Peer Review File

Article Information: https://dx.doi.org/10.21037/tlcr-21-728

Reviewer A

This is a timely and well-written manuscript on an important issue, addressing the need for increased specificity and sensitivity for detection of rare DNA mutations in liquid biopsies.

Comment 1: Minor comments: genes and alleles must be italicized.

Reply 1: Thank you for your comment. The appropriate changes have been made to the text.

Changes in the text: Changes were made throughout the text.

Comment 2: Please follow recommendations of the Minimum Information for Publication of Quantitative Digital PCR Experiments and include a table with the correspondent checklist.

Reply 2: Thank you for bringing this checklist to our attention. We have filled out the checklist and included it as a Supplemental Table.

Changes in the text: Added to document containing all tables for the manuscript.

Comment 3: Please correct the usage of "based off of..." in the text, e.g. the sentence in the "Experimental design" topic: "Samples were excluded based off of thermal cycling failure."

Reply 3: We appreciate your comment; the sentence was deleted due to changes to the structure and content of the paragraph.

Changes in the text: N/A

Comment 4: Topic "Precision and accuracy of both dPCR methods": most details shall rather be included in the legend of figure 2, instead in the text ("orange", "blue", "subfigure b", etc.) Restrict the text to results/conclusions.

Reply 4: Thank you for this insightful comment. We agree that this section would be better suited in the legend of Figure 2. The appropriate changes have been made to the text.

Changes in the text: In "Figure 2" legend (Pg. 33, Lines: 655-658, 659-664)

(b) MAF data individually plotted. Expected MAF were plotted on X-axis and the detection results were on Y-axis. This graph shows the degree of precision for endpoint and real-time dPCR through the distribution of datapoints with smaller distributions

indicating greater precision.

(d) Percent recovery plotted for each of the replications that were spiked with 0.06%, 0.2% or 2%. Red line indicates 100% recovery, blue markers are for endpoint dPCR results and orange markers are for real-time dPCR results. This graph shows both precision and accuracy. Precision is shown by the distribution of datapoints; accuracy is shown by the distance of datapoints to the 100% recovery line. The closer the values are to the 100% recovery line, the more accurate they are.

Comment 5: "were run" is repeated twice in the following sentence: "Additional chips at 0% and 0.06% were run for EGFR 19del samples were run to determine the LoD"

Reply 5: Thank you for pointing out this error. The appropriate changes have been made to the text.

Changes in the text: In "Materials and Methods" section, "Sample Batching" subsection (pg. 25, Lines: 522-523)

Additional chips at 0.00% and 0.06% were run for EGFR 19del samples to determine the LoD.

Comment 6: Use blue and orange in figure 1 (black and blue are hard to differentiate).

Reply 6: Thank you for bringing this point to our attention. We have changed the format of the graph so that it's easier to read.

Changes in the text: The graph was changed and added to the PDF containing all figures for this manuscript.

Comment 7: Discuss the application of the real-time dPCR for different kinds of samples, too (as formalin fixed paraffin embedded samples, see the work of Carvalho et al. 21 : https://doi.org/10.1016/j.mcp.2021.101745)

Reply 7: We appreciate you bringing this point to our attention. While real-time dPCR could be used on other kinds of samples, such as tissue samples, the benefit would not outweigh the convenience of using pre-established methods. For tissue samples, in particular, qPCR has been successfully used to test these samples. However, in a recently published paper by Xu et al., it found that for liquid biopsy samples, qPCR wasn't as sensitive; out of the 13 samples that were determined positive with tissue biopsy, qPCR only found 4 liquid biopsy samples positive. Real-time dPCR has been found to be more sensitive than both qPCR and endpoint dPCR, as discussed in this paper, finding 7 out of 13 samples positive. The benefit of real-

time dPCR, as explained in the text, is detecting low allele frequencies; because samples such as tissue samples don't require this degree of sensitivity, it would be more important to use this technology to fill the deficit in liquid biopsy testing.

Changes in the text: N/A

Reviewer B

Xu et all compared the assay performance of a digital PCR instrument with quantitative real-time PCR. Author concluded that the novel technology improved sensitivity with the clinical application potential.

The manuscript is NOT well prepared, among all the points that need to be further addressed , several key pointed need to clarified and discussed.

