


CORRESPONDENCE

Open Access



# Response to Brinkmann et al. “Re-assembly of 19th century smallpox vaccine genomes reveals the contemporaneous use of horsepox and horsepox-related viruses in the United States”

Ana T. Duggan<sup>1\*</sup> , Edward C. Holmes<sup>2</sup> and Hendrik N. Poinar<sup>1,3</sup>

This reply refers to the comment available at <https://doi.org/10.1186/s13059-020-02202-0>.

\* Correspondence: [duggana@mcmaster.ca](mailto:duggana@mcmaster.ca)

<sup>1</sup>McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, Hamilton L8S 4L9, Canada

Full list of author information is available at the end of the article

## Abstract

We thank Brinkmann and colleagues for their correspondence and their further investigation into these American Civil War Era vaccination strains. Here, we summarize the difficulties and caveats of work with ancient DNA.

## To the Editor:

We are delighted to see that the broad phylogenetic findings and interpretation as presented in our work “The origins and genomic diversity of American Civil War Era smallpox vaccine strains” have been reproduced by Brinkmann and colleagues (Brinkmann A, Souza ARV, Esparza J, Nitsche A, Damaso CR: Re-assembly of 19th-century smallpox vaccine genomes reveals the contemporaneous use of horsepox and horsepox-related viruses in the United States, in preparation) [1]. While the libraries we generated from the vaccination kits likely contain DNA fragments representing the entire genomic content, we question the ability to faithfully reconstruct these termini given the following, well-characterized aspects of ancient DNA (aDNA): (i) The DNA is heavily truncated to median fragments sizes less than 55 bp, (ii) is mixed with the DNA of many different species, and (iii) has damaged nucleotides, features which all complicate mapping and de novo assemblies [2, 3]. These factors exacerbate the fundamental difficulties of re-assembling genomic repetitive elements and regions [4], thereby creating de novo assemblies which are both considerably shorter than full genome length and prone to spurious contigs. We note that the starting material for these libraries were not purified vaccine materials, but rather scabrous material and old tin boxes. The resulting extracts produced complex metagenomic libraries, further increasing the likelihood of generating of chimeric contigs and chimeric genome



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

reassemblies. Furthermore, our libraries were biased by multiple sequencing strategies—including both shotgun and targeted enrichment—particularly important when considering novel insertions such as those proposed by Brinkmann et al. Assembly algorithms and software specifically designed for the challenges of aDNA are scarce, as such researchers typically err on the side of caution. Methodologies involving full or partial manual assemblies are avoided as they are necessarily subjective and therefore difficult to reproduce. On that final point, we note that while the increased length and diversity of terminal repetitive regions identified by Brinkmann et al. could be very interesting, within the sphere of aDNA research increased diversity and unexpected length variation are also red flags for chimeric sequence reconstruction.

#### Authors' contributions

The authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, Hamilton L8S 4L9, Canada. <sup>2</sup>Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Life and Environmental Sciences and School of Medical Sciences, University of Sydney, Sydney, NSW 2006, Australia. <sup>3</sup>M.G. DeGroot Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, DeGroot School of Medicine, McMaster University, Hamilton L8S 4L9, Canada.

Received: 28 October 2020 Accepted: 11 November 2020

Published online: 04 December 2020

#### References

1. Duggan AT, Klunk J, Porter AF, Dhody AN, Hicks R, Smith GL, Humphreys M, McCollum AM, Davidson WB, Wilkins K, Li Y, Burke A, Polasky H, Flanders L, Poinar D, Raphenya AR, Lau TTY, Alcock B, McArthur AG, Golding GB, Holmes EC, Poinar HN. The origins and genomic diversity of American Civil War Era smallpox vaccine strains. *Genome Biol.* 2020;21(175) <https://doi.org/10.1186/s13059-020-02079-z>.
2. Duchêne S, Ho SYW, Carmichael AG, Holmes EC, Poinar H. The recovery, interpretation and use of ancient pathogen genomes. *Curr Biol.* 2020;30(19):PR1215–31 <https://doi.org/10.1016/j.cub.2020.08.081>.
3. Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M. Genetic analyses from ancient DNA. *Annu Rev Genet.* 2004;38:645–79.
4. Tørresen OK, Star B, Mier P, Andrade-Navarro MA, Bateman A, Jarnot P, Gruca A, Grynberg M, Kajava AV, Promponas VJ, Anisimova M, Jakobsen KS, Linke D. Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic Acids Res.* 2019;47(21):10994–1006 <https://doi.org/10.1093/nar/gkz841>.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

