eLIFE
1st Floor
24 Hills Road
Cambridge CB2 1JP, UK
P 01223 855340
W elifesciences.org
T @elife

eLife's transparent reporting form

We encourage authors to provide detailed information *within their submission* to facilitate the interpretation and replication of experiments. Authors can upload supporting documentation to indicate the use of appropriate reporting guidelines for health-related research (see EQUATOR Network), life science research (see the BioSharing Information Resource), or the ARRIVE guidelines for reporting work involving animal research. Where applicable, authors should refer to any relevant reporting standards documents in this form.

If you have any questions, please consult our Journal Policies and/or contact us: editorial@elifesciences.org.

Sample-size estimation

- You should state whether an appropriate sample size was computed when the study was being designed
- You should state the statistical method of sample size computation and any required assumptions
- If no explicit power analysis was used, you should describe how you decided what sample (replicate) size (number) to use

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn't apply to your submission:

Our article is a Scientific Correspondence on (Mast 2020). The authors of Mast et al determined the sample sizes by choosing to include the data from Zhou et al. (N=9) and Michalovich et al. (N=40). We describe significant concerns regarding the composition of these sample sets in our correspondence. We performed our re-analysis of their data using all samples, as was described in their materials and methods/transparency reporting.

Replicates

- You should report how often each experiment was performed
- You should include a definition of biological versus technical replication
- The data obtained should be provided and sufficient information should be provided to indicate the number of independent biological and/or technical replicates
- If you encountered any outliers, you should describe how these were handled
- Criteria for exclusion/inclusion of data should be clearly stated
- High-throughput sequence data should be uploaded before submission, with a private link for reviewers provided (these are available from both GEO and ArrayExpress)

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn't apply to your submission:

The individual samples from Zhou 2020 and Michalovich 2019 each came from a unique individual patient or donor and can be accessed through the NCBI Sequence Read Archive. This means that samples in the same groups can be considered biological replicates. Similarly to the analysis performed in Mast et al., we used all case (N=9) and control (N=40) samples for this study.

Statistical reporting

- Statistical analysis methods should be described and justified
- Raw data should be presented in figures whenever informative to do so (typically when N per group is less than 10)
- For each experiment, you should identify the statistical tests used, exact values of N, definitions of center, methods of multiple test correction, and dispersion and precision measures (e.g., mean, median, SD, SEM, confidence intervals; and, for the major substantive results, a measure of effect size (e.g., Pearson's r, Cohen's d)
- Report exact p-values wherever possible alongside the summary statistics and 95% confidence intervals. These should be reported for all key questions and not only when the p-value is less than 0.05.

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn't apply to your submission:

The RNA-sequencing data processing including quality control, trimming, and alignment was performed as described by Mast et al using CLC Genomics Workbench. Differential expression analysis was performed using EdgeR according to the R-script supplied to us by the authors of Mast et al.. All alignment quality metrics and statistics produced by CLC genomics workbench are included as supplemental file (one per sample). The R script used for differential expression analysis is also supplied as a supplemental file.

(For large datasets, or papers with a very large number of statistical tests, you may upload a single table file with tests, Ns, etc., with reference to sections in the manuscript.)

Group allocation

- Indicate how samples were allocated into experimental groups (in the case of clinical studies, please specify allocation to treatment method); if randomization was used, please also state if restricted randomization was applied
- Indicate if masking was used during group allocation, data collection and/or data analysis

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn't apply to your submission:

Allocation of samples into groups was based on their COVID-19 positive or negative status, with all COVID-19 positive samples sourced from Zhou et al. and all control samples sourced from Michalovich et al.. We discuss significant issues regarding the experimental designation of control samples as "Healthy Controls" in the text.

Additional data files ("source data")

- We encourage you to upload relevant additional data files, such as numerical data that are represented as a graph in a figure, or as a summary table
- Where provided, these should be in the most useful format, and they can be uploaded as "Source data" files linked to a main figure or table
- Include model definition files including the full list of parameters used
- Include code used for data analysis (e.g., R, MatLab)
- Avoid stating that data files are "available upon request"

Please indicate the figures or tables for which source data files have been provided:

Additionally, unmodified count tables output from CLC genomics workbench are included in text format. We also supply several sorted count tables that include sample type information rather than SRA identifiers in excel format. These excel files include raw count data for the whole transcriptome, for genes of interest discussed in Mast et al., and for relevant rRNA genes.