

Comment

Proper Authentication of Ancient DNA Is Still Essential

Raphael Eisenhofer *  and Laura S. Weyrich

Department of Genetics & Evolution, Darling Building, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia; laura.weyrich@gmail.com

* Correspondence: raphael.eisenhoferphilipona@adelaide.edu.au; Tel.: +61-4-3582-4262

Received: 31 January 2018; Accepted: 19 February 2018; Published: 26 February 2018

Keywords: ancient DNA; microbiome; paleomicrobiology; microbial ecology

Santiago-Rodriguez et al. [1] report on the putative gut microbiome and resistome of Inca and Italian mummies, and find that Italian mummies exhibit higher bacterial diversity compared to the Inca mummies. However, contaminant taxa in their negative control account for most of the biological signal observed. In addition, they fail to properly apply field-standard ancient DNA authentication techniques to their data and self-plagiarize a previously published figure. Poor standards in paleomicrobiological research are currently plaguing the field, despite numerous warnings [2–4] and reviews [5–8] on best practice.

DNA contamination from museums, curators, scientists, soil, and even the laboratory can drive signals present in modern and ancient metagenomics data sets [3,7,9–16]. Therefore, explicit rules and standards to avoid falsely reporting contaminants in metagenomics datasets have been put forth [3,4,6–8,17]. These standards typically include sampling and extraction blank controls (e.g., tubes processed without the addition of biological samples) to monitor contaminant DNA and correctly attribute its contribution in subsequent analyses. In the study by Santiago-Rodriguez et al. [1], a non-template or blank control was included in their 16S analysis. However, the authors failed to explore the contaminant species within this control during their analysis of differences between Incan and Italian mummies. We explored the taxa present within their blank control (Supplementary_Dataset_2.txt from their publication) and compared it to those identified within the mummies. We found that laboratory contaminants present within their blank control are driving the differences between Incan and Italian mummies (Figure 1). For example, the five most abundant taxa identified in the Italian mummies (*Sphingomonadales*, *Pseudomonadales*, *Rhizobiales*, *Bacillales*, and *Clostridiales* species) are all found in the blank control. It is also worth noting that these taxa have previously been identified as common laboratory or reagent contaminants in numerous studies [3,14,15]. This strongly suggests that the cultural differences reported by the authors are likely the result of laboratory contamination and calls into question the validity of their subsequent analyses.

