**Checklist:**

Final revision date: Fri 28 Oct.

Point-by-point reply.

1 clean manuscript.

1 tracked changes.

All COI forms.

Check wordcount, figure count.

Check for table in MS word format.

Check figure file labels panels but no additional metatext.

Check correct final corresponding author is Tina.

Check for 40 word summary.

CC: rwhitley@peds.uab.edu

Dear Dr. Lawless,

Your manuscript has been reviewed by our editors and other experts in your field. The reviewers felt that it contains interesting information but in its present form is not acceptable for publication in The Journal of Infectious Diseases; however they did have comments, listed below, that require your consideration in the revision of your manuscript.

Editorial Comments:

As pointed out by the reviewers, the manuscript is complex and difficult to read. They have good suggestions for simplification. As Reviewer #3 points out, we suggest focusing on the virology data (without host GWAS or immunological data), emphasizing the two viral SNPs associated with outcome.

====================

Reviewers' Comments:

Reviewer #1:

This manuscript examines the viral genetic factors that are associated with prolonged RSV infection in children. Using the large INSPIRE cohort of children, the authors identified 19 children that were PCR+ for RSV in the upper respiratory track for 15 days or greater. The did not find any evidence for genetic factors in the children as being associated with the prolonged RSV infection. Examination of the RSV attachment (G) and fusion (F) gene sequences revealed that all of the children exhibiting prolonged RSV infection were infected with the RSV B strain and the authors identified two mutations within the G protein p.E123K/D and p.218T/S/L that were associated with prolonged RSV infection.

Overall, this manuscript is well written and the conclusions are well supported and justified by the data provided.

Specific comments:

1. The analyses of the host response is limited to IFN-g and type I IFN and does not provide much insight. I would suggest removing this section and data that are only reported in the text, as I do not think strong conclusions can be drawn from the data reported.

2. It is curious that all of the prolonged infections were with RSV B strain viruses. The analyses were limited to the G and F sequences (which is a bit surprising as one might expect that the polymerase (L) may also be important and worthy of looking at). Because the F and G proteins may interact, can the authors comment on any similarities noted in the F protein sequences? Were they all the same? were limited mutations found in the F protein in the cohort of 19 children that exhibited prolonged RSV infection compared to the larger cohort that did not?

=====================

Reviewer #2:

RSV is a major cause of respiratory infection in infants. Despite the disease burden, there is a paucity of viral genetic information related to prolonged RSV infection. Herein, Lawless et al using an infant cohort, identified 19 infants with prolonged RSV infection (>15 days between positive PCR tests). The virus (F and G proteins) and host (for previously identified mutations) were sequenced. Given the small cohort, no host traits were identified. However, the authors discovered two RSV strains with mutations within the G protein - p.E123K/D and p/P218T/S/L. These two variant genotypes have been circulating at low levels in the population for at least 30 years, suggesting these are not the result of recurrent mutational events. The authors describe potential functional and host responses; however, these were underwhelming due to the limitations of the study which are described in the discussion.

Comments:

\* Lines 154-155: Please explain why the analysis was focused on the F and G proteins. Was this because the F and G are the 2 major glycoproteins? Is it probable that mutations in the other RSV genes that affect replication and/or host immune modulation might contribute to the prolonged infections?

\* Lines 263 -273: there are several anti-G protein bnmAbs within positions 123-218 which encompass the CCD/CX3C motif (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5226777/, https://www.jimmunol.org/content/183/10/6338). Also, there are also CD4 epitopes within that region as well which should be noted here.

\* Regarding the functional assessment that the mutation within the HBD may promote cell attachment: While much of the literature asserts that RSV binds via surface GAGs, these studies have mostly been performed using immortalized cell lines that contain myriad of GAGs, including HS. However, recent work using primary cells has suggested that 1) RSV infects the apical surface of ciliated respiratory epithelial cells, and 2) these cultures do not have detectable HS on the surface. Please include a comment on the role/function of mutations within the putative HBD if heparin binding is not a major binding event in relevant cell lines. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2644177/, https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1005318, https://pubmed.ncbi.nlm.nih.gov/15613339/ )

\* Were Ct values normalized to input? Please describe how to interpret raw Ct values in this context.