Comment 1: Apparently, there also, are FDA approved NGS panels (e.g. Guardant Health 360 CDx) that has been broadly used , which shall be compared and discussed. It is therefore highly recommended comparing with available NGS panel based approach rather than only using the RT-PCR as the reference point

Reply 1: Thank you for your comment. While NGS is FDA approved, it functions more as a screening tool than a companion diagnostic tool. The issue withs NGS lies in its two ligation steps. Both reactions have an expected efficiency 20%. Together, the expected yield after both ligation steps is around 4%.

One of the issues with NGS is that it isn't as sensitive as real-time dPCR. In this study, we discuss that real-time dPCR can detect allele frequencies at 0.1% or below. This would be around 6 copies a particular mutation. However, if 6 copies are expected, it is unclear how many would be present in the NGS library prep. Due to the combined ligation efficiency being around 4%, it's not likely than any of the 6 total copies would be present.

Another issue with NGS is that it is used in a different context than real-time dPCR. As stated previously, NGS is more of a screening tool than a companion diagnostic tool. While NGS is good for calculating tumor burden, it is likely to get random background. This is due a combination of the annealing step as well as low copy numbers.

Lastly, while people are focused on improving the sensitivity of NGS, these efforts are made after the pre-analytical stage. However, as I discussed above, the issue lies in the library not including the mutation in question because of its low copy numbers. Any additional optimization after the pre-analytical stage won't improve the number of copies detected.

Changes in the text: In “Discussion Section”, “The advantage of real-time dPCR over NGS for CDx” subsection (Pg. 20-21, Line: 427-442)

Although there are many benefits to real-time dPCR, some would argue that next generation sequencing (NGS) is the future of companion diagnostics. While NGS assays are extremely high throughput and tremendous efforts were put in to improve its sensitivity by reducing the amount of false positive reads, however, the biggest drawback for NGS lies in its pre-analytical process. Library preparation for NGS has two ligation steps which adds sequencing adaptors to either end of fragmented DNA. This issue with these ligation steps is that each has an efficiency of around 20% (27, 28). Combined, the ligation steps have around 4% yield. As a result, if there are low copy numbers to be expected, such as 3-4 copies with 0.06% MAF, there is a very slim likelihood that these mutations will be present in the library. While there has been focus on improving the sensitivity of NGS, these measures are after the pre-analytical phase. Therefore, if the mutations are not present in the library, there is no way they are going to found after pre-analysis, regardless of the sensitivity after sequencing libraries are made. Additionally, NGS is prone to get random background due to a combination of the annealing step and low copy numbers. This worsens the sensitivity of NGS at low allele frequencies. While there are some benefits to NGS, such as calculating tumor mutation burden, for the purposes of companion diagnostics, it is not as sensitive as real-time dPCR at low allele frequencies.

Comment 2: Without the clear definition of the analytical cut off, it would be very difficult to compare the assay performance.

Reply 2: Thank you for addressing your concerns regarding the analytical cut off. The performance of the assays would not need to be compared between the different instruments as all the chips were run on the same instrument. Additionally, because the purpose of this paper is to compare and validate the instruments, and not the assays themselves, it is not necessary to compare the performance of the assay.

Changes in the text: N/A

Comment 3: The assay performance parameter and acceptance critical is not clear.

Reply 3: Thank you for your comment. The same chip was used to compare endpoint and real-time dPCR so there wouldn’t be an acceptance criterion for these chips. The only instance in which a chip wouldn’t be accepted if there were QC issues, which we see with the 19del data. Additionally, the purpose of this paper is to discuss the real-time dPCR instrument, not to evaluate the assays, therefore, having an assay performance parameter is not necessary.

Changes in the text: N/A

The manuscript needs extensive modification.

ABSTRCT

Comment 4: There is no method description in abstract, please include it.

Reply 4: Thank you for bringing this point to our attention. The appropriate changes have been made to the text.

Changes in the text: In the "Abstract" Section, "Methods" Subsection (Pg. 4, Lines: 82-88)

The novel real-time digital PCR thermal cycler was compared to an endpoint digital PCR instrument to determine the sensitivity and quantification accuracy of both instruments. Samples were all thermal cycled on the same instrument but were read on two separate instruments to collect real-time and endpoint data. Contrived samples for EGFR exon 19 deletion, T790M, and L858R as well as HER2 were tested. Different mutant allele frequencies and wild-type to mutant gene ratios were tested for EGFR and HER2, respectively.