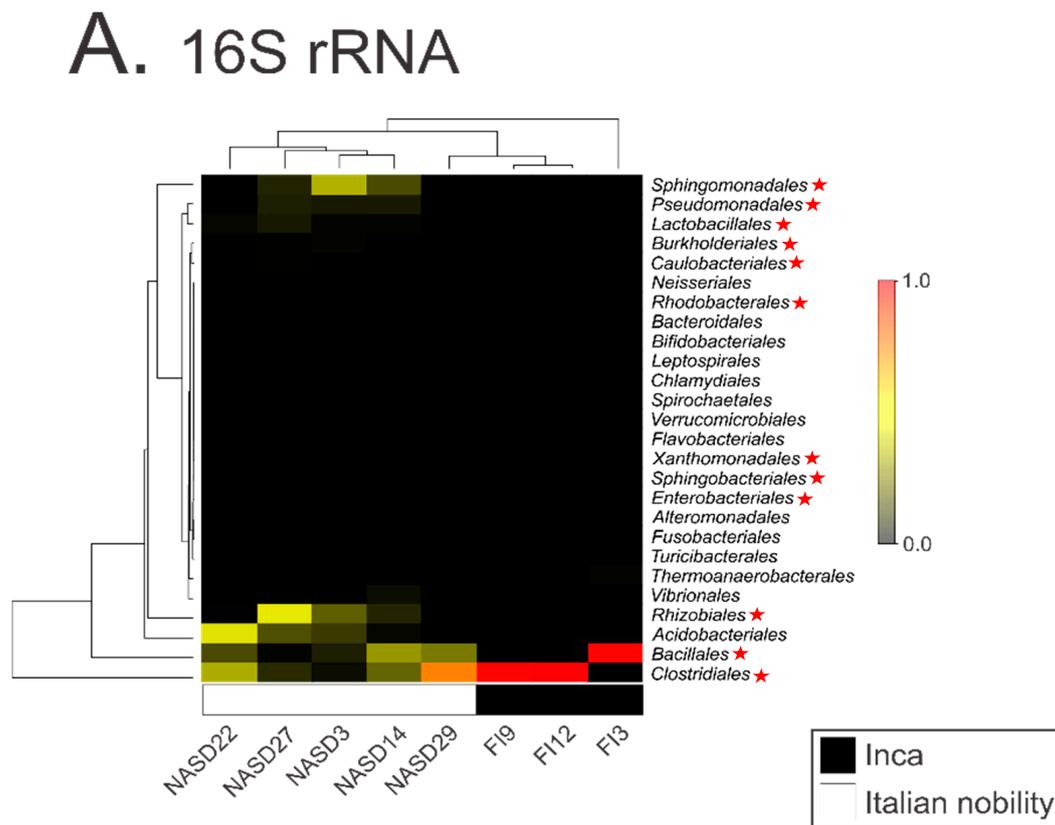


Figure 1. An altered reproduction of Figure 1A from Santiago-Rodriguez et al. [1] where taxa identified in the 16S rRNA blank control are identified in the mummy samples by red stars. The highest-abundance taxa identified in the 16S rRNA data are also present in the 16S rRNA blank control.

The authors then attempt to use MapDamage to assess the authenticity of their shotgun metagenomic ancient DNA; this tool is widely used within the paleomicrobiological field for detecting patterns of cytosine deamination that are characteristic of authentic ancient DNA [18]. Critically, the authors did not provide details as to how they ran the analysis; MapDamage calculates the deamination rate by comparing a reference genome to the mapped target sequences present in a given biological sample (i.e., the reference and target species are typically reported for the analysis). Despite this lack of information, the MapDamage plot provided by the authors in their supplementary information (Figure 2A) is identical to one in a previous publication by the team [4] (Figure 2B), suggesting that the authors self-plagiarized this figure and did not in fact run the analysis.

