Author’s response to reviews

Title: MOLECULAR AND EVOLUTIONARY CHARACTERIZATION OF NOROVIRUS GII.17 IN THE NORTHERN REGION OF BRAZIL

Authors:

Larissa Costa (prado.cris.lnc@hotmail.com)
Dielle Teixeira (dielleteixeira@gmail.com)
Ana Rodrigues Portela (anacarolineportela@gmail.com)
Ian Lima (ianlima@iec.gov.br)
Renato Bandeira (renatobandeira@gmail.com)
Edivaldo Júnior (edivaldojunion@iec.gov.br)
Jones Siqueira (jones_siqueira@ymail.com)
Hugo Resque (hugoresque@iec.gov.br)
Luciana Silva (lucianadamascenadasilva@gmail.com)
Yvone Benchimol Gabbay (yvonegabbay@iec.pa.gov.br)

Version: 1 Date: 13 Sep 2019

Author’s response to reviews:

Dear Editor,

BMC Infections Diseases

We greatly appreciate the comments from all reviewers on our manuscript titled “MOLECULAR AND EVOLUTIONARY CHARACTERIZATION OF NOROVIRUS GII.17 IN THE NORTHERN REGION OF BRAZIL.” We agree with all the concerns raised by the reviewers. We have incorporated the reviewer suggestions in this revised manuscript. We take this opportunity to thank the evaluators for the suggested revisions, which have improved the quality of the manuscript. Some modifications were made in the main text of the manuscript and are highlighted in yellow to meet all requirements.

Point by point responses to reviewers

Yuki Matsushima (Reviewer 1): The authors conduct a molecular epidemiological study of human norovirus GII.17_Kawasaki2014 in the northern Brazil. They show the increase of detection of the virus in the region in the 2015-2016, the clade classification of detected strains, time-scaled
evolutionary process of the GII.17_Kawasaki2014 and a possibility of antigenic changes in the virus associated with the epidemics. Although there were few reports associated with the epidemiological study of the GII.17_Kawasaki2014 in Brazil, they should address some issues before consideration for publication to this journal. Detailed comments are described below.

(1) Method, design of NoV primers.

Could designed primers (LNOV 5KF/5KR and LNOV 6KF/6KR) detect all genotypes of human noroviruses? The authors described human norovirus detection rate of 32.2% of all samples. However, it was unclear whether or not they used the LNOV primers for detecting and genotyping all human noroviruses. Additionally, the performances for primers appear to be investigated only by the GII.17_2014 genotype. They should show evidence that all genotypes could be detected with similar sensitivity and specificity, and these sequences are conserved in the primer regions. Again, please clearly mention primers used in the screening stage for human noroviruses, the genotyping stage and the specifically sequencing stage for the GII.17_2014 strains.

R: Initially, NoV was detected by immunochromatography or enzyme immunoassay (EIA). The positive samples were subjected to one-step RT-PCR to amplify the ORF1/ORF2 junction region using the MON431/MON 432 and G2SKR (557 bp) primers, which can detect all NoV genotypes. The amplicons were then sequenced. The samples classified as NoV GII.17 were again subjected to one-step RT-PCR using specific primers designed in our study (LNOV 5KF/5KR e LNOV 6KF/6KR). These primers specifically amplify the complete ORF2 region of NoV GII.17 during sequencing. This enables a more accurate phylogenetic analyses. As suggested, this information was added in the text: lines 95-104.

(2) Method, evolutionary analyses.

They compared AICM values to select the best clock and tree prior models in the Bayesian MCMC inference. However, Baele and colleagues (Mol Biol Evol. 2012 Sep;29(9):2157-67) showed path sampling (PS) and stepping-stone (SS) sampling methods improved the accuracy of model selection compared to the harmonic mean estimator (HME) and AICM estimations. They should use the PS/SS sampling method to select the models.

R: As requested by the review we use the PS / SS sampling method to select watch models. Lines 165-168.

(3) Results and figure 2.

It would be helpful to readers to show the number of detection during year not only for the GII.17_2014 but also for other genotypes (not combined).

R: Modifications were made, as requested. See Fig. 2

(4) Results, sensitivity assay. Please also show the concentration of the number of RNA copies per microliter.

R: As recommended we add the information requested and it is found in line 116-117.
(5) Results, Bayesian MCMC.

They conducted the Bayesian MCMC analyses with own selection of the GII.17_2014 sequences. However, a selection bias could produce the inaccuracy of the results in the evolutionary rates and divergent times. They should analyze using all sequences registered in the GenBank without the selection of strains to minimize the biases.