\* If possible, the authors should analyze type III IFN responses that are known to be more relevant in RSV.

====================

Reviewer #3:

With interest I read this paper "Viral genetic determinants of prolonged respiratory syncytial virus infection among infants in a healthy term birth cohort" by Lawless and colleagues. The authors analyzed viral and host factors related to prolonged infection in a cohort of healthy term infants with RTI (n=1949) including 325 with RSV infection (INSPIRE). Whole genome sequencing of RSV was performed in addition to host GWAS and immunological analysis of nose washes. No genetic host factor was identified. Two SNPs in RSV G, which encode a heparin sulfate binding part of the protein.

Comments:

1. This is a large very complex study in which complex mechanisms of RSV disease are studied. My main comment is that the paper does not read well, making it difficult to provide a complete review. Small pieces of results are shown while this requires a major integrated analysis. It might be easier for the reader if only virology data are described (without host GWAS or immunological data) to report the two viral SNPs associated with outcome.

2. Methods are insufficiently described, including the primary endpoint, prolonged infection. The Results section sometimes contains major pieces of background information, such as 4.6 in which RSV G function is explained based on literature without showing research data.

3. There are several issues with the primary endpoint. First, the primary endpoint is not well explained. Second, is prolonged infection a measure of severity? If not, why didn't the researchers study severity such as the need for hospital admission? Is prolonged infection defined as RTI symptoms for more than 15 days? Which symptoms? Would a runny nose be sufficient to qualify? Why did the authors use a dichotomous outcome instead of a continuous outcome (days of symptoms)? Third, only 19 children reached the primary endpoint. What is the power of the study? I feel this study is underpowered for independent associations.

=================

PLEASE NOTE:

With your revision, we also need the following information. Omission of any items will delay the processing of your manuscript.

1. Each author must complete the ICMJE Uniform Disclosure Form for Potential Conflicts of Interest, available at: http://www.icmje.org/conflicts-of-interest/

The corresponding author is responsible for collecting the forms from each author. The completed forms must be uploaded electronically as part of the revision.

2. Major Articles must be no longer than 3,500 words, and Brief Reports no longer than 2,000 words.

3. Insert limits: for Major Articles, 7 in print (tables and figures, with no more than 4 panels per figure), no more than 25 MB of online-only supplementary data; for Brief Reports, 2 in print (tables and figures, with no more than 4 panels per figure), no more than 25 MB of online-only supplementary data.

4. Use Word to make tables. Upload figures in their native file formats, i.e., the format in which the image was originally created. DO NOT EMBED FIGURES IN THE REVISED MANUSCRIPT.

5. If you have a multi-part figure, the journal prefers to receive these as a single file with panels labeled within the image. The figure files should include any labels or markers that are part of the figure itself, but not the figure number, title, legend, or notes; these labels should be provided with figure legend in the manuscript file.

6. Complete contact information for an alternate corresponding author in the event that the corresponding author is unavailable.

7. Source of funding and conflict of interest information, detailed for all authors, should be listed in the acknowledgment section of the manuscript.

8. If any change is made to the author listing, such as a reordering of authors or an addition of an author, an agreement form signed by all authors must be submitted to the editorial office before the revision can be processed. Additionally, please provide an explanation in your revision cover letter regarding the reasoning for changes in authorship.

9. Include on the title page of the manuscript a brief, 40-word-or-less summary of your article's main point. If accepted, this will be published under the article heading in the journal's table of contents.

Thank you for submitting your work to The Journal of Infectious Diseases.

Sincerely,

Martin S. Hirsch, MD

Editor

Richard Whitley, MD

Associate Editor

The Journal of Infectious Diseases

65 Landsdowne Street #412

Cambridge, MA 02139

E-mail: jid@idsociety.org