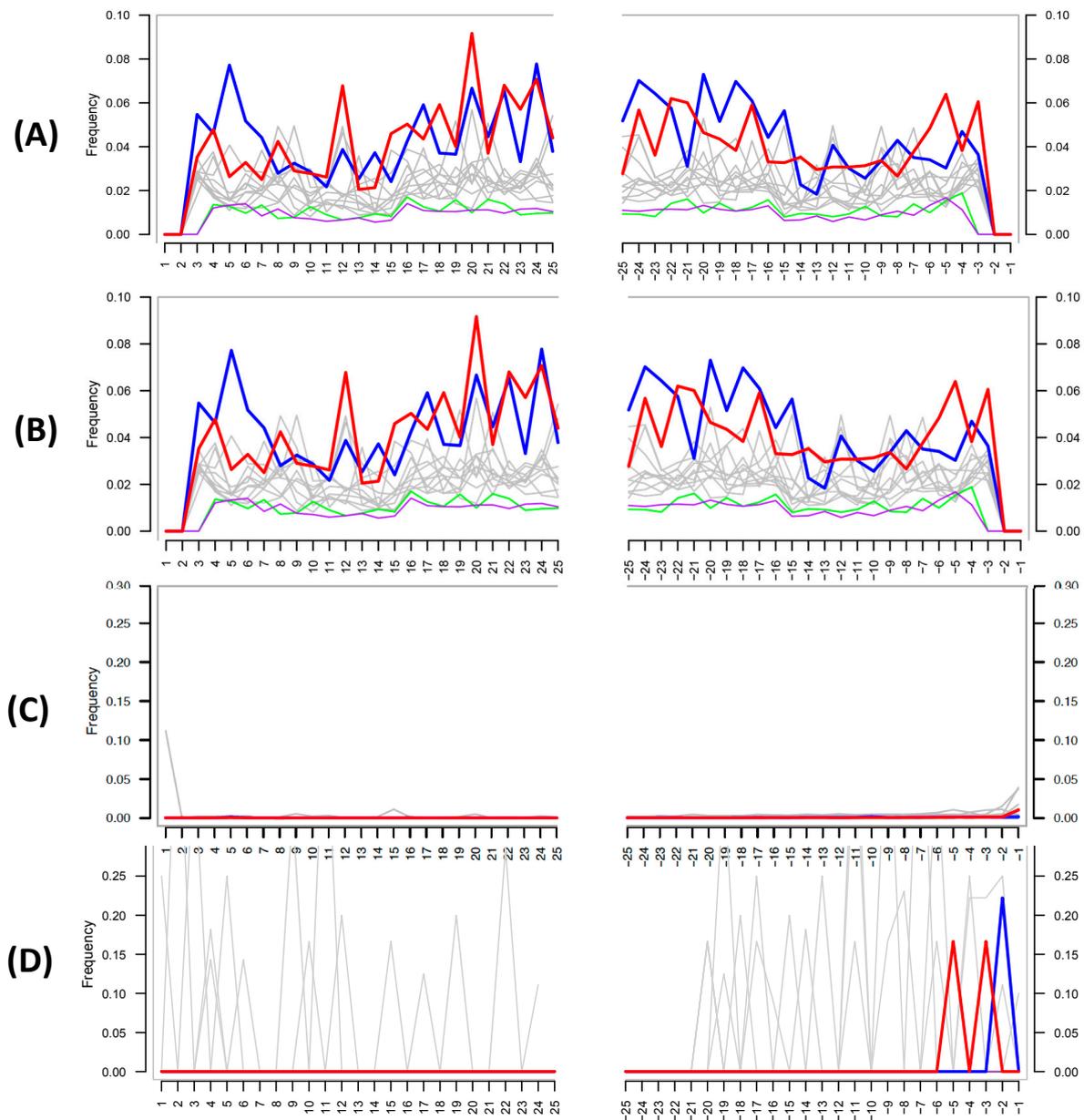


Figure 2. (A) MapDamage plot provided by the authors in their latest paper [1]. (B) MapDamage plot provided by authors in their previous publication [19]. Both plots are identical, and both show the absence of damage characteristic of authentic ancient DNA. (C) MapDamage plot obtained by using reads aligned from Italian mummy NASD14 from Santiago-Rodriguez et al. [1] against the *Sphingomonas sp. DC-6* genome (ASM71517v2). (D) Same as (C), except using *Vibrio parahaemolyticus* (ASM19609v1), a taxon not found in the authors' negative control. The lack of nucleotide misincorporation is indicative of modern DNA.

Despite this, the figure provided also does not support the authenticity of ancient DNA, as the expected C to T at the 5' and G to A substitutions at 3' ends of the DNA fragments are not present. The authors defend their lack of authentic ancient DNA signal by stating: "Damage-based ancient DNA authentication tools, such as mapDamage, may be incompatible with ancient microbiome studies unless a high sequencing coverage is reached". However, simulations and empirical data show that only a few thousand sequences from the genome of interest are required to assess the presence of cytosine deamination [6], and MapDamage has been successfully applied in several

paleomicrobiological studies [20–23]. To investigate if MapDamage could be appropriately applied to the Santiago-Rodriguez et al. data set [1], we downloaded the metagenomic reads from a mummy present within their study (NASD14) and identified species present in the sample using MALT and MEGAN [24,25] against a reference database containing >50-thousand bacterial and archaeal genomes obtained from NCBI Assembly. Similar to the authors' shotgun results, we identified ≈1.2 million reads assigned to *Sphingomonas* sp. DC-6 (Sphingomonadales), and 33,730 reads assigned to *Vibrio parahaemolyticus* (Vibrionales). We then mapped the metagenomic reads against these reference genomes with the BWA-backtrack (ALN) aligner [26]. The outputs were converted into SAM files then used as input for MapDamage, comparing the "ancient" *Sphingomonas* and *Vibrio* species to their respective reference genome. The resulting plots (Figure 2C,D) clearly illustrate no characteristic ancient DNA damage and are as expected for modern, likely contaminant, DNA. Given that *Sphingomonas* is one of the most abundant taxa in their data and is a known contaminant species, our reanalysis further strengthens the likelihood that contaminant DNA is driving their findings.

