

Using Ancient DNA Analysis in Palaeopathology: A Critical Analysis of Published Papers, with Recommendations for Future Work

C. ROBERTS^{a*} AND S. INGHAM^b

^a Department of Archaeology, Durham University, South Road, Durham DH1 3LE, UK

^b Department of Rheumatology, Clinical Sciences Building, City Hospital, Nottingham NG5 1PB, UK

ABSTRACT Despite an observable increase in the number of studies using ancient DNA analysis to diagnose disease in human remains, there remain issues to be addressed about the quality of the resulting publications. This paper describes the qualitative analysis of published papers that describe the detection of pathogenic DNA in human skeletal and mummified remains from archaeological sites. Its ultimate goal is to provide an overview of the main problematic issues in relationship to standards developed in molecular biology and to make recommendations for future work. Sixty-five papers published between 1993 and 2006 were surveyed and the quality of each was assessed using 15 criteria. Interesting results emerged. Of particular note was the high number of papers that did not acknowledge the use of even basic contamination control (90%) or procedures to validate results independently (85%). This study illustrates that attention to contamination control and authentication of results is needed in future research, if confidence in aDNA analysis in palaeopathology is to be increased. Additionally, methods of analysis must be described in published papers to ensure transparency in processes utilised to generate the data. Copyright © 2008 John Wiley & Sons, Ltd.

Key words: aDNA; palaeopathology; diagnosis; methods

Introduction

Despite rapid advances in genetic analysis since the 1953 discovery of the DNA double helix (Watson & Crick, 1953), it was not until 1989 that polymerase chain reaction (PCR) was successfully used to amplify ancient DNA in bone. Hagelberg & Sykes' (1989) achievements in the extraction and amplification of poor quality ancient DNA fragments from bone from several time periods led to significant implications for using ancient DNA to explore anthropological and archaeo-

logical questions. However, it would not be until the early 1990s and the detection of *Mycobacterium tuberculosis* complex DNA that PCR-based amplification methods started to be used to diagnose disease in skeletal and mummified remains (Spigelman & Lemma, 1993; Rafi *et al.*, 1994; Salo *et al.*, 1994). While studies such as these have increased, the use of ancient DNA analysis to explore the past history of disease still remains an 'emerging field of research' (Zink *et al.*, 2002: 141), but one with great potential to provide information about the origin, evolution and transmission of disease through time which had not been possible prior to aDNA analyses.

PCR and hybridisation systems have become commonplace techniques in ancient DNA

* Correspondence to: Department of Archaeology, Durham University, South Road, Durham DH1 3LE, UK.
e-mail: c.a.roberts@durham.ac.uk

laboratories over the last two decades. To date these are the only two methods that can be successfully used to analyse such small quantities of fragmentary ancient DNA. Hybridisation analysis is a long-established technique for genetic investigations; nevertheless, its favour with ancient DNA analysts is much reduced in comparison to its more sensitive counterpart, PCR (Brown, 2000). However, 'the question of authenticity of DNA isolated from human remains is the main concern' (Faerman *et al.*, 2000: 154). With only trace amounts of degraded DNA to analyse, the application of the extremely sensitive technique of PCR is necessary; the main problem, however, is that the degree of PCR sensitivity means it is highly prone to contamination by modern DNA present within its vicinity (Faerman *et al.*, 2000).

The destructive nature of bone and soft tissue sampling and analysis for ancient DNA analysis additionally causes concerns among physical anthropologists and museum curators (DeGusta & White, 1996). Many have objections to such destruction of valuable skeletal collections, and it is usual that special permission is required before sampling, or that sampling in some cases is not allowed. Due to the difficult nature of isolating and extracting ancient DNA from ancient human remains (DeGusta & White, 1996), many recommend that an aDNA viability analysis is completed (e.g. Kumar *et al.*, 2000; Haynes *et al.*, 2002; Reed *et al.*, 2003; Barnes & Thomas, 2006), along with experimental work on DNA survival following specific treatments and contamination (e.g. Waite *et al.*, 1997; Gilbert *et al.*, 2006; Bouwman *et al.*, 2006), before any samples are taken for extensive DNA analysis on valuable and irreplaceable collections of archaeological human remains. For example, the radiography of human remains might affect the survival of aDNA (although the authors are not aware of any studies being done), as could specific treatments applied to human remains in post-excavation processing, conservation and curation (e.g. washing, glues, storage conditions; see Nicholson *et al.*, 2002; Pruvost *et al.*, 2007). Discussions about whether aDNA actually survives in samples from human remains or not in certain parts of the world has seen heated debates at times (e.g. see Gilbert *et al.*, 2005a; Zink & Nerlich, 2005, on Egypt). While

aDNA viability research is important to prevent unnecessary damage to valuable human remains, the research proposed must also be viable and be testing a specific hypothesis or answering questions about the past that cannot be answered any other way. As Kaestle & Horsbaugh (2002) have stated, 'such studies must not be undertaken merely to demonstrate that surviving DNA is present in organic remains...'

Research using aDNA analysis of pathogens over the past 15 or so years has concentrated on: diagnosis of diseases that do not, or may not, leave visible changes recognisable in the skeleton (e.g. malaria: Taylor *et al.*, 1997; the plague: Drancourt *et al.*, 1998; Raoult *et al.*, 2000; Wiechmann & Grupe, 2005; *E. coli*: Fricker *et al.*, 1997; tuberculosis: Jankauskas, 1999), diagnosis of non-specific pathological lesions (e.g. Haas *et al.*, 2000), confirming a diagnosis based upon other criteria (Baxarias *et al.*, 1998; Taylor *et al.*, 2000; Mays & Taylor, 2002), and identification of the specific organism causing a disease (e.g. Zink *et al.*, 2004), with tuberculosis being the most common disease considered (e.g. Salo *et al.*, 1994; Gernaey *et al.*, 2001; Mays & Taylor, 2003). Most studies to date have also focused on individual skeletons or mummies rather than 'samples' (e.g. Taylor *et al.*, 2000) but occasionally there have been analyses of larger numbers of individuals (e.g. Faerman *et al.*, 1997; Mays *et al.*, 2001; Fletcher *et al.*, 2003a,b). Of particular interest and importance are the papers that (rarely) report negative results for ancient pathogenic DNA analysis (e.g. Bouwman & Brown, 2005), and the equally few papers that consider the pathological processes in a disease that will affect whether the organism's DNA will survive to be extracted from bones or teeth and amplified (e.g. see Von Hunnius *et al.*, 2007, on the poor likelihood of treponemal DNA surviving in bone in the later, compared with earlier, stages of the disease process).