R: Thank you for your suggestion. There are various NoV GII.17 sequences registered in Genbank. In this study, the selected database comprised 85 sequences (genome or complete ORF2). However, the related sequences share homology. Thus, to minimize the usage of computational resources, these sequences were clustered to find a representative for each group. We used all GII.17 sequences registered in the GenBank and the CD-Hit (Cluster Database at High Identity with Tolerance) to split sequences with 100% identity (Huang et al., 2010).

(6) Results, Bayesian MCMC.

The authors described introduction processes of the GII.17_2014 to the northern Brazil. However, I think the posterior probability values are likely low at the divergences points around the Brazilian strains as the nucleotide sequences have high similarities between the clade D strains of the GII.17_2014. Thus, I do not know whether or not the divergences around the Brazilian strains are highly accurate. In order to mention the justification of the introduce processes to Brazil, they should clearly show the related divergences at least have high posterior probability values.

R: Modifications were made, as requested.

(7) Results, evolutionary rates.

How was the number of strains included in the datasets for calculation of evolutionary rates, and what strains did they use for the analyses? Please show those. Moreover, they should exhibit the 95% HPD intervals for evolutionary rates.

R: As suggested, the information on the selection of samples for evolutionary analysis was included in the manuscript. Additionally, a spreadsheet with the sequence bank used for this study has been provided. See additional file 1.

(8) Results, structural analyses.

It was unclear how they determined binding residues of P2 domain to the HBGA fucose fraction. Did they perform docking simulation analyses using constructed structures and the fucose fraction, or did they superimpose the constructed structures to the template structures, including the fucose fraction? Please clarify those.

R: We performed the protein docking of P2 domain with the fucose fraction of HBG using the Autodock Vina program and verified the docking consistency by superimposing with the modeled structures. The detailed information was added in the revised manuscript in the lines 183-184 and 249-251.
(9) Discussion, evolutionary rates.

Parra et al. (PLoS Pathog. 2017 Jan 19;13(1):e1006136) showed lower evolutionary rates than the values by your analyses. Please discuss the inconsistency.
R: The modifications were made, as requested.

(10) Discussion, antigenic changes.

They discussed antigenic changes in the GII.17_2014 strains based on the reports associated with the epitopes to the GII.4 VP1. Under the condition, they could describe the changes of antigenicity between the GII.17_2014 and GII.4 genotypes. However, they could not mention the antigenic changes between the GII.17 strains since the low amino acid similarity between the GII.17 and GII.4 VP1 may induce the difference between the epitope positions to these genotypes under the alignments. Therefore, if they discuss the changes of antigenicity between the GII.17 strains, they should cite the reports associated with the epitopes to the GII.17 VP1.
R: Thank you for your question, the information was added. Please see lines 312-319.

(11) Figures 3 and 4.

They should show phylogenies using the maximum likelihood method, which can select the best substitution models in the datasets.
R: The modifications were made, as requested.

Minor comments;

(12) Background.

A polypeptide encoded in the ORF1 can not be determined as the amino acid length of 1738 since the insertion and deletion are induced in the ORF1 sequences of human noroviruses.
R: This information has been removed.

(13) Method, NoV genotyping.

Please correct 'the Genotyping Norovirus Tool' into 'the Norovirus Genotyping Tool'. In addition, please add the date accessed on the GenBank.
R: As recommended the sentence has been corrected. Please see line 151. The date of accessed on the GenBank was also includes, please see line 191-192.

(14) Tables 1 and 2.

Please correct commas in the percent values into dots.
R: As recommended the correction was made on Table 1 and 2.
The current manuscript contains some typing mistakes (no insertions of spaces and unnecessary insertions of spaces etc.). Please carefully check and revise the manuscript.
R: As recommended all text was reviewed.

Juan Degiuseppe (Reviewer 2): The present manuscript describes the emergence and evolutionary analysis of norovirus GII.17 strains detected in the northern region of Brazil.