INTRODUCTION

Comment 5: The clinical utilizes of the liquid biopsy is well recognized, please combine the fist and second paragraph of the introduction and significantly simplified it.

Reply 5: We appreciate and agree with your comment. We have made changes to combine both paragraphs in the text.

Changes in the text: In the "Introduction" section (Pg. 6, Lines: 109-126)

Due to the prevalent issue of temporal and spatial tumor heterogeneity for tissue biopsy, liquid biopsy is an important tool for companion diagnostics (CDx) to avoid false negative test results, especially for patients with multiple tumors and large tumors. Currently, liquid biopsy tests for CDx are not as sensitive as tissue biopsy. A previously published paper by Xu et al. found that out of 13 samples that were all determined positive by tissue biopsy, 4 were determined positive by Roche Diagnostics' Cobas-ARMS PCR kit and 7 were determined positive by Thermo Fisher Scientific's QuantStudio™ 3D Digital PCR System (1). While liquid biopsy assays are needed for CDx, there are still limitations regarding this technology. This mindset is reflected in the guidelines of the National Comprehensive Cancer Network (NCCN) for the use of liquid biopsy as the parameters around its use has changed twice within the past few years. The initial change in 2019 recommended the use of liquid biopsy for non-small cell lung cancer (NSCLC) patients for EGFR mutation detection only when tissue biopsy is not available (2). The second change only a year later states that liquid biopsy

can be ordered first—positive patients go directly to targeted therapy, negative patients will get another test from a tissue biopsy, if possible (3). While these changes reflect growing confidence in liquid biopsy, tissue biopsy is still needed on negative samples due to the lack of sensitivity of current liquid biopsy tests (4-5). Overcoming poor sensitivity is going to give liquid biopsy the strength it needs to be comparable with tissue biopsy as a CDx test.

Comment 6: I am not quite sure the third paragraph is necessary as it is note relevant the current manuscript which is NOT using the patient samples and please deleted it. The agreement between the tissue and liquid biopsy are associated with different factors such as tumor stage, shedding and treatment etc.

Reply 6: Thank you for voicing your concern. While this paper doesn't discuss patient samples, it is important to talk about these samples because the ultimate goal of the real-time dPCR instrument is to be used in a clinical setting for companion diagnostics. In a previous study, Wang et al. (2020) compared the viral detection of SARS-CoV-2 with endpoint and real-time dPCR. When retesting 'aged' samples which has previously been determined as positive, real-time dPCR had a positive rate of 86.36% while endpoint dPCR had a positive rate of 54.55%. Although the samples tested were degraded, there was a difference in the detection rate between endpoint and real-time dPCR. With this data as well as the results from the paper, it is easier to make the conclusion that real-time dPCR would have success in a clinical setting, not just in research.

Changes in the text: The third paragraph was incorporated into the first paragraph (see Comment 5 for changes in the text). An additional paragraph was added discussing the differences in quantification accuracy of real-time and endpoint dPCR for viral detection.

In the "Introduction" section (pg. 8, Lines: 166-169)

In different study, we have demonstrated improved sensitivity for viral detection. For SARS-CoV-2 LoD samples that were aged in -80°C conditions for 2 months, real-time dPCR determined 86.36% of samples as positive while endpoint dPCR only determined 54.55% of samples as positive suggesting a significant improvement of sensitivity by real-time dPCR (21).

Comment 7: How about the BioRad Droplet PCR system?

Reply 7: We appreciate you bringing this other method to our attention as we did not think to include it previously. The Bio-Rad Droplet Digital PCR System is equivalent to the Thermo Fisher Scientific chip-based endpoint dPCR system that was used in this study. Both systems use endpoint analysis and because Bio-Rad uses droplets, no real-time monitoring can be done because the individual droplets cannot be monitored.

Changes in the text: Information on the Bio-Rad ddPCR system was added to the Introduction.

In the "Introduction" section (Pg. 7, Lines: 144-148)

Another popular form of endpoint dPCR is droplet digital PCR (ddPCR). Bio-Rad's ddPCR™ technology (Bio-Rad, Hercules, CA) partitions PCR reaction mixture into ~20,000 droplets in a single step, then PCR amplification occurs in each droplet (17). The droplets, similar to chip-based dPCR, would contain 0 or 1 (or more) molecules of DNA or RNA and can be individually counted.