Many early papers using aDNA analysis to diagnose disease were guilty of not taking precautionary measures to prevent contamination. Because these studies were the examination of pathogenic rather than human DNA, it was wrongly assumed that contamination with modern DNA had no real bearing upon the resulting outcome (Bouwman & Brown, 2005: 704). This is erroneous, and many now agree that

careful contamination prevention and authentication of results are essential when analysing ancient DNA.

A number of critical papers suggesting that insufficient attention had been paid to contamination prevention and authentication were written many years ago in some cases (Richards *et al.*, 1995; Cooper & Poinar, 2000; Yang, 2003). Furthermore, although some researchers have focused on exploring and solving some of the problems in aDNA analyses (e.g. Malmström *et al.*, 2005; Sampietro *et al.*, 2006), and many papers might recognise that there are criteria for aDNA analyses, a considerable majority fail to respect the criteria. The recommendations in these critical papers must be implemented more often to authenticate data presented to the scientific and wider community.

The aim of the present study is to analyse the quality of published papers on pathogenic ancient DNA analyses in human remains, based on specific criteria.

Material and methods

A literature search of published papers reporting the diagnosis of disease in human remains using ancient pathogen DNA analysis was made, and each assessed for quality. Quality was assessed based on the description in the paper of the following methodological criteria that should be achieved. Some criteria (as indicated by *) were described originally by Cooper & Poinar (2000), with additional criteria adopted for this study which are described in a range of published literature. Although Cooper & Poinar's (2000) recommended criteria focused on all ancient DNA research, it was considered that they equally applied to ancient pathogenic DNA studies:

- Sterile sampling at the excavation level to limit contamination
- *The use of a 'physically isolated work area' for the research to avoid contamination; this means a dedicated isolated environment and not a building with large amounts of DNA being routinely amplified (Cooper & Poinar, 2000)
- The use of dedicated work areas for ancient DNA analysis

- Taking of DNA records for all personnel, and considering relevant past studies of aDNA in the laboratory
- The use of protective clothing in the laboratory situation (masks, gloves, coat, head covering)
- Removal of surface contamination from samples
- Analysis of survival of mtDNA and nuclear DNA
- *Negative control amplifications (multiple extraction and amplification controls to detect sporadic/low copy number contamination, and reporting of all contamination results)
- *Recording of appropriate molecular behaviour (i.e. PCR amplification success should be inversely related to product size; reproducible mitochondrial DNA results should be obtainable if pathogen DNA is detected; sequences should make phylogenetic sense)
- *Reproducibility of results (results should be repeatable from the same or different extractions of DNA from a sample)
- *Sequencing procedures (direct PCR sequences must be verified by cloning amplified products, cloning being an improvement on direct sequencing)
- *Independent replication of results in an independent laboratory (to discount intra-laboratory contamination and confirm data, especially if results are novel)
- *Assessment of preservation of other biochemical components of samples (indirect evidence of aDNA survival by assessing the relative extent of diagenetic change in amino acids and other residues)
- *Quantitation (to assess the copy number of the DNA target)
- *Analysis of associated remains, for example animal bones, to assess whether aDNA survives at the archaeological site.

Results

Sixty-five papers were identified from 29 journals, dating from 1993 to 2006. The exponential increase in the number of aDNA papers on the study of past disease over the last 15 years has fuelled the need for a universal set of explicit analytical criteria. However, as Gilbert *et al.* (2005b: 541) have suggested, criteria may not be

foolproof and must not replace 'thought and prudence when designing and executing ancient DNA studies'.

Although the general standard of research has improved through time, none of the papers considered in this study adhered to all the criteria outlined. Contrary to expectations, most of the 65 papers did not reveal improvements in quality over time, with a distinct absence of an increasing awareness for the need of contamination control and authentication, associated with the provision of a detailed methodology. Table 1 summarises the results.

Sterile sampling at the excavation level

Only 8% of papers highlighted the need for careful excavation of the bones to be used for aDNA analysis, and 90% did not discuss excavation procedures at all. No improvement through time was noted.

Physically isolated work areas

Fifty-two per cent of the papers described the use of physically isolated work areas for ancient DNA analysis. However, 42% of the papers did not denote whether isolated work areas were or were not used for pre-PCR and post-PCR work. There

was a slight increase in the use of isolated work areas in more recent research.

Dedicated work areas

There were 25% of papers that described the use of a dedicated ancient DNA laboratory for the work. A further 28% partly discussed this criterion, either through the description of a previous absence of any DNA work in the laboratory, or the presence of any bacterial/viral samples within the laboratory.

However, 45% of the papers still contained no reference to whether dedicated work areas were used, whilst a further 2% of papers confirmed the distinct absence of the use of a dedicated laboratory. It is possible to see a slight improvement in the frequency of the use of dedicated work areas in more recent papers, particularly after 2002, but once again this is a limited improvement.

DNA records

A recent development has been initiated by some researchers in the field to compile a modern and ancient DNA profile for each aDNA laboratory, providing a database for the cross-referencing of all ancient DNA sequences, both human and

Table 1. Summary of results: % of papers where criteria were not discussed (ND), fully discussed (FD), partly discussed (PD), or discussed but not addressed in the methodology (DNA)

	ND	FD	PD	DNA
Sterile sampling	90	8	2	—
Physically isolated work area	42	52	6	—
Dedicated aDNA laboratory	45	25	28	2
Laboratory DNA records	97	3	—	—
Protective clothing	65	18	17	—
Contamination removal	34	49	17	—
MtDNA isolation	89	11	—	—
Nuclear DNA isolation	66	29	5	—
Negative controls	11	29	60	—
Molecular behaviour	5	86	9	—
Reproducibility of results	20	49	31	—
Sequencing procedures	26	5	63	6
Independent replication	85	12	3	—
Preservation of biochemical components	95	5	—	—
Quantitation	93	5	2	—
Associated remains	92	2	6	—

pathogenic. The description of this criterion in papers shows that 97% did not discuss the provision of a genetics database as a means of contamination control and result authentication. Only 3% of articles partially address this criterion, often only noting some form of screening of their personnel. There was no improvement through time.

