DNA strand breakage during the generation of cell-free DNA resulting in decreased overall fragment size as observed in circulating tumor DNA (ctDNA). While the mechanism of shorter ctDNA is currently not fully understood, we felt that it was important to report negative data as both telomere dysregulation and shorter ctDNA have been previously linked to LFS and cancer.

Q5: In lines 127-145, similar to the last question, why would you group mutations in those categories? Take the splice site, for instance. Some of those sites might contribute positively to generating shorter fragments, while some of those sites might not, or even contribute negatively to generating shorter fragments. If so, this method of grouping would blur the real difference. Maybe you can try different angles of grouping those mutations. I understand that the conclusion is that broad TP53 mutation classes may not be useful for explaining the generation of short fragments, but due to the actual utility in future clinical use, I do think you can figure out a "not that broad classes" and take that as real cancer biomarkers.
â The TP53 variants were clustered based on their functional outcomes. This analysis was performed in collaboration with David Malkinâs laboratory and revisions of a separate manuscript is currently under review describing these observations. The data is also based on prior studies by Kato et al and Giacomelli et al who also assessed the functional outcomes of different TP53 variants (6,7). Given a large enough cohort, variants would be assessed independently as different variants within each mutation cluster are known to affect TP53 protein function to variable levels (hypomorphic vs loss of function missense variants).
â Specifically, splice site variants were grouped separately due to the uncertainty of the exact downstream effects of splice site variants on mRNA splicing and protein products of the TP53 gene. While a variant at a particular location may affect the cleavage of that locus by nucleases during cfDNA generation, this phenomenon is not unique to splice site variants as all variants and even SNPs in the genome would also affect this. The presence of a variant at a particular locus would also not affect the size of cfDNA globally. Splicing variants are not expected to affect the generation of cfDNA as splicing is a mechanism related to RNA synthesis. The downstream protein product of each TP53 splice site variant may lead to different changes in chromatin architecture based on the functional outcome of the splice variant. However, this has not been as extensively studied compared to missense TP53 variants and thus influenced our decision to group splice site variants together.

Q6: In line 176, how do you calculate sliding frequency? The genomic context? Detailed definitions shall be included.
â Dinucleotide frequencies were calculated using the position assessed plus the 5â preceding nucleotide. This was performed for every position across the span of the fragment including the 50bp flanking the fragment. We have clarified this in the Results and included additional information in the Methods. This was performed similar to methods previously described by Snyder et al (8).
â Results: - Page 6: âwe calculated the frequency of A/T and G/C dinucleotides across nucleosome spanning (167bp) cfDNA fragments (Figure 3B); 50bps flanking the fragment were also included for normalization.â
â Methods - Page 21 - 22âThe frequency of each dinucleotide were calculated along the region (position + 5â preceding nucleotide) and frequencies were normalized against the expected frequency of each dinucleotide based upon their prevalence across the genome.â

Q7: In line 177, why 167+-50bp and 167bp +-83bp in line 151? How do you pick the flanking size range?
â For our dinucleotide frequency analysis, we chose +/-50 bp flanking the nucleosome spanning fragment to ensure enough base-pairs were available for normalization and comparison with the nucleosome spanning region. This was also done by Snyder et al (8).
â For our nucleosome footprint analysis, 167+/-83 bp was not used. Instead, +/-83bp is used to reference the number of nucleotides that flank the center of a nucleosome spanning read which is 167bp (83bp on either side of the central nucleotide). We have clarified in the manuscript now as follows:
â Page 5 â 6: âIn TP53-wildtype, using a previously defined genome-wide set of ~13 million peripheral blood cell-derived nucleosomes (10), we observed an M-shaped distribution of fragment-ends, with peaks +/-83 bp from the center of the nucleosome, corresponding to a nucleosome spanning cfDNA fragment (~167bp) (28).â

Q8: In line 198, what do you mean by "disproportionate increase/ decrease"?
â We have removed the words âdisproportionate increase/decreaseâ and have instead quantified the change which is reflected in the Results section:
â Page 7: âNext, we split end-motifs into those that occur at frequencies above and below the expected random frequency (1/256 = 0.0039) in TP53-wildtype. Within TP53m-carriers, we observed an increased frequency in 63.0% (63/100) and a decreased frequency in 67.3% (105/156) of motifs that occur above and below the expected frequency (Chi-squared p-value < 0.001) compared to TP53-wildtype (Supplemental

Figure 9A). This was further accompanied by an increase in 82.5% (33/40) of A/T rich end-motifs (3+ A/T) that occur above (Chi-squared p-value = 0.009) and a decrease in 84.4% (49/58) of G/C rich end-motifs (3+ G/C) that occur below (Chi-squared p-value < 0.001) the expected frequency (Supplemental Figure 9B).â

Reviewer #2 (Remarks to the Author): expertise in epigenetics and chromatin accessibility

In this paper, Wong et al. investigate cell-free DNA from patients with Li-Fraumeni syndrome. Circulating cell-free DNA in the plasma of cancer patients contains cell-free tumour DNA (cfDNA) derived from tumour cells. Sequencing of cell-free DNA in the blood of cancer patients (liquid biopsy) provides attractive opportunities for early diagnosis. In this paper, Wong et al. comprehensively investigated the cfDNA of 82 Li-Fraumeni and 30 healthy TP53-wild-type patients using shallow whole genome sequencing. However, the paper is confusing, and lacks biological insights.

