General comments

The paper "The Nextflow pipeline nf-core/metadenovo for reproducible annotation of metatranscriptomes and more" presents a new pipeline for de novo assembly and annotation of metatranscriptomes. While some alternatives do exist, they seem to be not maintained. In the view of rapid interest growth in metatranscriptomics in recent years, this pipeline promises to be a valuable addition to the scientific community and will allow bioinformaticians to easily generate reproducible results without a need to manually run multiple different tools.

The pipeline is available, documented and easy to launch. Thus, it will be a useful instrument even for researchers without advanced computational skills. However, I encountered a couple problems during the run (see below).

The manuscript itself is well-structured and fairly easy to follow. The authors demonstrate different aspects of the developed pipeline using 3 completely different datasets and compare their results to the original studies. Seeing these examples and use-cases can be quite interesting for readers doing similar studies. The text is supported by a substantial amount of graphical material. In my opinion, some figures can be moved to the supplementary material and the quality of some plots can be improved (see below). Below I provide a few suggestions that in my opinion may improve this manuscript.

Major comments

1. I tried to run the pipeline several times on different machines, but stumbled across a few problems:
    a. GitHub and nf-core pages provide different input CSV file structure, only one worked for me in the end.
    b. During my first run, the pipeline refused to work with .fastq files, telling me they need to be gzipped. I think this should be easily fixable.
    c. Finally I encountered an error saying "ERROR ~ No such file or directory: /XXX/YYY/eukulele" (the pipeline was run in /XXX/YYY/). I cannot say whether it's my problem or not, but I expected an nf-core pipeline to run smoothly without any prerequisites.

   Thus, I recommend performing a bit more testing in different environments. Usability is the key aspect in tool-based publications.

Minor comments

1. I suggest doing one more proofreading as I stumbled across typos, or, for example, inconsistencies in using small and capital letters in the assembler names.

2. The authors describe existing pipelines in the Discussion section. I think it would be useful for the reader to give this information in the Introduction.
3. In the “Dataset descriptions” subsection the authors describe how these datasets were analysed, I think it should be moved to somewhere else within the Methods section.
4. One of the possible assemblers the pipeline offers is Megahit. Although Megahit is undoubtedly a great assembler, to the best of my knowledge it was mostly tested on metagenomic and genomic data (as stated on their GitHub page). While in this work it seems to show decent results, I wonder whether there were any benchmarks made previously for metatranscriptome assemblies using Megahit? In any case, some justification for using Megahit would add value to the paper.
5. If I’m not mistaken, the authors mention their pipeline is capable of working with metagenomic data as well (at least documentation says so). However, there is barely any information in the manuscript of how metagenomic reads can be processed and whether a user can combine metaT and metaG data.
6. As long-read technologies gain popularity, I think it might be useful to add support for long-read data (e.g. rnaSPAdes supports hybrid mode). I am not sure whether these technologies are used a lot specifically for metatranscriptome sequencing, it can be a nice addition for the future.
7. Fig. 3A color scheme can be more colorbl-friendly. While I understand the struggle of selecting multiple colors for a bar-plot, some predefined schemes exist in various plotting libraries.
8. It can be quite useful for a potential user to provide a plot of RAM consumption over the time (maybe just a single example).
9. I don’t really understand what the X axis means on Fig. 4A from its description (and whether there should be any label).
10. Fig 4b (and following). I don’t get why the authors add “…”. Also, meaning of * becomes clear only from the caption, thus making this figure a bit less self-explanatory. I’d suggest adding this detail right to the figure.
11. Coloring of Fig. 7C is extremely flashy and should be changed (also not colorblind friendly). I suggest using some gradient scheme to indicate expression changes.
12. The authors might consider moving some plots to the supplementary material (e.g. Fig. 5) and keeping only some examples (will leave this decision to the authors).
13. I recommend adding a supplementary table with all the software versions used for the analysis.
14. Citations are clickable but do not work (lead to Zotero instead of the reference list).
15. Of note, if rnaSPAdes was used, I believe it should be cited with Bushmanova et al., 2019 rather than a broader Prjibelski et al., 2020 (will leave up to the authors).