Use of protective clothing

There were 65% of papers that did not discuss protective clothing as a precaution against contamination. Moreover, of the remaining 35% of papers, only 18% fully described the use of protective clothing worn by staff. Presumably, the finding that the other 17% of papers indicated that investigators wore gloves only satisfied the anti-contamination procedures in place at that particular laboratory. There were no improvements to methodological practice in more recent papers.

Removal of surface contamination

There were 49% of papers that acknowledged the need for, and practice of, contamination removal before DNA sampling. However, an additional 34% of papers did not mention contamination removal. In relation to these findings, one cannot be sure whether contamination removal did occur and was simply not described in the paper, or whether it was not conducted at all. There was a small improvement in the use of surface contamination removal in more recent papers.

Survival of mitochondrial and nuclear DNA

A promising 29% of papers recorded the amplification of human nuclear DNA alongside the pathogenic DNA as a means of standardised confirmation. However, this must be seen in the context of the (majority) 66% of papers that did not discuss the amplification of human nuclear DNA. In contrast, only a limited number of papers actually published information regarding the amplification of mitochondrial DNA (11%),

with 89% of papers having no mention of mitochondrial DNA at all. There appeared to be no improved amendments to methodological practice associated with more recent papers.

Control amplifications

Negative controls with extraction blanks were confirmed to be the most popular method of control amplification; 29% of papers clearly described the use of negative contamination control, with a further 60% partially describing negative controls. It is clear that, although many use negative controls, most do not state the number applied.

Recording of appropriate molecular behaviour

As a template for appropriate molecular behaviour, the amplification value of <300 bp was evaluated. Eighty-six per cent of papers reproduced the expected amplification product during the course of their investigations. Only 9% of the papers partially described appropriate molecular behaviour, with a mere 5% of papers investigated not discussing the fragment size. This indicates that such a feature was an expected part of the method used, even in the early days of ancient DNA analysis. There was no alteration in procedures through time.

Reproducibility of results

Reproducibility of results has been increasingly used over time, indicating that many researchers are now appreciating the necessity for result authentication. Forty-nine per cent of the papers noted that a PCR sensitivity test had been carried out. However, only 31% of papers addressed this criterion. No alteration of procedures through time was noted.

Sequencing procedures

There were 63% of papers that partly discussed DNA sequencing as a means of DNA

authentication. In the majority of these cases the authors appear to have only carried out direct sequencing (with 5% also undertaking cloning; this was presumably due to prohibitive costs and the time needed). Despite the obvious usefulness of sequencing, 26% of investigators did not clearly state whether or not they used the method. There appear to be no improvements to practice that can be associated with more recent papers.

Independent replication of results

Independent replication of data in a second laboratory was not discussed in 85% of the papers. This is considered a very important and simple stage in ancient DNA authentication that should have been easily followed from the first ancient DNA initiative. Only 12% of the papers described independent replication. However, there was a marked recent improvement in the number of papers that have begun to include the practice of independent replication of results as part of standard of authentication.

Biochemical preservation

A very high percentage (95%) of papers did not describe the analysis of other biochemical residues, presumably because of the unknown application of such procedures to pathogenic DNA. However, 5% of the papers (more recent in publication date) covered some biochemical analysis (from amino acid racemisation to mycolic acid extraction/derivation). No significant change with time was noted.

Quantitation

The possibility of sporadic instances of contamination due to low copy numbers was only discussed in 5% of articles, with 2% of these providing no further specific details on the matter. Certainly, the remaining 93% of papers did not feel the need to discuss such issues within the context of their work.

Analysis of associated remains/soil samples

Many of the papers considered (92%) did not mention the use of associated remains or soil samples as a means of contamination control. Only two papers recorded the use of associated faunal remains as a contamination precaution. In reality, the 6% of papers that were recorded as partially covering this criterion analysed soil samples from the investigation site as a means of identifying any pathogenic contaminating DNA at the burial location. There was no correlation between alteration in procedures and the date of publication.

Discussion

Overall, the results of this study showed that some criteria were better adhered to than others. Sterile sampling (10%), amplification of MtDNA (11%), independent replication of results (15%), biochemical preservation (5%), quantitation (8%) and analysis of associated remains (8%) were infrequently described in papers, although some criteria were better fulfilled (e.g. negative controls, molecular behaviour, reproducibility and sequencing). These results have implications for future research.

Improved conditions during excavation of human remains to limit contamination are critical – as Herrmann & Hummel (1994: 62) stated, 'the vast majority of specimens do not derive from sterile places'. Humans were disposed of in a variety of environments, incorporating many potential contaminants, alongside those of human contact. Furthermore, if sterile sampling is in place then this might prevent or inhibit post-excavation degradation of preserved DNA; if this is not implemented, then an increased risk of contamination being the source of DNA sequences is possible. For example, Pruvost *et al.* (2007) found that freshly excavated and non-treated unwashed bone contained six times more DNA and had twice as many authentic DNA sequences as bones treated with standard procedures. They also found that DNA preservation varied for different bones of the same c.3200-year-old aurochs excavated in two seasons 57 years apart; no DNA amplification was

possible for the washed museum-stored bones, but the freshly excavated bones yielded authentic sequences. Therefore, for example, the use of sterile gloves should be routine on all excavations (Brown, 2000) to limit contamination of ancient DNA samples. Researchers should be aware that, as many samples for aDNA analysis are taken many years after excavation and original analysis and curation, the potential is great for contamination from a variety of sources. With respect to ancient pathogen DNA, contamination from environmental micro-organisms is a real risk (Gilbert *et al.*, 2004, 2005c). Furthermore, a knowledge of the history of the handling of the human remains sampled *before analysis* is critical, and without this information 'it is very difficult to comment objectively on the reliability of results'. Some research has suggested that there are certain analytical methods that will identify contamination, such as real-time quantitative PCR (Pruvost & Geigl, 2004).

