Reviewers' comments:  
  
Reviewer #1 (Remarks to the Author):  
  
The manuscript by Miller et al with title “Tracking the horizontal transfer of plasmids in Shigella sonnei and Shigella flexneri using phylogenetics” describes the transfer of plasmids among Shigella species in an Australian collection of samples. The study examines the plasmid presence and absence among primarily among S. sonnei and infers the co-evolution of the chromosome and multiple plasmids. Other studies on the rates of evolution among the plasmids and the core chromosome. The core chromosome is not well described in this part of the studies, nor is the impact of missing data, such as the lack of pINV plasmid in the majority of the S. sonnei isolates. Additional observed studies on the potential interspecies transfer of the pKRS100 plasmid among S. sonnei and S. flexneri provides insights into plasmid distribution. Overall, the study is interesting, but strictly observational at this point. While observational studies can be extremely valuable, there are many biases in this study that are either not addressed or not considered. The most significant of those are the lack of the pINV plasmid is the majority of cultured S. sonnei, which is not discussed in this study and can impact these findings significantly. How many isolates lack the plasmid and what can this indicate about the distribution of the plasmid? Additionally, it is not clear, if these plasmids have the direct ability to mobilize and transfer, and some direct measure of these rates would be useful for validation of these studies. The lack of critical details make the interpretation of the data difficult and brings doubt that the conclusions are sound and applicable.  
  
Major Comments  
  
The authors assume that the rates of evolution are the same for the chromosome and each the plasmids (line 101). Additionally, later in the study the authors use the assumption that the metabolic impact is the same for each of the plasmids, which is unlikely and is not supported by the data the authors provide, such as potential rate of transfer (Fig S1). This is a fundamental and fatal flaw of this study. Unless directly examined the assumption that a plasmid that contains antimicrobial resistance genes will be under the same pressures for replication and retention for the chromosome or plasmids that lack resistance markers is unsupported. The inherent lack of the pINV plasmid in the majority of cultured S. sonnei isolates makes these studies incomplete.  
  
An additional issue with the observational nature of the study is the assumption that the isolate being examined is the only isolate in that subject/source. This is being demonstrated to be significantly short sighted as deeper culture studies are being completed and metagenomic studies are being examined in detail. The assumption that the observed isolates are the only isolates and represent the complete diversity within the patient is simplistic. Additionally, Shigella has been demonstrated to not have a binary association with disease, as both the MALeD and GEMS have demonstrated Shigella associated with individuals without clinical disease. This can directly alter the outcomes of this study, both in terms of the observation, as well as the potential sources of the plasmid transfer.  
  
These two points undermine the findings of the purely observational nature of the work as currently presented. The authors need more direct observations of plasmid evolution, transfer rates among species and for some plasmids evidence that mobilization into other species is possible.  
  
Figure 1 seems to attempt to determine directionality of plasmid transfer. This observational aspect of the study is unsupported without greater detail in terms of how many isolates contain the plasmid. (i.e. the resolution of the current study is lacking to support the conclusions that authors draw). Additionally, there does not appear to be evidence in the literature that has examined plasmid transfer of these molecules (do the each contain transfer and mobilization genes? How are they moving – the same question can be made for discussion of pKSR100)  
  
The discussion of pKSR100 has subsampling, which is not well described. For the S. sonnei, this is not much of an issue as many studies have demonstrated the monomorphic nature of the chromosome; however, the sampling of S. flexneri is not simple or straightforward with multiple serogroups and phylogenomic groups (as demonstrated by Bengtsson/Baker et al 2022). Application of a single rate across the genomic variation of this species would likely lead to incorrect conclusions. The subsampling methods must be described in detail and the genomic variability of S. flexneri considered (for example to certain phylogroups transfer more readily? Were all phylogroups represented equally? Etc).  
  
There is an observation bias in this study that is not recognized and needs to be addressed. For example, in the discussion of pKSR100 the authors suggest a rapid expansion of certain clades that acquired this plasmid. This is somewhat out of context without knowing the other selective pressures such as treatment with antibiotics and patient comorbidities, coinfections and other clinical features. Additionally, as with the other plasmids, how is this plasmid determined to be the same? It is doubtful that the pKRS100 plasmid is identical in all species and isolates, so the authors should identify how much variation would be acceptable to still call the plasmid the same? Is the replicon enough, as it is for many plasmid groups, or are there plasmid features that are required, or is it simply a nucleotide similarity, and what is that threshold.  
  
Minor Comments  
  
The authors indicate that the pINV is considered part of the Shigella species core genome (line 126). Does that statement only consider the conserved core components of the plasmid, as it is clear in all genomic studies where the plasmid has been examined that pINV is not one single entity and variability exists.  
  