We thank this Reviewer for their comments. They have brought up important points that should be studied more extensively but are not within the scope of this study. Cell-free DNA fragmentomics is a relatively new field of study and there remains many avenues of investigation into the mechanisms involved in cfDNA release and how cell-type specific differences may influence this. Figure 1 was included as a graphical abstract to introduce each of the fragmentomic concepts. Relevant literature has been cited in each section should the reader require additional information to understand the underlying concepts. We have made changes to the manuscript to increase the clarity and highlight previous literature in the field that is pertinent to understanding our study.

Major comments

1. Were the healthy and Li-Fraumeni patient plasma samples indeed collected in the exact same way? This will be important to note for any changes seen. We need to be sure this is not due to collection biases and is really due to real differences.

â The healthy control subjects were collected in Streck tubes using the same protocol as all samples collected at the Princess Margaret Cancer Center. We have investigated the effects of collection tube and additional variables (see Reviewer #1 comments) and found that they were not associated with the levels of short cfDNA fragmentation in our cohort. Our healthy controls were also consistent with healthy controls from previously published studies (see Reviewer #1 comments).

2. The observations should be put into context. What are the other cfDNA studies? What have they found?

â Studies investigating cell-free DNA in individuals with hereditary cancer syndromes are lacking due to the relative rarity of these syndromes in comparison to sporadic cancer. Our group has previously published on mutation and methylation analysis of cell-free DNA in LFS for early detection (9). This is the first comprehensive study investigating differences in cell-free fragmentation in individuals with hereditary cancer syndromes. However, it is well established that other disease states such as cancer and auto-immune disease can alter cell-free DNA fragmentation. These studies have been cited throughout the manuscript.

3. There is very little biological insight into the data. The authors make a throwaway statement âOther factors such as germline genetics, epigenetic, or lifestyle may also contribute to the variability observed between patientsâ â but they do not investigate this at all. The authors say âwe observed a decrease in A/T motifs and an increase in G/C motifs in LFS-AC samples, which may be due to differential expression and activity of nucleases in cancer cells (Supplemental Figure 9D)â â the authors make such statements like these, but do not go on to explore anything further about what nucleases might be the reason.

â Fragmentomics is a new and developing field. Thus, there are still many questions that require further investigation that are not fully within the scope of this study. We agree that this analysis has raised several new exciting avenues for biological investigation, several of which we are pursuing in the context of LFS model organisms as well as larger cohorts within and outside of LFS. To encourage others to build upon these leads, we constrained the scope of this work on reporting the most salient features to differentiate LFS patients from non-LFS controls, with potential diagnostic implications.

â It is known that the epigenetic landscape of the cell-of-origin influences the genomic origins and lengths of cfDNA fragments that are released into circulation (8,10â12). However, there have not been studies investigating how variation in the epigenetic landscape within cell populations and individuals affects the release of cfDNA and the global cfDNA landscape. Our study shows that germline events can influence the cfDNA landscape.

â Germline genetics have been shown to affect the cfDNA landscape (13,14). In our study, we also show that variants in TP53 affect the cfDNA landscape. However, these studies are within a monogenic context, and it is currently unknown how variants in other genes and/or the combinatorial effects of germline variation may also affect the cfDNA landscape. This is an important question to ask as liquid biopsy increases in prevalence within clinical spaces. However, these analyses are not within the scope of this study.

â Studies have also shown that lifestyle events such as obesity and exercise can also affect the amounts of cfDNA in circulation and the cell-type make-up of the cfDNA landscape

(15,16)(17,18). Given this information, we postulate that some degree of the differences observed within our LFS cohort may be attributed to these lifestyle differences that may be present. However, this information is not included within our clinical abstraction and thus cannot be tested rigorously. It is also not within the scope of this study.
â These observations are supported by the fact that within healthy individuals, there is slight variation in the cfDNA landscape between individuals and even, temporally within the same individual. One of the strengths of cfDNA analysis is that it provides a snap shot of the cellular physiology within an individual in real time as the half-life of cfDNA within circulation is predicted to be only 1-2 hours (19).
â We have added additional text to clarify:
â Page 5: âOther factors such as germline genetics, epigenetic, or lifestyle may also contribute to the variability observed between patients, similar to in TP53-wildtype individuals (27).â
â Page 7: âLastly, comparing LFS-H to LFS-AC (fold change over healthy non-LFS controls), we observed a decrease in A/T motifs and an increase in G/C motifs in LFS-AC samples, which may be due to the differential expression and activity of nucleases previously described in cancer cells (Supplemental Figure 9D)(30).â