An isolated work area to minimise the risk of contamination is important, so ancient DNA work should be carried out in specialised laboratories (Bramanti *et al.*, 2003; Bouwman & Brown, 2005). It has been deemed essential to physically separate pre-PCR and post-PCR work spaces, with analytical instruments never being exchanged between these locations (Herrmann & Hummel, 1994; Yang, 2003: 171). Such isolation is needed to prevent contamination of pre-treated samples. Certainly, Yang (2003: 171) indicated that the 'pre-PCR laboratory should have a UV-filtered ventilation system and positive pressure airflow' to ensure isolation. Dedicated work areas are also important so that the workspace used for ancient DNA analysis is separate from other DNA research. In no circumstance should ancient DNA studies take place in a laboratory situation where modern DNA is examined, particularly if it is of the same pathogenic background (Stoneking, 1995). The possibilities of cross-contamination would be too high, since pathogenic aDNA can be contaminated by modern pathogenic DNA (Bouwman & Brown, 2005).

The DNA record of all personnel and past studies in a laboratory needs attention. It is clear that it would be helpful if each aDNA laboratory set up a database of the genetic fingerprint of all

laboratory workers (e.g. see Sampietro *et al.*, 2006), alongside a history of all pathogenic material that has been studied within the laboratory (including any genetic details). This database should be cross-referenced whenever any aDNA is isolated, and especially in the generation of unexpected results; this is to ensure cross-contamination has not occurred. Disposable facemasks, head and foot coverings, laboratory coats and gloves should be in use every time an individual enters the aDNA work area (Bramanti *et al.*, 2003: 109; Yang, 2003: 171). Laboratory coats must be strictly limited to the specific laboratory and must be frequently washed (Bramanti *et al.*, 2003: 109), although disposable gowns are better.

Modern mitochondrial and nuclear DNA can penetrate deep into bones and teeth during routine handling during and after excavation (Cooper *et al.*, 2004: 431), but the depth of penetration will vary considerably according to the bone or tooth being sampled and its preservation. Some have advocated the removal of surface contamination (Haas *et al.*, 2000; Fletcher *et al.*, 2003a), but whether contamination is necessarily always removed during these processes is debatable. However, several methods have been tested as being successful for the removal of external contamination (Herrmann & Hummel, 1994). For example, Richards *et al.* (1995) undertook a series of tests on deliberately contaminated, fragmented pig bones. Only the non-handled bone and the bone cleaned with shot blasting provided a sample of the endogenous pig DNA, whilst the handled sample showed that modern DNA successfully contaminated the ancient fragments. The samples that were soaked in 0.5% sodium hypochlorite generated equal amounts of pig and human DNA, suggesting that, although this may not be the most efficient means of removing contamination, it is still a useable method when it comes to highly fragmented and poorly preserved remains (Richards *et al.*, 1995). Stoneking (1995) has suggested that, at the very least, the removal of the outer surface layer of bone or teeth must be carried out so as to reduce the possibility of external contamination of the DNA sample. There has been some progress in removing the outer, potentially contaminated, layer of bones or teeth. There are even approved

methods that can be used if the sample is too small for the standard abrasive surface removal. One such method is known as pre-treatment, and uses ultraviolet wavelengths to disrupt the DNA bonds present on the surface of bone i.e. the layer where modern impurities could settle, thus preventing amplification of any contaminant DNA before sampling occurs (Herrmann & Hummel, 1994). While on the basis of the papers surveyed here, there has been a chronological trend in increased surface contamination awareness and removal into the 21st century, it is not commonly appreciated that 'removing the surface of an ancient sample may not totally remove modern contaminants' (Cooper, 1997: 1001). In the case of bone, in particular, where the structure is very porous, deeply penetrating modern DNA contamination is almost inevitable (Gilbert *et al.*, 2006). Gilbert *et al.* (2005c) assessed levels of contamination in teeth and bones from an Italian cemetery and found both were readily contaminated and were difficult, if not impossible, to decontaminate. Furthermore, Kemp & Smith (2005), in their tests of the effect of bleach on the survival of endogenous DNA, found that the DNA was stable even to extreme treatments. Thus, the processes behind ultimate contamination and the mechanisms for decontamination are not well understood; removal of surface contamination therefore may not be the answer in many cases.

It is not only important to assess whether DNA survives in a bone or tooth sample, but also whether the sample analysed has DNA that is mitochondrial or nuclear. Such information has a bearing on the observation of contamination, as one individual should only have two copies of nuclear DNA (Yang, 2003: 172). Hence, if more distinctive fragments are observed in the final sequencing, it could be concluded that contamination had occurred. Little progress has been made regarding the importance of whether mtDNA or nuclear DNA is present in a sample. This is surprising given that Cooper (1997: 1001) highlighted how multiple copy mtDNA survives for greater periods of environmental stress than single copy nuclear DNA. This is most likely because mtDNA has a greater numerical presence to start with (being a multi-copy component). The statistical chance of some mtDNA surviving

is therefore greater. In terms of contamination, monitoring mtDNA means a more solid basis for contamination detection. If more than one type of mtDNA is recovered, then contamination is likely (Yang, 2003).

Negative controls are essential to monitor possible contamination of extraction reagents (Stoneking, 1995: 1259; Yang, 2003: 172). In order to provide statistical authentication, as many negative controls as DNA extracts must be set up. In most of the papers surveyed here there was no recognition that multiple negative controls needed to be present in order to produce valid detections (Yang, 2003: 171). Certainly, the sensitivity of the method should not be authenticated without a sufficient number of negative controls, as individually they have a reduced capability of picking up low levels of contamination (Yang, 2003).

With respect to appropriate molecular behaviour, several papers argued that the PCR was inversely proportional to the product size (Cooper & Poinar, 2000), with the optimal fragment size being approximately 200–300 bp. However, the size of the DNA fragment that is amplifiable depends on many factors such as its age and its burial conditions. Nevertheless, the shorter the target fragment, the more DNA there is potentially available for amplification (Poinar *et al.*, 2006). Any results obtained during an investigation must be reproduced both from the same and different samples (Cooper & Poinar, 2000). Multiple DNA extractions are often carried out, 'so that concordance of the results from different extracts' could serve as an additional check on the authenticity of results (Stoneking, 1995: 1259).