RESULTS;

Comment 8: 20ng input is quite highly for digital PCR assay, do you have any data for 5ng or in ng input to demonstrate the assay performance?

Reply 8: Thank you for this question as we did not directly discuss it in the paper. There is no data for DNA inputs below 20ng because at the lowest allele frequency we tested, 0.1% MAF, there are 6 copies expected. If the input DNA was lower than 20ng, the expected number of copies would go down as well; fractions of copies would be expected at these lower input amounts.

Changes in the text: In the "Discussion" section, "Other considerations for real-time dPCR on

CDx applications" subsection (Pg. 21, Lines: 444-451)

While the LoD is low, it may seem that the input of DNA, at 20ng, is high. However, with the lowest detected MAF at 0.06 or 0.03%, depending on the mutation, 3-4 copies of mutant alleles are expected. If the input amount of DNA were to decrease, fractions of copies would be expected making it impossible to detect mutations. In addition, it was found in a previously published paper that as the input amount of DNA decreases, the amount of background signal would increase in comparison to the signal (29). This would make it even more difficult to discern between real and false positive signal and to cut a threshold. For these reasons, 20ng is an appropriate amount of input DNA for real time dPCR rare allele assays.

Comment 9: How to define the cut off?

Reply 9: We appreciate your question regarding the cut-off value. While we do recognize that a cut-off value will eventually be needed, because the study was done with cell-line samples and not clinical samples, a cut-off value cannot be determined. For a cut-off value to be determined, either clinical samples and a comparator assay are needed or clinical outcomes from patient samples are needed. Additionally, the focus of this study was not on assay development or the limit of detection for the assays but instead on the performance of the

instruments. With these two points in mind, determining the cut-off value is important but outside of the scope for this particular study.

Changes in the text: N/A

DISCUSSION

Comment 10: It is recommended to add subheading in the discussion part.

Reply 10: We appreciate your recommendation and agree that subheading would be beneficial in the discussion section. The appropriate changes have been made to the text.

Changes in the text: In the discussion section, added subheadings

"Improving sensitivity allows real-time dPCR to benefit CDx in multiple avenues" (Pg. 19, Line: 386)

"The advantage of real-time dPCR over NGS for CDx" (Pg. 20, Line: 426)

"Other considerations for real-time dPCR on CDx applications" (Pg. 21, Line: 443)

"Final conclusions" (Pg. 22, Line: 472)

Comment 11: I am not quite sure the ADC drug part is relevant to the topic, please delete it .

Reply 11: We appreciate you bringing up your concern with the ADC drug portion of the discussion section. This portion was included to show that for an emerging field of therapy, a lower threshold than what was discussed in the paper will be required. This is due to ADC drugs working with IHC 1+ and 2+ samples. Because of the number of copies expected in these samples, the threshold, in general, will have to be lower than current clinical standards. Additionally, for patients who would be using ADC drugs, they are often late-stage patients who have undergone multiple lines of therapy and who have multiple and/or large tumors. Because of the reasons discussed in the paper, liquid biopsy would be the method to detect mutations for these patients. Therefore, due to the threshold having to be lowered and tissue biopsy not being a viable option for these patients, liquid biopsy tests with high sensitivity at low allele frequencies is necessary for this technology. These parameters fit with the technical capabilities of real-time dPCR which is why the section was included in the discussion section.

Changes in the text: N/A

FIGURE & tables

Comment 12: Figure 1: I am not quite sure it is the meaningful data to present the LOD , also, it shall be plot in the median.

Reply 12: Thank you for bringing your confusion to our attention. The figure was plotted to show that while both instruments detected the rare allele, compared to the negative control sample, the real-time dPCR values for rare allele frequency do not overlap with the negative control results. The boxplots display the median as the middle line of the box.

Changes in the text: N/A

Comment 13: Table 2: The precent of CVs are quite high in both platforms. what is the purpose of presented this table?

Reply 13: Thank you for asking questions regarding Table 2. As stated in the manuscript, in the "Results" section, "Precision and accuracy of both dPCR methods" subsection, the CVs are expected to be high because the allele frequencies being tested are very low. The CVs have been removed as they may confuse the reader.

Changes in the text: N/A