To conclude, a reanalysis of Santiago-Rodriguez et al.'s [1] findings strongly suggest that the observed signal is due to laboratory contamination of modern bacterial species. The authors also failed to compare their data to their extraction blanks controls, did not include shotgun metagenomic extraction blanks, and did not authenticate their ancient DNA using MapDamage. Paleomicrobiology is a new and rapidly growing field of research, with little room for plagiarized figures and blatant disregard for best-practice methods. In light of these findings, we suggest either heavy corrections or retraction of the article to prevent further erosion of the scientific integrity of paleomicrobiological research.

Author Contributions: R.E. analyzed the data, created the figures, and wrote the manuscript. L.S.W. provided feedback on the manuscript.

Conflicts of Interest: The authors work in the field of paleomicrobiology and want science published in this space to be of sufficient scientific rigor.

References

1. Santiago-Rodriguez, T.M.; Fornaciari, G.; Luciani, S.; Toranzos, G.A.; Marota, I.; Giuffra, V.; Cano, R.J. Gut Microbiome and Putative Resistome of Inca and Italian Nobility Mummies. *Genes* **2017**, *8*, 310. [[CrossRef](#)] [[PubMed](#)]
2. Weyrich, L.S.; Llamas, B.; Cooper, A. Reply to Santiago-Rodriguez et al.: Was luxS really isolated from 25- to 40-million-year-old bacteria? *FEMS Microbiol. Lett.* **2014**, *353*, 85–86. [[CrossRef](#)] [[PubMed](#)]
3. Salter, S.J.; Cox, M.J.; Turek, E.M.; Calus, S.T.; Cookson, W.O.; Moffatt, M.F.; Turner, P.; Parkhill, J.; Loman, N.J.; Walker, A.W. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **2014**, *12*, 87. [[CrossRef](#)] [[PubMed](#)]
4. Eisenhofer, R.; Cooper, A.; Weyrich, L.S. Reply to Santiago-Rodriguez et al.: Proper authentication of ancient DNA is essential. *FEMS Microbiol. Ecol.* **2017**, *93*. [[CrossRef](#)] [[PubMed](#)]
5. Cooper, A.; Poinar, H.N. Ancient DNA: Do It Right or Not at All. *Science* **2000**, *289*, 1139. [[CrossRef](#)] [[PubMed](#)]
6. Warinner, C.; Herbig, A.; Mann, A.; Yates, J.A.F.; Weiß, C.L.; Burbano, H.A.; Orlando, L.; Krause, J. A Robust Framework for Microbial Archaeology. *Annu. Rev. Genom. Hum. Genet.* **2017**, *18*, 321–356. [[CrossRef](#)] [[PubMed](#)]
7. Llamas, B.; Valverde, G.; Fehren-Schmitz, L.; Weyrich, L.S.; Cooper, A.; Haak, W. From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR Sci. Technol. Archaeol. Res.* **2017**, *3*, 1–14. [[CrossRef](#)]
8. Kim, D.; Hofstaedter, C.E.; Zhao, C.; Mattei, L.; Tanes, C.; Clarke, E.; Lauder, A.; Sherrill-Mix, S.; Chehoud, C.; Kelsen, J.; et al. Optimizing methods and dodging pitfalls in microbiome research. *Microbiome* **2017**, *5*, 52. [[CrossRef](#)] [[PubMed](#)]
9. Shen, H.; Rogelj, S.; Kieft, T.L. Sensitive, real-time PCR detects low-levels of contamination by *Legionella pneumophila* in commercial reagents. *Mol. Cell. Probes* **2006**, *20*, 147–153. [[CrossRef](#)] [[PubMed](#)]