Direct sequencing is often considered the most specific form of target analysis. Its acceptance was reflected in most papers that described the type of sequencing used as being direct sequence analysis. However, many argue that, 'direct PCR sequences must be verified by cloning amplified products to determine the ratio of endogenous to exogenous sequences' (Cooper & Poinar, 2000: 1140). Cloning amplified products is therefore essential for the verification of PCR sequencing via the determination of the endogenous to exogenous sequence ratio (Cooper & Poinar, 2000). Bower *et al.* (2005) noted that the number of

clones to be sampled is also important in order to determine a reliable consensus sequence. They estimated that 20 need to be sampled to be 95% confident of identifying the most abundant sequence present at 70% in an ancient sample; sequencing a few clones can be more error-prone than direct sequencing. The findings of this study regarding sequencing and cloning are possibly due to what must only be imagined as the lower cost and reduced time for direct sequencing over the extra cloning (Brown & Brown, 1992: 15). Consequently, it appears logical that to ensure future authentication, cost be placed to one side.

Clearly, it is extremely important that independent replication of results in a second different laboratory is used as a means of assessing the quality and authenticity of the data. Extraction, amplification and sequencing should occur in other (independent) laboratories to ensure that no laboratory is contaminating the sample or misinterpreting the data (Cooper & Poinar, 2000: 1140; Yang, 2003: 172). This study shows that replication in an independent laboratory is still very much underused, although Gilbert *et al.* (2004) comment that if a sample is contaminated before analysis, split, and one half sent to another laboratory, both laboratories will produce a contaminant sequence. In the past many suggested that independent replication should not be carried out due to potential cost or logistical problems (Stoneking, 1995), specifically relating to the perhaps unnecessary destruction of valuable remains, and the extended time period that would be required for testing. Additionally, they argued that safety precautions would not allow the removal of bone samples to another laboratory, and many said, 'to require such independent analysis would cause more problems than it would solve' (Stoneking, 1995: 1260). It seems that, even after a lapse of over ten years, many scientists still hold the same view, with little being done to support the use of independent laboratories. In many papers the results were reproduced several times within the same laboratory, but the exact figure was often not stated, nor was the number of different extracts that had been sampled. Although independent replication is a more time-consuming and costly procedure, it is absolutely essential to

discover intra-laboratory error and errors in interpretation.

Analysing other biochemical components of a sample can be a means of indirect evidence for DNA preservation (Cooper & Poinar, 2000), although the application of this to pathogenic DNA remains unknown. With respect to quantitation, investigators should be wary that low copy numbers of the starting DNA template could lead to sporadic incidences of contamination and must therefore be noted in any results published (Cooper & Poinar, 2000: 1140; Yang, 2003: 172). Recently recommendations have been made for the use of associated remains and soil samples as a means of contamination control, especially as soil micro-organisms are believed to be a source of contaminant DNA (Gilbert *et al.*, 2004). The remains/soil are treated in exactly the same manner as the human sample to ensure that contamination in any part of the analysis is detected. In most situations, nearby faunal remains are collected at the site. For example, Richards *et al.* (1995: 292) obtained animal remains 'directly from an excavation at Guildford Castle to use as controls in contamination tests'. As Cooper (1997: 1002) stated, 'Powerful supporting evidence can also be provided by associated faunal remains'. In reality, the use of associated animal remains did not feature in most papers. This is surprising considering that, early on in ancient DNA studies, the use of associated animal remains was recorded as a potential basis of negative control (Herrmann & Hummel, 1994: 63).

Heralded as a significant advance in the study of palaeopathology in the early 1990s, the application of ancient DNA analysis to questions about the origin and evolution of disease still holds problems. This study has shown that many of the criteria on which the quality of papers were considered were not described, and there was an alarming absence of improvement in the quality of papers through time. Even as recently as 2002, Spigelman *et al.*, on the first pathogenic DNA paper published, wrote, 'this research received some criticism as technical procedures were not described fully and results were not confirmed' (2002: 393). Despite this acknowledgement, many researchers still seemed reluctant to address in writing the problems and solutions to laboratory contamination, and to authenticating

data with ultimate effects on interpretation (Bouwman & Brown, 2005). Furthermore, Cooper and Poinar (2000: 1139) said that, 'The need to authenticate results became obvious in the mid-1990s when a series of high-profile studies were shown to be unrepeatably'. In spite of these recommendations, the use of quality control still remains sparse and inconsistent among even respected journal publishers. In particular, this study highlights the fact that only a few criteria are routinely described in papers (use of negative controls, describing appropriate molecular behaviour, reproducing results, discussing sequencing procedures, and removal of surface contamination). Finally, it was readily noted that 'few papers report negative results in the search for pathogen aDNA' (Bouwman & Brown, 2005: 712). Theoretically, several of the papers investigated during this study should have obtained negative detections, as the pathogens allegedly identified were unlikely to have survived the conditions in which the human remains were found. It has been emphasised by Bouwman and Brown (2005: 711) that although 'some of these reports are undoubtedly correct', many lack the rigorous validation of assurance.

One of main problems overall, however, was that there was generally poor communication of information to the reader of the published papers. In many instances it was simply impossible to deduce whether or not certain validation criteria had been addressed at all, or whether they had merely not been described. It is accepted that restrictions on inclusion of detailed methodological information may be in place for some journals, but if this is a problem then it needs to be rectified in all future papers if the scientific community are to continue to support ancient DNA studies. Until such improvements in communication occur, and pathogen aDNA studies fully mature, one must view all results with caution. As Cooper & Poinar (2000: 1141) stated, 'If ancient DNA research is to progress... it is essential that journal editors, reviewers and granting agencies, and researchers... subscribe to the criteria'.

Conclusion

The aim of this study was to provide a critical review of the quality of published ancient

pathogenic DNA papers, with a particular focus on investigating the methods used in the detection of pathogen DNA. The study focused on the formation of a set of criteria with which to review the papers. Few papers adhered to many of the criteria, and there appeared to be no improvement through time.

Interesting findings transpired, particularly relating to the unexpectedly large percentage of papers that did not even discuss some of the more basic criteria. Of particular note was the extremely high percentage (85%) of papers that did not acknowledge the use of independent replication of results, despite the ready acceptance that such duplication is the only means by which intra-laboratory error can be dismissed. Although already widely accepted that the lack of independent replication was a problem among ancient DNA studies, it appears that many are unaware of just how serious this problem is (although note Gilbert *et al.*'s (2004) comments).