The resolution of the figures makes them different to interpret. Each should be larger and labels should also be larger so a reader can clearly understand what is going on.  
  
The title of the paper is misleading as it includes both S. sonnei and S. flexneri, but the majority of the analysis has been focused on S. sonnei and S. flexneri is an afterthought with the exception of the pKSR100 study, and the inclusion of S. flexneri in that section is problematic.  
  
  
  
Reviewer #2 (Remarks to the Author):  
  
Title: Tracking the horizontal transfer of plasmids in Shigella sonnei and Shigella flexneri using phylogenetics  
  
  
Overall comments  
  
This manuscript describes a method that reconstructs chromosome-plasmid associative histories through a Bayesian phylogenetic analysis incorporating the horizontal acquisition of plasmids independent of the chromosome, but crucially, does not ignore chromosomal evolution during the same period. This analysis utilizing co-divergence appears to improve our understanding of plasmid evolutionary rates over time, an improvement in understanding the evolution of antibiotic resistance since plasmids remain an important source of resistance gene transfer.  
  
There are significant issues that prevent me from recommending acceptance of this paper at this time, outlined below, particularly because they hinder interpretation. The figures are inconsistently labeled, poorly formatted, and is lacking critical information. Significant revision of the results and resubmission may bring it to the editorial standard of Nature Communications, or authors may consider an updated submission to a journal with a more technical focus. In particular, expansion of results regarding pKSR100 would be timely given the current rise in XDR Shigella affecting vulnerable populations. Alternatively, the article is about drivers of antibiotic resistance but does not detail the antibiotic resistance genes associated with these strains; some demonstration of how these improved co-divergent phylogenies provide clarity to the dissemination of antibiotic resistance, for example macrolide resistance gene mphA, would be timely.  
  
Regarding impact, the authors do not adequately discuss what is novel or unexpected about their findings. The discussion correctly asserts that this method and the questions it addresses are important for understanding antibiotic resistance, but fail to discuss this in the context of the phylogenetic reconstructions found in the paper. Some discussion of the rise in plasmid abundance relative to the timing of the observed rise of MSM-associated pKSR100 plasmids, for example, could provide some validation of results that is currently lacking. In particular, while this interspecies communication of the plasmid is known to occur, the frequency of this transfer being observed is not, particularly over time and regarding the overall frequency of pKSR100 observation.  
  
Specific comments  
  
Title  
• The method here is more about a reconstruction than tracking. It infers plasmid transfer without using historical isolates.  
  
Abstract  
• The abstract is overall clear and well constructed.  
• Here, and in several places in the manuscript, the authors state their method tracks “how plasmids move between bacterial lineages.” The plasmids move by transconjugation, more specific language here is warranted since the method infers or predicts intra- and inter-species transfer events rather than the physical method.  
• The abstract should mention CoalPT.  
  
Introduction  
• The introduction is generally well constructed and written, but some areas would benefit from improved organization and focus.  
• Line 36: “The spread of drug-resistant plasmids within a bacterial population increases the chance of AMR genes disseminating to other bacterial species in the same ecological niche.” Do the citations from the previous sentence carry over here? If so, I think this is fine; if not, it should be softened to be “thought to” or cited separately.  
• Line 39: This paragraph has the thesis statement, “Shigella [spp.] are are key exemplar pathogen to develop new methodological approaches to study the movement of plasmids between bacterial lineages.” While I agree with the thesis, the paragraph does not mention which of the plasmids are known to have transferred between species or subtypes, including pKSR100. The authors should consider strengthening their case for using Shigella spp. as their organism of choice with some information on why we expect interesting plasmid transfer dynamics on a species level here.  
  
Results  
• The results need considerable improvement to be considered for publication. Additionally, I have concerns that the model suffers from considerable underestimation of variance in historical predictions. Figures are inconsistently or improperly labeled. There are key moments in the paper where additional measurements are necessary. Some figures are not cited in the paper.  
  