4. The âchromatin architectureâ part is especially confusing. I cannot understand how genome-wide fragmentation profiles can give insights into chromatin architecture. I do not understand even what chromatin architecture is being referred to here â nucleosomes? Open chromatin? 3D genome organization? It would help if there are perhaps some citations of previous work done in this aspect.
â We have provided Figure 1 as a graphical schematic on the concepts of fragmentomics and how cfDNA fragmentation can be used to infer the epigenetic landscape (open vs closed chromatin, nucleosome bound vs naked DNA, transcription factor binding activity). These concepts have been well established by studies that have been cited throughout the text.
â We have used the term âchromatin architectureâ to denote the 3-dimensional chromatin conformation within a cell that influences the release of cfDNA into circulation. These signatures are often tightly tied to the epigenetic signatures associated with regulating cell-identity. We have provided additional text to clarify this in the Results:
â Page 8: âChromatin is 3-dimensionally arranged in relation to the genome as a mechanism of regulating gene expression and cell identity. This 3-dimensional chromatin architecture influences the representation of cfDNA fragments (size and location) in circulation due to the differential accessibility of nucleases throughout the genome (10).â

5. The authors say âour observations suggest that TP53m-carriers exhibit differences in cfDNA fragmentation that may be driven by abnormal chromatin organization in specific tissuesâ but given they do not look at the original cancers that the cfDNA came from, I do not see how this association or correlation can be made.
â These observations are independent of cancer formation. It is known that individuals with LFS exhibit baseline differences in transcriptional and DNA methylation signatures compared to non-LFS individuals (20,21). These studies have only studied specific cell-types, however, it is presumed that these differences may be ubiquitous, though cell-type specific, throughout the body as LFS individuals harbor constitutional TP53 variants. Further studies should be considered to investigate systemic and cell-type specific differences in LFS vs TP53-wildtype individuals. However, those studies are out of scope for this study.

6. Same comment for the nucleosome positioning â I am also very confused as to how this work was done. How did the authors find their nucleosome positioning? The p53 ChIP-Seq in the database may not reflect the actual p53 binding sites in the cancer or cfDNA in these specific patients. Also, nucleosome is not the same as p53 transcription factor binding sites.
â These concepts are highlighted in Figure 1 as well as the relevant literature that has been cited. Nucleosome-bound DNA is protected from degradation which results in increased read representation in regions of consistent nucleosome occupation. This is especially pronounced surrounding transcription factor binding sites (22â24). Nucleosome positioning was not used interchangeably with TFBS as those are distinctly different concepts. However, they are linked in the context of cfDNA read representation and in silico prediction of transcription factor binding activity. We have clarified this in the text:
â Page 9: âUsing Griffin, we inferred the nucleosome positioning surrounding p53 binding sites using p53 CHIP-seq data from the Gene Transcription Regulation Database (GTRD) (34)(35).â
â While true that the p53 binding sites may not fully represent all of the binding sites in the cancer or cfDNA in the patients in our study, it is not within the scope of this study or budget to have performed CHIP-seq on tumor and buffy coat samples from patients in our study. Instead, we used the GTRD, who have curated published CHIP-seq datasets from multiple tissue/cell line sources to create a list of meta-peaks for p53 which is reflective of general p53 binding (25â27).
7. Same comment for the section on âincreased chromatin accessibility at developmental and cancer-associated lociâ â I do not understand how the authors âinferred the nucleosome positioning at the top 10,000 binding sites (meta peaks) for 335 transcription factors (TFs)â from their cfDNA WGS data. The

authors say âwe explored nucleosome positioning at cancer-associated sites using ATAC-seq peaks from 23 TCGA cancer typesâ â these may not be relevant to the specific patient samples examined, as ATAC-Seq is known to be different in different cell types and different people. To do this analysis, it should be done in the same patient (i.e. matched cfDNA and patient cancer ATAC-Seq).
â See comments above regarding p53 and nucleosome positioning.

8. For the prediction of phenoconverters, this is very exciting results, but it would be best if an independent cohort could be obtained and the predictor be applied to test, to check if there the original predictor was over-fitted or not. If the predictor is good, then that means it will be able to predict with an independent cohort.
â Li-Fraumeni syndrome is a rare syndrome. Our cohort represents the largest cohort of patients with comprehensive cfDNA profiling to date in the literature. This cohort has also exhausted the patient cohort present at both of our institutions (adult and pediatric). In order to accrue a validation cohort, this would require additional years of sample collection and collaboration with additional institutions. This study is also a proof-of principle study that will require additional follow-up studies, both in larger cohorts (we have started a prospective study to acquire more samples) and in model organisms (led by Dr. Malkinâs lab).
â To reduce this limitation, we split our cohort into a held-out test and cross-validation cohorts and performed iterative downsampling and 10-fold cross-validation to train our model. This is stated in our Results, Methods, and in Supplemental Figure 15C.