In the latter part of the 20th century, aDNA analysis for diagnosing infectious disease to answer significant questions about the origin and evolution of disease was noted as a field with huge potential for interdisciplinary cooperation (Brown, 2000: 468). Indeed, non-morphological approaches to the study of human remains have increased over the last two decades, with the percentage of molecular-based papers more than doubling from 2% to 5% in the *American Journal of Physical Anthropology* alone (Stojanowski & Buikstra, 2005). However, more traditional analyses in that journal still overshadow non-morphological applications. Furthermore, molecular-based studies on human remains 'ha[ve] not only increased in visibility, but ha[ve] also dramatically increased in [citation] impact' (*ibid*: 104). Today, with advances in techniques of ancient DNA analysis, the potential for ongoing ancient pathogen research is extremely positive. However, this study shows that there is still an apparent inconsistency in methodology between many of the investigations that are carried out. 'Furthermore, the bulk of the ancient DNA work done to date deals with single individuals, or with remains that are widely separated in space and/or time' (Stoneking, 1995), a statement true of today. Such practices do not appear to have been modified much at all for the 21st century, and

there are justifiable complaints by many practising bioarchaeologists and archaeologists, along with curators of collections, that the destruction of valuable material should only be allowed if legitimate research questions are being asked, and not simply because 'we can'. Not only is there huge potential for ancient pathogenic DNA studies, but potential for the destruction of a valuable and ethically sensitive archaeological and anthropological resource. For future standards to be improved, scientists must begin by asking whether such an invasive technique is really necessary (Stoneking, 1995), and what interest and answers could be gained by carrying out this analysis? To move forward, population studies to access global patterns of disease are recommended as key foci, with specific questions to ask and hypotheses to test. This is contrary to the, what often appears to be, haphazard analysis of individual skeletons and mummies that have been the target in the past. While 'case studies' can be useful and interesting in their own right, especially when data from many are brought together, a 'population' approach provides palaeopathology with a more realistic view of patterning of disease in past society.

The majority of the recommendations that result from this study, such as independent replication, isolated work areas and appropriate excavation procedures, are all very simple practices that can easily be implemented without great inconvenience. A number of these proposals are of such a basic nature that they could (and should) have been applied during the early years of aDNA investigations. The criteria outlined below summarise the recommendations for future research in this field; the reader is also directed to the above discussion with respect to some criteria which are useful but where their limitations should be considered (*).

- Use of correct contamination prevention procedures during excavation
- Removal of surface contamination from samples*
- Use of separate pre-PCR and post-PCR laboratories
- Use of a dedicated ancient DNA laboratory
- Comparison of data with a DNA database for personnel working in the laboratory and previous DNA studies
- Use of protective clothing within the laboratory
- Noted analysis of mtDNA or nuclear DNA
- Use of appropriate negative controls
- Description of appropriate molecular behaviour
- Report on whether the results were reproduced
- Description of sequencing procedures*
- Independent replication of results in another laboratory*
- Report of survival of other biochemical components in the sample
- Recognition of the use of quantitation
- Use of associated remains/soil samples as controls.

If these criteria are applied, or at least considered, in all future studies, then there should be a reduction in the number of questions regarding the authenticity of results from analysis of pathogenic ancient DNA. There have been suggestions that, if specific methodological criteria can be agreed by biomolecular archaeologists and these criteria are published, then one reference to those criteria would suffice. Whether this is feasible will rely on the 'biomolecular archaeology community' debating whether this can happen. If there was agreement, this could also solve the problem of journal editors restricting the space available to describe methodological details in papers. As Gilbert *et al.* (2005b: 541) reminded us, researchers in this field should also 'take a more cognitive and self-critical approach (and) ... explain, in sufficient enough detail to dispel doubt, how the data were obtained, and why they should be believed to be authentic'. A balanced approach between using specific criteria and common sense would surely be advocated by most researchers. However, we should not forget Gilbert *et al.* (2005b: 541) who suggested that criteria may not be foolproof and must not replace 'thought and prudence when designing and executing ancient DNA studies'.

In bioarchaeology, we have seen the development and adoption of standards for basic data collection from human skeletal remains over the last 15 years (Buikstra & Ubelaker, 1994; Brickley & McKinley, 2004). As the analysis of ancient pathogenic DNA has contributed greatly to our understanding of the evolution and history of disease, it is hoped that there will be a similar

adoption of methodological standards. This will provide all interested parties, with a range of 'biomolecular knowledge' with more faith in, and understanding of, the research published.

Acknowledgements

The authors are very grateful for comments given by the audience when the content of this paper was presented at the Annual Meeting of the Paleopathology Association in Anchorage, Alaska, in 2006, and the 1st South American Paleopathology Association Meeting in Rio de Janeiro in 2005. Jane Buikstra and Alicia Wilbur of Arizona State University also kindly commented on an earlier draft of this paper. We are also particularly grateful to the two anonymous reviewers of the paper who have provided very constructive and helpful comments, directed us to other relevant papers, and ultimately improved the paper's message.

The British Academy and the Royal Society provided Conference Grants to the first author (CR) for attendance at the Brazil meeting. The original research for this paper formed the thesis component for the MSc Palaeopathology, Durham University, by the second author (SI).