10. Witt, N.; Rodger, G.; Vandesompele, J.; Benes, V.; Zumla, A.; Rook, G.A.; Huggett, J.F. An Assessment of Air As a Source of DNA Contamination Encountered When Performing PCR. *J. Biomol. Tech.* **2009**, *20*, 236–240. [[PubMed](#)]
11. Naccache, S.N.; Greninger, A.L.; Lee, D.; Coffey, L.L.; Phan, T.; Rein-Weston, A.; Aronsohn, A.; Hackett, J.; Delwart, E.L.; Chiu, C.Y. The Perils of Pathogen Discovery: Origin of a Novel Parvovirus-Like Hybrid Genome Traced to Nucleic Acid Extraction Spin Columns. *J. Virol.* **2013**, *87*, 11966–11977. [[CrossRef](#)] [[PubMed](#)]
12. Lusk, R.W. Diverse and Widespread Contamination Evident in the Unmapped Depths of High Throughput Sequencing Data. *PLoS ONE* **2014**, *9*, e110808. [[CrossRef](#)] [[PubMed](#)]
13. Adams, R.I.; Bateman, A.C.; Bik, H.M.; Meadow, J.F. Microbiota of the indoor environment: a meta-analysis. *Microbiome* **2015**, *3*, 49. [[CrossRef](#)] [[PubMed](#)]
14. Lauder, A.P.; Roche, A.M.; Sherrill-Mix, S.; Bailey, A.; Laughlin, A.L.; Bittinger, K.; Leite, R.; Elovitz, M.A.; Parry, S.; Bushman, F.D. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome* **2016**, *4*, 29. [[CrossRef](#)] [[PubMed](#)]
15. Glassing, A.; Dowd, S.E.; Galandiuk, S.; Davis, B.; Chiodini, R.J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* **2016**, *8*, 24. [[CrossRef](#)] [[PubMed](#)]
16. Perez-Muñoz, M.E.; Arrieta, M.-C.; Ramer-Tait, A.E.; Walter, J. A critical assessment of the “sterile womb” and “in utero colonization” hypotheses: implications for research on the pioneer infant microbiome. *Microbiome* **2017**, *5*, 48. [[CrossRef](#)] [[PubMed](#)]
17. Key, F.M.; Posth, C.; Krause, J.; Herbig, A.; Bos, K.I. Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends Genet.* **2017**, *33*, 508–520. [[CrossRef](#)] [[PubMed](#)]
18. Jónsson, H.; Ginolhac, A.; Schubert, M.; Johnson, P.L.F.; Orlando, L. mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinform. Oxf. Engl.* **2013**, *29*, 1682–1684. [[CrossRef](#)]
19. Santiago-Rodriguez, T.M.; Fornaciari, G.; Luciani, S.; Dowd, S.E.; Toranzos, G.A.; Marota, I.; Cano, R.J. Taxonomic and predicted metabolic profiles of the human gut microbiome in pre-Columbian mummies. *FEMS Microbiol. Ecol.* **2016**, *92*, fiw182. [[CrossRef](#)] [[PubMed](#)]
20. Ziesemer, K.A.; Mann, A.E.; Sankaranarayanan, K.; Schroeder, H.; Ozga, A.T.; Brandt, B.W.; Zaura, E.; Waters-Rist, A.; Hoogland, M.; Salazar-García, D.C.; et al. Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Sci. Rep.* **2015**, *5*, 16498. [[CrossRef](#)] [[PubMed](#)]
21. Weyrich, L.S.; Duchene, S.; Soubrier, J.; Arriola, L.; Llamas, B.; Breen, J.; Morris, A.G.; Alt, K.W.; Caramelli, D.; Dresely, V.; et al. Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature* **2017**, *544*, 357–361. [[CrossRef](#)] [[PubMed](#)]
22. Valtueña, A.A.; Mitnik, A.; Key, F.M.; Haak, W.; Allmäe, R.; Belinskij, A.; Daubaras, M.; Feldman, M.; Jankauskas, R.; Janković, I.; et al. The Stone Age Plague and Its Persistence in Eurasia. *Curr. Biol.* **2017**, *27*, 3683–3691.e8. [[CrossRef](#)]
23. Vågene, Å.J.; Herbig, A.; Campana, M.G.; García, N.M.R.; Warinner, C.; Sabin, S.; Spyrou, M.A.; Valtueña, A.A.; Huson, D.; Tuross, N.; et al. Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat. Ecol. Evol.* **2018**, *2*, 520–528. [[CrossRef](#)] [[PubMed](#)]
24. Herbig, A.; Maixner, F.; Bos, K.I.; Zink, A.; Krause, J.; Huson, D.H. MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* **2016**, 050559. [[CrossRef](#)]
25. Huson, D.H.; Beier, S.; Flade, I.; Górska, A.; El-Hadidi, M.; Mitra, S.; Ruscheweyh, H.-J.; Tappu, R. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput. Biol.* **2016**, *12*, e1004957. [[CrossRef](#)] [[PubMed](#)]
26. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]