Minor comments
1. Define cfDNA in the abstract. Also, explain what is cell-free DNA more clearly in the introduction, with more citation of the relevant literature.
â We have included additional text in the abstract to clarify what cell-free DNA is:
â Page 2: âOur group has previously shown that LFS patients harbor shorter cell-free DNA, fragments of DNA released by cells into bodily fluids, fragmentation; independent of cancer status.â
â We have included additional text in the Introduction to clarify what cell-free DNA is:
â Page 3: âCell-free DNA (cfDNA) are short fragments of DNA released, mainly by apoptotic cells, into bodily fluids (7).â
â Due to wording limits, fragmentomic concepts are introduced in the Results section as required to understand each analysis and concept. Relevant citations are also all included throughout the text which highlight studies associated with each concept. This decision was made to reduce redundancy in repeating concepts in both the Introduction and Results section and also to maintain the citation list within the guidelines of Nature Communications.
2. In the introduction, it is important to state they also looked at 30 healthy TP53-wild-type controls.
â We have now added in the Introduction:
â Page 3: âWe comprehensively characterize the cfDNA fragmentomic landscape of a cohort of 169 blood samples collected from 82 TP53m-carriers and 30 TP53-wildtype individuals (Figure 1A/B).â
Reviewer #3 (Remarks to the Author): expert in TP53 genomics

In this manuscript, Wong and colleagues aim to assess if mutations in the tumor suppressor TP53, which are prevalent in cancer and when occurring in the germline cause the cancer-predisposing Li-Fraumeni Syndrome (LFS), is associated with specific patterns of circulating cell-free DNA (cfDNA) fragmentation, with potential diagnostic/prognostic value. To this aim, the authors got liquid biopsies from a cohort of cancer-free/past-cancer/active-cancer LFS patients and healthy wtTP53 individuals and performed cfDNA fragmentomics to detect nucleosome-free/degradation-prone and nucleosome-protected/degradation-resistant DNA fragments.

Data analysis showed that TP53 missense/splice/LOF LFS mutations were associated, independent of cancer status, with reduced cfDNA fragment size. Analyses of fragments' end and bp composition indicated increased cleavage of nucleosome-associated sequences, resulting in fragments with specific sequence patterns. Interestingly, mapping LFS fragment profiles with annotated TFBS identified a set of BS for developmental TFs whose accessibility correlated with cancer development. Furthermore, fragment profiling could also detect, in active-cancer LFS, increased accessibility in chromatin regions that were found more accessible also in advanced cancer from TCGA ATAC-seq datasets. The authors also developed ML algorithms validating the diagnostic/prognostic value of cfDNA profiles in LFS. These results led to the identification of cfDNA markers of cancer-free LFS, and also of possible nucleosome alterations, associated with TP53 mutations, correlating with cancer development, with possible diagnostic/prognostic value in LFS and potentially also in somatic cancers driven by TP53 mutations. However, I have some concerns that should be addressed to improve the MS.

We thank this Reviewer for their review and comments. We have performed additional analyses to address this Reviewerâs comments and to further validate our findings. Further analyses assessing the impact of TP53 mutation on gene expression have been performed using data from the TCGA as well as in-house data. These analyses have been included as additional supplemental figure panels. We have also included Figure 9 as a graphical conclusion figure to improve readability. We hope that these analyses

and additions have addressed this Reviewerâs concerns.

1) Data are very dense and not always clearly described. In some cases the use of different terms to describe the same readout is confusing. Overall, the text should be revised to improve readability. As one example, it should be briefly introduced what the parameters âmidpoint coverageâ, âcentral coverageâ, âamplitudeâ, âvariabilityâ, and âmeta peaksâ represent. The explanation that decreased âcentral coverage = increased accessibility/bindingâ, âindicative of binding activityâ, should be anticipated from line 309 to the section describing nucleosome positioning analyses (lines 277-281). Also, whenever possible, biologically meaningful terms should be used instead of technical parameters (e.g. replacing âdecreased central/midpoint coverageâ with âincreased accessibilityâ).
â We have carefully reviewed the Results section and have made changes in the text to clarify this and to increase the ease of reading and understanding. These changes have been most extensive in the nucleosome placement sections of the manuscript.
â A sentence has been added earlier in the Results section introducing and linking the parameters of the nucleosome positioning analysis to biologically meaningful terms:
â Page 9: âThe coverage at the p53 binding site (midpoint coverage) is indicative of binding activity (decreased central coverage = increased accessibility/binding), the amplitude is indicative of the robustness of the nucleosome placement surrounding the binding site within a sample, and the variability (standard deviation) is indicative of the robustness of nucleosome positioning throughout the cohort (36).â

2) To support the relevance of the identification of TFBS whose accessibility correlates with cancer development in LFS, the authors should assess whether the TFs targets are upregulated in mutant p53 cancers, by analyzing available LFS and TCGA RNAseq datasets.
â Similar observations have previously been made in iPSC derived osteoblasts from LFS and non-LFS individuals (20). We have performed additional analyses looking at TF and developmental gene sets in TP53-inactive vs TP53-intact tumors across all of the TCGA cancer types (28) and in LFS and non-LFS-derived osteosarcomas and glioblastomas (29). While we did not consistently observe the same TFs, this difference is likely due to the tissue and cellular origin dependent context of transcriptional regulation. Our cfDNA dataset is derived primarily from normal lymphocytes and other white blood cell lineages whereas the TCGA and tumor datasets are non-blood derived and are primarily from malignant cell populations. Despite this, we did observe an enrichment of developmental TF gene sets in TP53-inactive and LFS-derived tumors.
â Page 10: TFs with increased chromatin accessibility, in TP53m-carriers, were associated with embryonic organ development (GATA3, GATA4, GATA6, PRDM1, ASCL2, TCF7, and NKX3.1) which is consistent with previous observations in LFS-derived osteoblasts compared to TP53-wildtype (11). To further validate these observations, we compared the transcriptional profiles of TP53-inactive and TP53-intact cancers across 26 cohorts in The Cancer Genome Atlas (TCGA)(44) and in LFS-derived and sporadic osteosarcomas and glioblastoma (45). In both contexts, we observed an enrichment in developmental and transcription factor gene sets in TP53 inactive and LFS-derived tumors compared to TP53 wildtype and non-LFS derived tumors (Supplemental Figure 13B/C).
3) In Fig.1 the authors included a schematic of the whole study. To improve readability of results, graphical abstract(s) of conclusions should also be included.
â We have now included Figure 9 which outlines the conclusions of the study.