References

- Barnes I, Thomas M. 2006. Evaluating bacterial pathogen DNA preservation in museum osteological collections. *Proceedings of the Royal Society B* **273**: 645–653.
- Baxarias J, Garcia A, Gonzalez J, Perez-Perez A, Tundo BG, Garcia-Bour CJ, Campillo DD, Turbon ED. 1998. A rare case of tuberculosis gonarthropathy from the middle ages in Spain: an ancient DNA conformation study. *Journal of Paleopathology* **10**(2): 63–72.
- Bouwman AS, Brown T. 2005. The limits of biomolecular palaeopathology: ancient DNA cannot be used to study venereal syphilis. *Journal of Archaeological Science* **32**: 703–713.
- Bouwman AS, Chilvers ER, Brown KA, Brown TA. 2006. Brief communication: identification of the authentic ancient DNA sequence in a human bone contaminated with modern DNA. *American Journal of Physical Anthropology* **131**: 428–431.
- Bower MA, Spencer M, Matsumura S, Nisbet R, Ellen R, Howe CJ. 2005. How many clones need to be sequenced from a single forensic or ancient DNA sample in order to determine a reliable consensus sequence? *Nucleic Acids Research* **33**: 2549–2556.
- Bramanti B, Hummel S, Chiarelli B, Herrmann B. 2003. Ancient DNA analysis of the delta F508 mutation. *Human Biology* **75**: 105–115.
- Brickley M, McKinley J (eds). 2004. *Guidelines to the standards for recording human remains*. Institute of Field Archaeologists Paper 7: Reading.
- Brown K. 2000. Ancient DNA applications in human osteoarchaeology: achievements, problems and potential. In *Human Osteology in Archaeology and Forensic Science*, Cox M, Mays S (eds). Greenwich Medical Media: London; 455–472.
- Brown T, Brown K. 1992. Ancient DNA and the archaeologist. *Antiquity* **66**: 10–23.
- Buikstra JE, Ubelaker D. 1994. *Standards for data collection from human skeletal remains*. Arkansas Archeological Survey Research Seminar Series 44: Fayetteville, Arkansas.
- Cooper A. 1997. Reply to Stoneking: ancient DNA – how do you know when you have it? *American Journal of Human Genetics* **60**: 1001–1002.
- Cooper A, Poinar HN. 2000. Ancient DNA: do it right or not at all. *Science* **289**: 1139–1141.
- Cooper A, Drummond AJ, Willerslav E. 2004. Dispatch ancient DNA: would the real Neanderthal please stand up? *Current Biology* **14**: 431–433.
- DeGusta D, White TD. 1996. Technical report on the use of skeletal collections for DNA analysis. *Ancient Biomolecules* **1**: 89–92.
- Drancourt M, Aboudharam G, Sigoli M, Dutour O, Raoult D. 1998. Detection of 400 year old *Yersinia pestis* DNA in human dental pulp: an approach to the diagnosis of ancient septicemia. *Proceedings of the National Academy of Science USA* **95**: 12637–12640.
- Faerman M, Jankauskas R, Gorski A, Bercovier H, Greenblatt ChL. 1997. Prevalence of human tuberculosis in a medieval population of Lithuania studied by ancient DNA analysis. *Ancient Biomolecules* **1**: 205–214.
- Faerman M, Nebel A, Filon D, Thomas MG, Bradman N, Ragsdale BD, Schultz M, Oppenheim A. 2000. From a dry bone to a genetic portrait: a case study of sickle cell anaemia. *American Journal of Physical Anthropology* **111**: 153–163.
- Fletcher HA, Donoghue HD, Holton J, Pap I, Spigelman M. 2003a. Widespread occurrence of *Mycobacterium tuberculosis* DNA from 18th–19th century Hungarians. *American Journal of Physical Anthropology* **120**: 144–152.

- Fletcher HA, Donoghue HD, Taylor GM, Van der Zanden AGM, Spigelman M. 2003b. Molecular analysis of *Mycobacterium tuberculosis* DNA from a family of 18th century Hungarians. *Microbiology* **149**: 143–151.
- Fricker EJ, Spigelman M, Fricker CR. 1997. The detection of *Escherichia coli* DNA in the ancient remains of Lindow man using polymerase chain reaction. *Letters in Applied Microbiology* **24**: 351–354.
- Gernaey AM, Minnikin DE, Copley MS, Dixon RA, Middleton JC, Roberts CA. 2001. Mycolic acids and ancient DNA confirm an osteological diagnosis. *Tuberculosis* **81**: 259–354.
- Gilbert MTP, Cuccui J, White W, Lynnerup N, Titball R, Cooper A, Prentice M. 2004. Absence of *Yersinia pestis*-specific DNA in human teeth from five European excavations of putative plague victims. *Microbiology* **150**: 341–354.
- Gilbert MTP, Barnes I, Collins MJ, Smith C, Eklund J, Goudsmit J, Poinar H, Cooper A. 2005a. Notes and comments. Long-term survival of ancient DNA in Egypt: response to Zink and Nerlich (2003). *American Journal of Physical Anthropology* **128**: 110–114.
- Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I. 2005b. Assessing ancient DNA studies. *Trends in Ecology and Evolution* **20**: 541–544.
- Gilbert MTP, Rudbeck L, Willerslev E, Hansen AJ, Smith C, Penkman KEH, Prangenberg K, Nielsen-Marsh CM, Jans ME, Arthur P, Lynnerup N, Turner-Walker, Biddle M, Kjølbye-Biddle B, Collins MJ. 2005c. Biochemical and physical correlates of DNA contamination in archaeological human bones and teeth excavated at Matera, Italy. *Journal of Archaeological Science* **32**: 785–793.
- Gilbert MTP, Hansen AJ, Willerslev E, Turner-Walker G, Collins M. 2006. Insights into the processes behind the contamination of degraded human teeth and bone samples with exogenous sources of DNA. *International Journal of Osteoarchaeology* **16**: 156–164.
- Haas CJ, Zink A, Molnar E, Szeimies U, Reischl U, Marcsik A, Ardagna Y, Dutour O, Pálfi G, Nerlich AG. 2000. Molecular evidence for different stages of tuberculosis in ancient bone samples from Hungary. *American Journal of Physical Anthropology* **113**: 293–304.
- Hagelberg E, Sykes B. 1989. Ancient bone DNA amplified. *Nature* **342**: 485.
- Haynes S, Searle JB, Bretman A, Dobney KM. 2002. Bone preservation and ancient DNA: the application of screening methods for predicting DNA survival. *Journal of Archaeological Science* **29**: 585–592.
- Herrmann B, Hummel S. 1994. *Ancient DNA: Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical and Forensic Specimens*. Springer Verlag: New York.
- Jankauskas R. 1999. Tuberculosis in Lithuania: palaeopathological and historical correlations. In *Tuberculosis: Past and Present*, Pálfi G, Dutour O, Deák H (eds). Golden Book Publishers and Tuberculosis Foundation: Budapest and Szeged, 551–558.
- Kaestle FA, Horsbaugh KA. 2002. Ancient DNA in anthropology: methods, applications, and ethics. *Yearbook of Physical Anthropology* **45**: 92–130.
- Kemp BM, Smith DG. 2005. Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. *Forensic Science International* **154**: 53–61.
- Kumar SS, Nasidze I, Walimbe SR, Stoneking M. 2000. Brief communication: discouraging prospects for ancient DNA from India. *American Journal of Physical Anthropology* **113**: 129–133.
- Malmström H, Storå J, Love D, Holmlund G, Götherström A. 2005. Extensive human DNA contamination in extracts from ancient dog bones and teeth. *Molecular Biology and Evolution* **22**: 2040–2047.
- Mays S, Taylor GM, Legge AJ, Young DB, Turner-Walker G. 2001. Palaeopathological and biomolecular study of tuberculosis in a medieval skeletal collection from England. *American Journal of Physical Anthropology* **114**: 298–311.
- Mays S, Taylor GM. 2002. Osteological and biomolecular study of two possible cases of hypertrophic osteoarthropathy from medieval England. *Journal of Archaeological Science* **29**: 1267–1276.
- Mays S, Taylor GM. 2003. A prehistoric case of tuberculosis from Britain. *International Journal of Osteoarchaeology* **13**: 189–196.
- Nicholson GJ, Tomiuk J, Czarnetzki A, Bachmann L, Pusch CM. 2002. Detection of bone glue treatment as a major source of contamination in ancient DNA analyses. *American Journal of Physical Anthropology* **118**: 117–120.
- Poinar HN, Schwarz C, Qi J, Shapiro B, MacPhee RDE, Buigues B, Tikhonov A, Huson DH, Tomsho LP, Auch A, Rapp M, Miller W, Schuster SC. 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* **311**: 392–394.
- Pruvost M, Geigl E-M. 2004. Real-time PCR to assess the authenticity of ancient DNA amplification. *Journal of Archaeological Science* **31**: 1191–1197.
- Pruvost M, Schwarz R, Bessa Correia V, Champlot S, Braguier S, Morel N, Fernandez-Jalvo Y, Grange T, Geigl E-M. 2007. Freshly excavated fossil bones are best for amplification of ancient DNA. *Proceedings of the National Academy of Sciences* **104**: 739–744.
- Rafi A, Spigelman M, Stanford J, Lemma E, Donoghue H, Zias J. 1994. DNA of *Mycobacterium leprae*

