Department of Microbiology

University of Manitoba

45 Chancellors Circle, Room 414D/Laboratory 118, Buller building

Winnipeg, MB Canada R3T 2N2

204-290-9548

Miguel.Uyaguari@umanitoba.ca

September 28th, 2021

Dear Editor and Reviewers,

We thank you for having spent the time to read our work and give comments towards improving our manuscript. We have made edits to address your concerns.

Please find below a point-by-point response to the comments received. For readability purposes, the comments will be in bold.

We hope this version of the manuscript satisfies the comments raised by the reviewers and our manuscript can be published on peerJ.

Sincerely,

Miguel Uyaguari, PhD

Assistant Professor/Indigenous Scholar

Department of Microbiology

University of Manitoba

45 Chancellors Circle, Room 414D/Laboratory 118, Buller building

Winnipeg, MB Canada R3T 2N2

204-290-9548

On behalf of all authors.

**REVIEWER 1**

**Basic reporting**

**There were no major issues with regard to basic reporting. I found the article clear for the most part with some grammatical errors in a few places:**

**Line 163-165 and Table 1: I believe you need to use the exact name of the primers and probes from the papers you got them. I don’t think you should change the name from what it is in the paper you cited. Some of the papers you cited have multiple primers and probes (and each primer and probe may have different efficiencies) for each organism and it would cause confusion on which primer or probe you are using.**

Thanks for your comment. We agree. We have revised the way we refer to primers and probes. Now, their names are the same as those in the papers referenced.

**Figures 2-4: What is the difference between a and b figures? Please explain more clearly in caption or titles.**

Thanks for your observation. We have included additional clarification in the captions for these figures.

**Line 370: typo**

**Line 418: 'caution' instead of 'cautiousness'**

These comments have been addressed. Thanks.

**Experimental design**

**I believe the article had several challenges when it comes to experimental design which would need to be fixed for publication. One of the main challenges is the lack of replication in the experimental design which makes the study less rigorous. Previous studies have looked at this topic using 5-10 samples whereas this study only used 2 samples per season. Replication is important in this field as concentrations are highly variable in wastewater samples. Hence, comparisons cannot be be accurately made without more replication. I also don't believe you can compare concentrations received from 2 different samples if you used different viral capture and extraction approaches. Especially if you have not determined the relative viral capture efficiencies associated with each approach. Further line by line comments are stated below:**

**Line 96: Was it an actual virus with genome and capsid that were spiked in or was it just the RNA genome? If it is only the RNA genome, this will only give you efficiency of collecting genomes not collecting viruses which have different sizes. Also it is an RNA target which might have different capture efficiency from a DNA target.**

For the validation experiments, we used Armored RNA Internal Process Control (IPC), a commercial product from Asuragen. The genome size of this IPC is 1000 bp (link is included). There is an Armored DNA IPC, which contains 247 bp. The artificial sequences of Armored RNA and Armored DNA are overlapping; however, the RNA sequence is longer. Both artificial sequences contain ssDNA. The validation experiments here presented used reverse transcriptase and DNA polymerase for RNA and DNA quantification targeting this artificial ssDNA. These enzymes were the same type of enzymes used for the quantification of enteric viruses here presented.

<http://asuragen.com/wp-content/uploads/2021/05/Armored_RNA_IPC_Brochure.pdf>

**Line 114-116: Isn’t it possible that there are microbes on the solid waste or debris that was filtered by the funnel? Did you quantify the viruses on those? If so, please report those concentrations.**

The presence of microorganisms in the solids filtered out was likely and we have added to the manuscript some comments on the matter. However, this was not further investigated.

**Line 97: Please clarify what exactly was not used in the study.**

**Line 135-143: What is the typical recovery from this process? Maybe you could cite a paper that discusses this.**

**Line 228: why was this only done for raw sewage and not activated sludge and effluent? Wouldn’t there be different efficiencies in the different types of samples? the efficiency associated with capturing viruses from SC is also very important.**

**Line 343-344: Because the concentration methods are different, I don’t believe this (comparing between SC and RS) is a good comparison. Differences could be due to differences in extraction and concentration recovery efficiencies. I suggest doing recovery efficiencies for SC and all other treatment samples.**

Thanks for your observation. We have performed the necessary validation experiments for all sample types used in this study and included the recovery efficiencies in the manuscript.

**Figures 2-6: how are you making box plots with duplicate samples from each event and treatment? Did you take more than two samples? Or are you using the qPCR data as replicates? Could you clarify what exactly are the data points used to make these box plots. Could you also state the limit of detection.**

Thanks for your observation. We have now detailed what each data point of the box plot represents, as well as the cut-off Ct values to the manuscript.

**Line 330-333: Could also be resolved with more replication.**

**Line 352-355: You would need more sampling replicates than you have to compare between seasons.**

We agree. A longer and/or more thorough study with more samples being collected each season would certainly be beneficial. There have been restrictions in the facility due to the COVID-19 pandemic to continue doing a longer time series.

**Line 394-403: I think you have too few points for these analyses. Maybe these can be cut.**

We agree. More replicates would resolve this issue as well. We have removed the pertinent discussions.