Although the aspects that authors mention are relevant and represent useful data for understanding the evolutionary dynamic of noroviruses, there are some considerations that must be addressed to improve it. First, there are several misspellings, grammatical errors, and awkward sentence structuring throughout the text. Therefore, language editing by a native English speaker or experienced English editor would benefit the manuscript.
R: As recommended the article was sent to Editagen for English review (Please see annex)

Authors should describe better the epidemiological data from the samples tested: how many samples come from the sporadic panel and how many related to acute gastroenteritis outbreaks, the ages of these patients, if they were hospitalized, and so on. Also, it would be better if authors state if samples were tested for other bacterial or viral enteropathogens (i.e., Group A Rotavirus, Shigella sp) for understanding if NoV GII.17 strains were associated with co-infections or if they are associated with some age group or type of sample (i.e., sporadic vs. outbreak) in particular. In this manner, a full description of the NoV genetic diversity must be addressed to understand the global prevalence of GII.17 strains considering that authors remark its circulation and therefore have analyzed them.
R: As recommended we add more information in the topic “Samples collections”, please see lines 80-89.

It is not clear why authors needed to use a new pair of primers designed for this study. These new primers are used for NoV detection in samples or for amplification and further sequencing of the NoV positive ones? Also, it does not seem that these new primers went under proper validation assays. They were designed taking a strain from Hong Kong as reference when probably it would be more useful to use one of the Brazilian circulating strains. Also, as it is not clear the aim of this new strategy, if they were used for detection, they should be validated with a set of human samples besides the referred controls. If the new set of primers were only designed to amplify a partial genome segment for sequencing purposes, it is not needed to validate the assay. Authors should clarify this aspect.
R: The primers designed in our study (LNOV 5KF/5KR e LNOV 6KF/6KR) specifically amplify NoV GII.17, which were used only for sequencing. There was a need to design new primers as there are no primers for amplifying the complete NoV GII.17_2014 VP1 with optimum efficiency. More information on this was included in the revised manuscript in the lines 95-104.

The primers were used only for sequencing and several parameters were analyzed:
- We sequenced and genotyped GII.17 samples as positive controls;
- We performed RT-PCR with a temperature gradient to verify the best annealing temperature;
- We determine the limit of detection, which measures the lowest virus concentration a primer can identify in a sample;
- The assays were performed by 2 different analysts under the same reaction conditions during the period of 48 h. The objective was to obtain the results to evaluate parameters, such as repeatability, reproducibility, accuracy, sensitivity, and specificity. These parameters were compared using the Cohen Kappa Index.

This information is described under the subsection “Design and validation of NoV primers
GII.17_2014.”
R: We have performed protein modeling of VP1 to verify the mechanism underlying the binding of VP1 and human HBGA. The changes in VP1 protein is reported to affect the binding. We have incorporated more information in the lines 56-59.

Other considerations:
Background section:
It is desirable to avoid the structure description paragraph and describe the NoV genetic diversity (i.e., authors state the characteristics of NoV GII.4 strains when no references on the genetic diversity where mentioned before). In this section, authors should center in the aspects that are relevant for this manuscript.
R: As recommended we add more relevant information to the manuscript.
It is not clear the purpose and the meaning of the statement ‘Whereas most GenBank sequences…’ (page 4, line 21).
R: As recommended the text was reviewed.

Methods section
The Genotyping Norovirus Tool 1.0 is not really updated currently. Authors should consider using https://norovirus.phiresearchlab.org for better results.
R: As recommended we have changed. Please see line 151-152.
The description of the 138 sequences retrieved from GenBank that were used for phylogenetic analysis should be part of the Evolutionary analysis subsection.
R: As recommended we have changed. Please see line 149.
It is not clear why authors conduct protein modeling analysis considering that in the Background section there is no reference about epitopes or how changes could implied an impact in the spreading pattern.
R: We have performed protein modeling of VP1 to verify binding mechanisms to the structure of human HBGA, as it has been shown that changes in this protein can cause changes in this binding. So we add more information about it in the lines 56-60.

Results section
When there is evidence of regional circulation activity, it is not clear why authors do not compare their detected strains with the Brazilian and the Argentine ones.
R: As recommended we added the Brazilian and Argentine strains in the analysis.

Discussion section
Statements already mentioned in the Background section should be eliminated to avoid duplicated information.
R: Modifications were made, as requested.
Limitations are not addressed (i.e. why the number of genotyped strains are below 50%, the results of the validation process and that no clinical samples were used to test this new pair of primers). The strength of every research relies in identifying the weaknesses and the assumptions that were made.
R: As recommended these limitations were added. Please see line 326-329.

Figures

Authors should state when analysis were conducted on partial or complete sequences. Argentine strain is described in the Background and Discussion sections but is not taken into account in the phylogenetic or evolutionary analysis.

R: Modifications were made, as requested