- detected by PCR in ancient bone. *International Journal of Osteoarchaeology* 4: 287–290.
- Raoult D, Aboudharam G, Crubezy E, Larrouy G, Ludes B, Drancourt M. 2000. Molecular identification by 'suicide PCR' of *Yersinia pestis* as the agent of medieval Black Death. *Proceedings of the National Academy of Sciences USA* 97: 12800–12803.
- Reed FA, Kontanis EJ, Kennedy KAR, Aquadro CF. 2003. Brief communication: ancient DNA prospects from Sri Lankan highland dry caves support an emerging global pattern. *American Journal of Physical Anthropology* 121: 112–116.
- Richards MB, Sykes BC, Hedges REM. 1995. Authenticating DNA extracted from ancient skeletal remains. *Journal of Archaeological Science* 22: 291–299.
- Salo WL, Aufderheide AC, Buikstra JE, Holcomb TA. 1994. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proceedings of the National Academy of Sciences USA* 91: 2091–2094.
- Sampietro ML, Gilbert MTP, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C. 2006. Tracking down human contamination in ancient human teeth. *Molecular Biology and Evolution* 23: 1801–1807.
- Spigelman M, Lemma E. 1993. The use of polymerase chain reaction to detect *Mycobacterium tuberculosis* in ancient skeletons. *International Journal of Osteoarchaeology* 3: 137–143.
- Spigelman M, Matheson C, Lev G, Greenblatt C, Donoghue HD. 2002. Confirmation of the presence of *Mycobacterium tuberculosis* specific DNA in three archaeological specimens. *International Journal of Osteoarchaeology* 12: 393–401.
- Stojanowski CM, Buikstra JE. 2005. Research trends in human osteology: a content analysis of papers published in the *American Journal of Physical Anthropology*. *American Journal of Physical Anthropology* 128: 98–109.
- Stoneking M. 1995. Ancient DNA: how do you know when you really have it and what you do with it? *American Journal of Human Genetics* 57: 1259–1262.
- Taylor GM, Rutland R, Molleson T. 1997. A sensitive polymerase chain reaction method for the detection of *Plasmodium* species DNA in ancient human remains. *Ancient Biomolecules* 1: 193–203.
- Taylor GM, Widdison S, Brown IN, Young D, Molleson T. 2000. A medieval case of lepromatous leprosy from 13–14th century Orkney, Scotland. *Journal of Archaeological Science* 27: 1133–1138.
- Von Hunnius T, Yang D, Eng B, Wayne JS, Saunders SR. 2007. Digging deeper into the limits of ancient DNA research on syphilis. *Journal of Archaeological Science* 34: 2091–2100. DOI: 10.1016/j.jas.2007.02.007.
- Waite ER, Child AM, Craig OE, Collins MJ, Gelsthorpe K, Brown TA. 1997. A preliminary investigation of DNA stability in bone during artificial diagenesis. *Bulletin de la Société Géologique de France* 168: 547–554.
- Watson JD, Crick FHC. 1953. Molecular structure of nucleic acids. *Nature* 171: 737–738.
- Wiechmann I, Grupe G. 2005. Detection of *Yersinia pestis* DNA in two early medieval skeletal finds from Aschheim (upper Bavaria, 6th century AD). *American Journal of Physical Anthropology* 126: 48–55.
- Yang DY. 2003. Contamination controls and detection in ancient DNA studies. *Acta Anthropologica Sinica* 22: 163–173.
- Zink AR, Reischel U, Wolf H, Nerlich AG. 2002. Molecular analysis of ancient microbial infections. *FEMS Microbiology Letters* 213: 141–147.
- Zink AR, Sola C, Reischel U, Grabner W, Rastogi N, Wolf H, Nerlich AG. 2004. Molecular identification and characterization of *Mycobacterium tuberculosis* complex in ancient Egyptian mummies. *International Journal of Osteoarchaeology* 14: 404–413.
- Zink AR, Nerlich AG. 2005. Notes and Comments. Long-term survival of ancient DNA in Egypt: reply to Gilbert *et al.* *American Journal of Physical Anthropology* 128: 115–118.