**Line 376: Were there any culturable assays measured in this study? I think another limitation that can be stated is the use of only molecular assays which do not measure viability and can cause overestimation of risk.**

No, we only used culture-independent methods in this study. Like you said, this course of action certainly skewed our evaluations towards overestimation of risk. We have included this point in the discussion. Thanks for your observation.

**Validity of the findings**

**Since this study does not have good replication, it greatly impacts validity of findings and comparisons in the discussion. More comments below:**

**Line 227-229 and 303-307: This recovery efficiency is pretty low. Is it because you pre filtered your sample?**

We have updated this part of the results with yields from our recent validation experiments.

**Line 325-329: From your results, concentrations of indicators decreased in AS but then increased in EF after. How did this happen? Especially for viruses that cannot grow in wastewater. If your reasoning is the hydraulic retention time, could you explain that more.**

In the manuscript, we have now expounded on our line of reasoning behind hydraulic retention time being responsible for the AS – EF discrepancies.

**Line 356-359: Possibly because in that country, the winters aren’t as cold as Winnipeg hence less viral transmission.**

This is possible. However, given the various patterns of seasonal SaV detection worldwide (detailed in the apropos discussion), a purely temperature-based explanation to SaV presence may be insufficient.

**Figure 4: The concentration of E. Coli being lower than crass and Pmmv is very strange. Is this the usually concentration for this particular E. coli marker in raw sewage? If so, I believe other markers would be better such as EC 23S (https://doi.org/10.1021/es302222b) or even the culturable marker.**

Compared to other similar studies (now referenced in the apposite discussions), our *uidA* yields were much lower. This reinforces the notion introduced earlier that larger fractions (i.e., bacteria or microeukaryotes) may have been lost during the sample treatment process compared to a smaller fraction (viruses).

**Additional comments**

**I believe if more replication is done and viral capture efficiencies are evaluated for all samples, the paper can be accepted.**

Despite our preference for more samples being collected each season (and for more seasons), further replications are not possible at this time due to logistical reasons and budget assigned to this project. We have, however, performed the validation experiments as mentioned earlier. Samples used for validation experiments were obtained through our co-author Dr. Qiuyan Yuan as a special request.

**REVIEWER 2**

**Basic reporting**

**Thank you for allowing me to review your manuscript “Quantitation of human enteric viruses as alternative indicators of fecal pollution to evaluate wastewater treatment processes”. This novel and well written article provides findings on important research into using human enteric viruses as indicators of contamination in wastewater.**

**I thought the manuscript used clear and professional language throughout. The introduction provided a clear outline of background information on usage of E.coli as a means of detection in wastewater but that only using E.coli excludes the possibility of the presence of other contaminants. Relevant prior literature and citations were included.**

**All figures were relevant to the content of the article and the raw data is easily accessible and all links such as those to the Github pages were accessible. The results were directly related to the main hypothesis and self-contained in the article. The main conclusion that AdV, crAssphage and PMMV be used as viral indicators of water quality was supported in the results section.**

**I do I have some minor comments on language errors I found throughout the manuscript as well as a question about a figure:**

**- Line 39-40: “This indicate that..” should be “indicated”**

**- Line 370 – “The NESTP” Should be removed**

**- The authors use the word “quantitated” throughout the manuscript. I would suggest using “quantified” instead.**

Thanks for your observations. These comments have been addressed.

**- Figure S2B – It seems to be off the page or at the very edge of the page. Ensure that this is reasonable for publication.**

We had configured Tableau to print the figures themselves (as opposed to printing a predetermined area of the workspace, which might crop out anything beyond a set border), so we are confident the figures will be included in its entirety and will be suitable for publication.

**Experimental design**

**The design of the experiment seems robust with proper techniques utilized and proper citations for certain methods.**

**I did have a couple questions related to statistics –**

**1. Using the Spearman rank correlation analysis was the average of all gene copies taken (average of event 1-4) and compared in this way? Do you think using this average value skewed results if there are differences between samples taken in the fall compared to winter?**

Yes, and unfortunately, yes. Reviewer 1 had also mentioned some shortcomings of the study due to the lack of replicates, and more samples being collected over the season may possibly reduce, although not entirely, the skewness.

**2. Were any statistical tests carried out to compare samples collected in Fall vs Winter?**

There were only tests to compare samples between individual events (e.g., AdV in event 1 versus AdV in event 2), not between pooled seasonal samples (e.g., AdV in the fall season (events 1 & 2) versus AdV in the winter season (events 3 & 4)).

**3. Line 207 – The authors discuss the statistical analysis. Please identify what you are actually comparing here and throughout the results. Are you comparing the average GCN of one sample type to another across each viruses? It was difficult to follow and determine what was statistically significant.**

We agree, the results for this section could certainly use an overview of the quantities being compared. This overview has now been added to the manuscript.

**Validity of the findings**

**All the data is easily accessible, and findings are reliable. I thought the conclusion was well written and clearly outlined the end results. It is apparent that this work will benefit future work monitoring water quality.**