GENOME WATCH

Beyond the palaeomicrobiology

Helena Seth-Smith

This month's Genome Watch highlights the power of palaeomicrobiology in extracting detailed information about the genomes of ancient microorganisms.

Phylogenetic analysis of the genomes of currently circulating bacterial strains allows us to infer a great deal about the genomes of their common ancestors. Palaeomicrobiology offers an alternative approach by enabling direct examination of these ancestral genomes. Studies in this field can potentially provide unprecedented insights into major evolutionary questions about bacterial mutation rates, genome degradation and the impact of mobile elements.

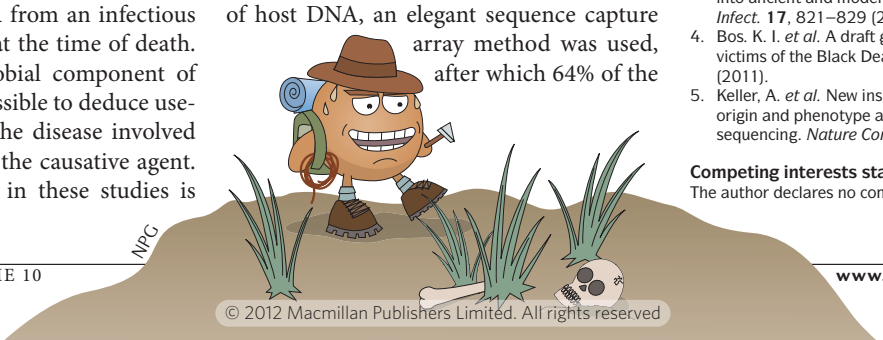
Palaeomicrobiology is derived from palaeobiology, which has seen the use of preserved hair samples to obtain partial genome sequences of extinct mammals, such as ancient humans (~4,000 years old)[1] and woolly mammoths (~20,000 years old)[2]. However, determining genome sequences from ancient DNA (aDNA) is problematic, as the material is often scarce and difficult to process in a way that maintains its sterility; thus, the likelihood of contamination is high. DNA also degrades over time, and in samples >1,000 years old it is usually fragmented into <200-base-pair (bp) pieces. Furthermore, deamination of bases occurs over time, resulting in miscalls during sequencing and the false identification of genomic variations. Thus, experiments using aDNA must be designed with these limitations in mind.

The principle behind palaeomicrobiology is that particular tissues within a skeleton or corpse may retain DNA from an infectious agent that was present at the time of death. By analysing the microbial component of the aDNA, it may be possible to deduce useful information about the disease involved and, more importantly, the causative agent. An additional obstacle in these studies is the fact that the microbial aDNA fragments make up only a fraction of the total aDNA, so enrichment and amplification are likely to be required.

Pathogenic bacteria, viruses and parasites have all been detected in ancient samples, and several studies over the past 20 years have focussed on *Mycobacterium tuberculosis* (reviewed in REF. 3). Previously, evidence of tuberculosis infection in corpses relied on the presence of bone lesions resulting from tuberculosis-associated osteomyelitis. In the palaeomicrobiology studies, aDNA was extracted from ancient human tissue samples (including lung, lymph nodes and bone) and a 17,000-year-old bison skeleton, and the *M. tuberculosis* DNA within this aDNA was detected using PCR. Further investigations used spoligotyping (which gives detailed typing information by amplifying specific genomic regions) and mycolic acid profiles to infer the ancestry of modern *M. tuberculosis* lineages[3]. The combined efforts of these studies have helped palaeoepidemiologists in determining the historical global distribution of tuberculosis.

The identity of the causative agent of the Black Death plague in the fourteenth century has been controversial, although *Yersinia pestis* has been strongly implicated. The power of palaeomicrobiology was highlighted recently in a paper describing a draft genome sequence of *Y. pestis* obtained from victims of the Black Death[4]. aDNA was extracted from the dental pulp of four people who perished during 1348–1349 in London, UK. To enrich *Y. pestis* DNA from a background of host DNA, an elegant sequence capture array method was used, after which 64% of the next-generation sequencing reads mapped to a reference *Y. pestis* genome. This resulted in a mean genome coverage of greater than 20×, with an average read length of only 55 bp, reflecting the extensive DNA degradation. Phylogenetic analysis of the genomic data indicated that the ancient *Y. pestis* falls at the ancestral node shared by modern *Y. pestis*, with only 97 chromosomal single-nucleotide polymorphisms (SNPs) compared with this reference strain. With so few differences between the modern and plague-causing strains, it is likely that environmental and host factors were key issues in the devastation of the Black Death.

Future investigations will be able to focus on the host and the infectious agent together. For example, total DNA sequencing of bone from the Tyrolean Iceman produced good coverage of the nuclear genome along with the sequences of associated microbiota, which indicated that he was infected with *Borrelia burgdorferi* when he died[5]. With ongoing advances in DNA recovery, high-throughput genome sequencing and sequence capture techniques, many more such studies should now be possible.



Helena Seth-Smith is at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.
e-mail: microbes@sanger.ac.uk

doi:10.1038/nrmicro2768
Published online 12 March 2012

1. Rasmussen, M. *et al.* Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757–762 (2010).
2. Miller, W. *et al.* Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* **456**, 387–390 (2008).
3. Donoghue, H. D. Insights gained from palaeomicrobiology into ancient and modern tuberculosis. *Clin. Microbiol. Infect.* **17**, 821–829 (2011).
4. Bos. K. I. *et al.* A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* **478**, 506–510 (2011).
5. Keller, A. *et al.* New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Commun.* **3**, 698 (2012).

Competing interests statement
The author declares no competing financial interests.