Minor concerns:

- In lines 288-298, please briefly describe the rationale for analyzing CTCF.
â CTCF binding sites are known to be evolutionarily conserved and to be stable even between different cell types (30). This information has been added to the manuscript including a citation:
â Page 9: âWhile nucleosome placement and positioning were consistent at evolutionarily conserved CCCTC-binding factor (CTCF) sites (39) and the TSS of 3,804 previously reported housekeeping genes (37),â

- In lines 281-284, the authors describe decreased midpoint coverage and increased amplitude at TSS of p53 targets but not housekeeping genes (Figure 6B, Supplemental Figure 12B). However, in lines 288-290, they state that TP53m-carriers displayed decreased central coverage at the TSS of housekeeping genes (Supplemental Figure 12A-B). Please clarify this apparent contradictory conclusion.
â We have modified this sentence as it was only partially correct. While we did not see increased amplitude or variability, we did observe increased accessibility at CTCF and housekeeping genes in TP53m-carriers. This has been clarified in the Results text:
â Page 9: âSimilar observations were also found at the transcription start site (TSS) of p53 target genes (n = 384; ANOVA p-values: midpoint = 0.284, amplitude <0.001; variability = 1.79x; Figure 6B, Supplemental Figure 12B)â
â Page 9: âWhile nucleosome placement and positioning were consistent at evolutionarily conserved CCCTC-binding factor (CTCF) sites and the TSS of 3,804 previously reported

housekeeping genes (37), we did observe increased accessibility at these sites in TP53m-carriers (Supplemental Figure 12A-B).â

- Lines 295-295: a steep drop-off of fragments <90bp in healthy TP53-wildtype at open-associated chromatin sites is reported, but no statistically significant difference is highlighted in Fig.S12H. Please clarify. Also, it should be better described how the authors reach the conclusion that, in TP53m-carriers, a decreased fraction of total reads suggests increased degradation.
â Figure S12H does not show a statistically significant difference as the p-value = 0.071. We have removed the statement suggesting statistical significance and have increased the clarity of this section.
â Page 10: âAt open-associated chromatin sites, we observed a steep drop-off of fragments less than 90bp in TP53-wildtype which was not observed in TP53m-carriers (Supplemental Figure 12H). TP53m-carriers also showed a decreased fraction of total reads spanning both open and closed chromatin sites (Supplemental Figure 12I). Read coverage of cfDNA within a genomic region is dependent on the activity and accessibility of nucleases at that specific site (41). Thus, the combination of decreased coverage and a higher prevalence of shorter cfDNA fragments suggests that the increased accessibility observed in TP53m-carriers may be due to increased fragmentation and degradation.â

- Nucleosome positioning (midpoint and amplitude) was not found to be significantly different across different TP53 mutation classes (Supplemental Figure 12C-F). Please clarify what groups were compared in statistical analyses. It seems that some mutations behave differently (e.g. LOF impacts amplitude only in p53 TSS, Supplemental Figure 12E-F). Please discuss these differences.
â The significance values were calculated comparing to TP53-wildtype samples. Multiple pair-wise comparisons using the Kruskal-Wallis test within our LFS cohort shows not statistical difference between the mutation clusters. We have included these p-values in the manuscript:
â Page 9: âNucleosome profiles were not significantly different across different TP53 mutation classes (Kruskal-Wallis p-values: p53 binding sites midpoint = 0.576, amplitude = 0.578; TP53 target genes midpoint = 0.435, amplitude = 0.547; Supplemental Figure 12C-F).

- Some Figure panels are not referenced in the text (e.g. 3C-D). Please check them all.
â We have reviewed the manuscript and now ensured all figure panels are referenced in the text.
- In some Figure panels, some annotation is lacking/cropped (e.g. asterisks indicating statistically significant differences in Fig. S12). Please review all figure panels.
â We have gone through the figure panels and made adjustments as necessary.

- It would be helpful to indicate, in Fig.7A and S13A plots, the TFs described in lines 310-314.
â Figures 7A and S13A have been revised to call out the specific transcription factors mentioned in the text.
References
1. Szymanski JJ, Sundby RT, Jones PA, Srihari D, Earland N, Harris PK, et al. Cell-free DNA ultra-low-pass whole genome sequencing to distinguish malignant peripheral nerve sheath tumor (MPNST) from its benign precursor lesion: A cross-sectional study. Abbosh C, editor. PLoS Med. 2021 Aug 31;18(8):e1003734.