• Section beginning at line 97 and Figure 1:  
o Much of the text here is too small to be legible.  
o The numbers at the bottom at Figures A, B, and C are assumed to be years, but it is never explicitly stated.  
o Figures 1D, E, and F have text too small to be readily interpreted.  
o Figure 1D is not referenced in text.  
o Are the colors in A, B, C, D, E, and F are shared across the figure?  
o It appears from Figure 1 that spB is not frequently transferred, which does not match analysis in the text.  
o spA is predicted to have the highest transfer rate, how confident is this prediction? Is there a suitable statistical test for the predicted distribution vs spC?  
o How many tips are present on the phylogenetic trees? Do tips represent sets of strains? It does not appear to be the 789 isolates mentioned in the materials.  
o High variability is observed between 2016 and 2020 for the isolates, yet in Figure 1E, predictions before 2015, it assumes low variability and a general smooth trend for these plasmids that does not account for the probability afterwards.  
- It appears that the model may not account for variability in observed data well since the year to year variance observed is high but the predicted variance is low.  
o There are some clades/branches that appear unrelated to the others here. How much of the higher branches of the tree are missing?  
- One solution is to use the typing scheme here to identify which strains of S. sonnei are being analyzed: Hawkey, et al. 2021. Nature Communications. Global population structure and genotyping framework for genomic surveillance of the major dysentery pathogen, Shigella sonnei. [https://www.nature.com/articles/s41467-021-22700-4](https://protect-au.mimecast.com/s/OIlLC81ZRASjBJJkLinHM0t?domain=nature.com)  
o There are some branches with no circle, but still showing plasmid transfer events (see upper left of Fig 1A). Are these predicted S. sonnei strains that are predicted to have existed, but for whom there are no extant descendants?  
o Removing the grey lines from Figures 1ABC and Figure S2 would be clearer regarding the individual histories of the plasmids. Currently, there are many distracting lines in each figure, plus, since the base dendrogram of S. sonnei is the same in each case helpful year markers, they are easily compared.  
o There is a brief period in Figure 1B where one of the plasmids spends time in the lower half-clade, of which no extant branches have the plasmid. What accounts for this behavior in the model?  
  
• Section beginning on Line 128 and Figure 2  
o For Figures 2A and B, please provide specific units for y axis values, such as “point mutations per year and site.”  
o What do the error bars represent in Fig 2A?  
o Consider re-labeling “Network” as “CoalPT” since it is the output of the model.  
o Provide a test of significance for the differences in plasmid mutation rate means and chromosome.  
o For Figures 2 CDE, are the X-axes values years?  
o It is not immediately obvious the density trees have been improved for spA and spC, is a metric available for scoring these?  
o Figures S3, S4 and S5 use these simulated datasets to validate the model’s performance for population size and plasmid transfer rate, please provide a test of how well the model performed for S4 and S5. S3 seems appropriate.  
  
• Section beginning on line 170 and Figure 3  
o Figure 3C is referenced after Figure 3D  
o Which species are represented in Figure 3A and B, and where? It is impossible to interpret in its current form.  
o The higher order nodes are again cut off, but cannot be directly compared in this instance.  
o It seems that posterior support for events above 0.5 are shown as in Figure 1, but it is not stated in the legend.  
o Figure 3B appears to be a truncated phylogenetic tree to show plasmid relatedness. Are the descending branches independent? Is the figure as presented a minimum number of rearrangements? Without having any idea of how these are rooted, it’s impossible to interpret the relative differences in topology.  
  
• Discussion  
o The variability observed in the five-year period versus the relatively tight confidence before this (Fig 1E) requires some comment either stating it is real or requires improvement, or addressing potential overconfidence in results from the five year sampling period.  
- The modeling suggests that the sampled period has little variance in observed results though there is considerable variability within the sampled five-year period.  
o More information discussing the predictions made by the model and what they mean regarding S. sonnei and S. flexneri is needed. For example, the first mphA Shigellae were found in samples dating back to the 1980s, but the authors only see a rise in contribution of pKSR100 since 2010. This predates Baker et al. 2015 Lancet by 5 years, and suggests that the original report was timely in detection, less a neglected problem.  
o While they predict transfer between species, such assessments are not new or novel. Assessments of their timing and frequency are new, but these are not discussed, particularly with the pKSR100-driven outbreak across bacterial species.  
- For instance, what is the likelihood of a new strain appearing with pKSR100 from a previously non-carrying isolate, and from which species will it originate?  
  
• Materials and Methods  
o The equations and their implementation appear to be appropriately documented.  
o Notes for additional information has been included in the sections above.  
o The reviewer did not complete an exhaustive review of the algorithm and its implementation, but they appear to be documented and organized sufficiently.  
  
  
  
Concluding remarks  
  
While this manuscript may mark a considerable improvement to our ability to reconstruct plasmid evolutionary histories, there are significant problems in the presentation and discussion of results that prevent recommendation for publication. Further, the behavior of the model prior to the observed period requires some addressing in text if the authors are confident in the confidence intervals. The manuscript also uses data from a current outbreak as part of their results, but does not discuss how the output of their model is relevant to understanding that outbreak.