2. Zhu G, Guo YA, Ho D, Poon P, Poh ZW, Wong PM, et al. Tissue-specific cell-free DNA degradation quantifies circulating tumor DNA burden. Nature Communications. 2021 Apr 13;12(1):2229.

3. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature. 2019 Jun;570(7761):385â9.

4. Ivanov M, Baranova A, Butler T, Spellman P, Mileyko V. Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation. BMC Genomics. 2015 Dec;16(S13):S1.

5. Sanchez C, Roch B, Mazard T, Blache P, Dache ZAA, Pastor B, et al. Circulating nuclear DNA structural features, origins, and complete size profile revealed by fragmentomics. JCI Insight. 2021 Apr 8;6(7):e144561.

6. Kato S, Han SY, Liu W, Otsuka K, Shibata H, Kanamaru R, et al. Understanding the functionâ structure and functionâmutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci USA. 2003 Jul 8;100(14):8424â9.

7. Giacomelli AO, Yang X, Lintner RE, McFarland JM, Duby M, Kim J, et al. Mutational processes shape the landscape of TP53 mutations in human cancer. Nat Genet. 2018 Oct;50(10):1381â7.

8. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. Cell. 2016 Jan 14;164(1):57â68.

9. Wong D, Luo P, Oldfield LE, Gong H, Brunga L, Rabinowicz R, et al. Early Cancer Detection in Li-Fraumeni Syndrome with Cell-Free DNA. Cancer Discovery [Internet]. 2023 Oct 24 [cited 2023 Dec 16]; Available from: https://aacrjournals.org/cancerdiscovery/article/doi/10.1158/2159-8290.CD-23-0456/729598/Early-Cancer-Detection-in-Li-Fraumeni-Syndrome

10. Jiang P, Sun K, Peng W, Cheng SH, Ni M, Yeung PC, et al. Plasma DNA end motif profiling as a fragmentomic marker in cancer, pregnancy and transplantation. Cancer Discov. 2020 Feb 28;CD-19-0622.

11. Lo YMD, Han DSC, Jiang P, Chiu RWK. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. Science. 2021 Apr 9;372(6538):eaaw3616.

12. Sun K, Jiang P, Cheng SH, Cheng THT, Wong J, Wong VWS, et al. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. Genome Res. 2019 Mar;29(3):418â27.

13. Chan RWY, Serpas L, Ni M, Volpi S, Hiraki LT, Tam LS, et al. Plasma DNA Profile Associated with DNASE1L3 Gene Mutations: Clinical Observations, Relationships to Nuclease Substrate Preference, and In Vivo Correction. The American Journal of Human Genetics. 2020 Nov;107(5):882â94.
14. Han DSC, Ni M, Chan RWY, Chan VWH, Lui KO, Chiu RWK, et al. The Biology of Cell-free DNA Fragmentation and the Roles of DNASE1, DNASE1L3, and DFFB. The American Journal of Human Genetics. 2020 Feb;106(2):202â14.

15. Fridlich O, Peretz A, Fox-Fisher I, Pyanzin S, Dadon Z, Shcolnik E, et al. Elevated cfDNA after exercise is derived primarily from mature polymorphonuclear neutrophils, with a minor contribution of cardiomyocytes. Cell Reports Medicine. 2023 Jun;4(6):101074.

16. Neuberger EWI, Sontag S, Brahmer A, Philippi KFA, Radsak MP, Wagner W, et al. Physical activity specifically evokes release of cell-free DNA from granulocytes thereby affecting liquid biopsy. Clin Epigenet. 2022 Dec;14(1):29.

17. Camuzi Zovico P, Gasparini Neto V, VenÃ¢ncio F, Soares Miguel G, GraÃ§a Pedrosa R, Kenji Haraguchi F, et al. Cell-Free DNA as an Obesity Biomarker. Physiol Res. 2020 Jun 30;515â20.

18. Hopkins MK, Koelper N, Caldwell S, Dyr B, Dugoff L. Obesity and no-call results: optimal timing of cell-free DNA testing and redraw. American Journal of Obstetrics and Gynecology. 2021 Sep;225(3):351.

19. Stejskal P, Goodarzi H, Srovnal J, HajdÃºch M, Van âT Veer LJ, Magbanua MJM. Circulating tumor nucleic acids: biology, release mechanisms, and clinical relevance. Mol Cancer. 2023 Jan 21;22(1):15.

20. Lee DF, Su J, Kim HS, Chang B, Papatsenko D, Zhao R, et al. Modeling Familial Cancer with Induced Pluripotent Stem Cells. Cell. 2015 Apr;161(2):240â54.

21. Samuel N, Wilson G, Lemire M, Id Said B, Lou Y, Li W, et al. Genome-Wide DNA Methylation Analysis Reveals Epigenetic Dysregulation of MicroRNA-34A in TP53 -Associated Cancer Susceptibility. JCO. 2016 Oct 20;34(30):3697â704.

22. Doebley AL, Ko M, Liao H, Cruikshank AE, Kikawa C, Santos K, et al. Griffin: Framework for clinical cancer subtyping from nucleosome profiling of cell-free DNA [Internet]. Oncology; 2021 Sep [cited 2021 Dec 6]. Available from: http://medrxiv.org/lookup/doi/10.1101/2021.08.31.21262867

23. Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. Nat Genet. 2016 Oct;48(10):1273â8.

24. Ulz P, Perakis S, Zhou Q, Moser T, Belic J, Lazzeri I, et al. Inference of transcription factor binding from cell-free DNA enables tumor subtype prediction and early detection. Nat Commun. 2019 Dec;10(1):4666.

25. Kolmykov S, Yevshin I, Kulyashov M, Sharipov R, Kondrakhin Y, Makeev VJ, et al. GTRD: an integrated view of transcription regulation. Nucleic Acids Research. 2021 Jan 8;49(D1):D104â11.

26. Yevshin I, Sharipov R, Kolmykov S, Kondrakhin Y, Kolpakov F. GTRD: a database on gene transcription regulationâ2019 update. Nucleic Acids Research. 2019 Jan 8;47(D1):D100â5.

27. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human

genome. Nature. 2012 Sep;489(7414):57â74.
28. Donehower LA, Soussi T, Korkut A, Liu Y, Schultz A, Cardenas M, et al. Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas. Cell Reports. 2019 Jul;28(5):1370-1384.e5.

29. Villani A, Davidson S, Kanwar N, Lo WW, Li Y, Cohen-Gogo S, et al. The clinical utility of integrative genomics in childhood cancer extends beyond targetable mutations. Nat Cancer. 2022 Dec 30;4(2):203â21.

30. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. Cell. 2007 Mar;128(6):1231â45.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Regarding the revised version of the manuscript, most of my comments were addressed. However, I have a few more comments.

- Definition shall be more accurate. For instance, in lines 116-117, the cell-free DNA molecules shall not be defined as "short fragments". Many latest studies have proved the utility of long cfDNA as clinical biomarkers.
- Some logic is not convincing enough. For example, in lines 126-128, the author hypothesized that germline events can affect cfDNA fragmentation based on previous research discussing the different fragmentation patterns in cases with DNASE1L3 mutation. However, DNASE1L3 is a nuclease that has a direct effect on the fragmentation process. The resulting altered fragmentation profile from TP53 mutation is not comparable to DNASE1L3. In a word, the premise the author listed can not effectively support his hypothesis.
-In lines 156-159, the author described a serial cohort consisting of 122 samples from 36 mutation carriers. However, only 16 carriers are mentioned. How about the other 20 carriers?
- In lines 245-247, I don't see a decreased fragment-end frequency at the peaks in Figure 1A.

Reviewer #2

(Remarks to the Author)
My concerns have mostly been addressed, with two exceptions:

Point 4 response – “This 3-dimensional chromatin architecture influences the representation of cfDNA fragments (size and location) in circulation due to the differential accessibility of nucleases throughout the genome (10).” => I examined reference (10) cited in the main text which is Jay Shendure’s work in Cell, 2016. In this paper, Shendure and colleagues show that nucleosome positioning and certain transcription factor binding sites (such as CTCF) can be inferred from cell-free DNA. One sentence in the abstract says “The cfDNA nucleosome occupancies correlate well with the nuclear architecture, gene structure and expression observed in cells, suggesting that they could indicate the cell-type of origin.” However, I believe the “nuclear architecture” they are referring to is the nucleosome positioning and transcription factor binding sites. Nowhere in the article do they do infer any conclusion that may be related to Hi-C, HiChIP, ChIA-PET, or any of the other 3D chromatin architecture datasets. As such, this sentence should be revised for clarity.

Point 5 response – Given the authors have acknowledged such a study is out of scope, then, I think the sentence “our observations suggest that TP53m-carriers exhibit differences in cfDNA fragmentation that may be driven by abnormal chromatin organization in specific tissues” should be removed.

Reviewer #3

(Remarks to the Author)
the authors addressed all the concerns.
here is only a minor point regarding the cfDNA dataset. authors should clarify its origin early in the manuscript, and it should be explained how the origin (normal lymphocytes and other white blood cell lineages..) of cfDNA was established.

Author Rebuttal letter:

We thank the Reviewers for their thorough review of our revised manuscript. See below for a point-by-point response:

Reviewer #1 (Remarks to the Author):

Regarding the revised version of the manuscript, most of my comments were addressed. However, I have a few more comments.

- Definition shall be more accurate. For instance, in lines 116-117, the cell-free DNA molecules shall not be defined as "short fragments". Many latest studies have proved the utility of long cfDNA as clinical biomarkers.
ï· We define the meaning of âshort fragmentsâ as cfDNA fragments >150bp in length at the first mention within the text. We also explicitly state we are only investigating the role of short cfDNA fragments in this study. Long cfDNA fragments have been proven to have utility but is not within the scope of this study.
ï· To improve accuracy, we have changed mention of âshort fragmentsâ to âshort cfDNA fragmentsâ throughout the text.
- Some logic is not convincing enough. For example, in lines 126-128, the author hypothesized that germline events can affect cfDNA fragmentation based on previous research discussing the different fragmentation patterns in cases with DNASE1L3 mutation. However, DNASE1L3 is a nuclease that has a direct effect on the fragmentation process. The resulting altered fragmentation profile from TP53 mutation is not comparable to DNASE1L3. In a word, the premise the author listed can not effectively support his hypothesis.
ï· We have removed this premise from the introduction and instead have emphasized that LFS patients have been found to have differential methylation and gene expression independent of cancer status and the role of the chromatin landscape on the generation of cfDNA.
-In lines 156-159, the author described a serial cohort consisting of 122 samples from 36 mutation carriers. However, only 16 carriers are mentioned. How about the other 20 carriers?
ï· The remaining carriers were cancer negative at all timepoints. This has been added to the text.
- In lines 245-247, I don't see a decreased fragment-end frequency at the peaks in Figure 1A.
ï· Figure 3A, bottom panels show the absolute difference in frequency which clearly show a dramatic drop in the frequency at +/-83bp from the nucleosome peak (0bp). This is also shown in the middle panels which display a Z score compared to the cohort of healthy TP53-wildtype controls.

Reviewer #2 (Remarks to the Author):

My concerns have mostly been addressed, with two exceptions:

Point 4 response â âThis 3-dimensional chromatin architecture influences the representation of cfDNA fragments (size and location) in circulation due to the differential accessibility of nucleases throughout the genome (10).â => I examined reference (10) cited in the main text which is Jay Shendureâs work in Cell, 2016. In this paper, Shendure and colleagues show that nucleosome positioning and certain transcription factor binding sites (such as CTCF) can be inferred from cell-free DNA. One sentence in the abstract says âThe cfDNA nucleosome occupancies correlate well with the nuclear architecture, gene structure and expression observed in cells, suggesting that they could indicate the cell-type of origin.â However, I believe the ânuclear architectureâ they are referring to is the nucleosome positioning and transcription factor binding sites. Nowhere in the article do they do infer any conclusion that may be related to Hi-C, HiChIP, ChIA-PET, or any of the other 3D chromatin architecture datasets. As such, this sentence should be revised for clarity.
ï· We have included an additional reference (1) which compared cfDNA fragmentation profiles from Hi-C data derived from lymphoblastoid cell lines.

Point 5 response â Given the authors have acknowledged such a study is out of scope, then, I think the sentence âour observations suggest that TP53m-carriers exhibit differences in cfDNA fragmentation that may be driven by abnormal chromatin organization in specific tissuesâ should be removed.
ï· We have altered this sentence to instead focus on the known epigenetic and transcriptomic differences observed in LFS individuals and how that may be driving the differences in cfDNA fragmentation observed in our study:
o âAs cfDNA fragmentation is known to be driven by the chromatin architecture of the cell-of-origin, our observations suggest that TP53m-carriers exhibit differences in cfDNA fragmentation that may be driven by epigenetic and transcriptomic differences compared to TP53-wildtype individuals (11, 12).â

Reviewer #2 (Remarks on code availability):

I am not a bioinformatician by training (my lab has bioinformaticians but I myself do not feel

confident to review code).

Reviewer #3 (Remarks to the Author):

the authors addressed all the concerns.
here is only a minor point regarding the cfDNA dataset. authors should clarify its origin early in the manuscript, and it should be explained how the origin (normal lymphocytes and other white blood cell lineages..) of cfDNA was established.
ï· We have added an additional statement in the Introduction:
o âCell-free DNA (cfDNA) are fragments of DNA released, mainly by apoptotic cells, into bodily fluids such as lymphocytes into the blood plasma (2).â
o âWe comprehensively characterize the cfDNA fragmentomic landscape of a cohort of 169 blood plasma samples collected from 82 TP53m-carriers and 30 TP53-wildtype individuals (Figure 1A/B).â
References
1. S. Cristiano, A. Leal, J. Phallen, J. Fiksel, V. Adleff, D. C. Bruhm, S. Ã. Jensen, J. E. Medina, C. Hruban, J. R. White, D. N. Palsgrove, N. Niknafs, V. Anagnostou, P. Forde, J. Naidoo, K. Marrone, J. Brahmer, B. D. Woodward, H. Husain, K. L. van Rooijen, M.-B. W. Ãrntoft, A. H. Madsen, C. J. H. van de Velde, M. Verheij, A. Cats, C. J. A. Punt, G. R. Vink, N. C. T. van Grieken, M. Koopman, R. J. A. Fijneman, J. S. Johansen, H. J. Nielsen, G. A. Meijer, C. L. Andersen, R. B. Scharpf, V. E. Velculescu, Genome-wide cell-free DNA fragmentation in patients with cancer. Nature 570, 385â389 (2019).

2. D. Zhu, H. Wang, W. Wu, S. Geng, G. Zhong, Y. Li, H. Guo, G. Long, Q. Ren, Y. Luan, C. Duan, B. Wei, J. Ma, S. Li, J. Zhou, M. Mao, Circulating cell-free DNA fragmentation is a stepwise and conserved process linked to apoptosis. BMC Biol 21, 253 (2023).

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.
The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/