Note: this is an expanded version of the entry entitled "Ancient DNA" in the 2002 McGraw-Hill, Yearbook of Science and Technology, pp. 9-14.

The Rise and Fall and Rise of Ancient DNA Studies

Anne D. Yoder and Thomas Delefosse

Department of Cell and Molecular Biology
Northwestern University Medical School
Chicago, IL 60611, USA

Address correspondence to:
Anne D. Yoder
Department of Cell and Molecular Biology
303 E. Chicago Ave., Ward Building
Chicago, IL 60611-3008
e-mail: ayoder@nwu.edu
phone: 312-503-0005
fax: 312-503-7912

The Rise and Fall of Ancient DNA Studies

Within the past decade or so, biochemical methods have advanced to the degree that DNA sequences can be retrieved from surprisingly ancient sources. This means that DNA can be copied and sequenced from individuals representing extinct populations, or even from individuals representing entire species that are now extinct. This practice of retrieving DNA from long-gone sources has typically been dubbed the study of ancient DNA (or aDNA for short). The introduction of the Polymerase Chain Reaction (PCR) into the molecular biologist's tool box is what has made this extraordinary feat possible. PCR is a biochemical technique whereby free nucleotides, DNA primer sequences, and a polymerase (*Taq*) are employed to amplify DNA in vitro. Theoretically, one can begin with a single target DNA molecule and end with millions of copies that can then easily be sequenced with routine methods.

Almost immediately after the introduction of the PCR method, a multitude of investigators began exploiting the technique for amplifying DNA from ancient sources. In the early 1990's, there was an explosion of studies reporting DNA sequences from extinct organisms, some of which were of extraordinary antiquity. The interested reader could find reports of DNA amplified from the Tyrolean Ice Man, ancient mummies, frozen mammoths, a 7000-thousand-year-old human brain, Miocene leaf fossils, and millions-of years-old amber-entombed insects --- just to name a few. Understandably, given this vast range of antiquity and organismic diversity, these studies generated intense interest in the method as well as in the science to which it was applied. This interest even permeated the general public, culminating in the blockbuster movie, *Jurassic Park*.

The bubble of enthusiasm was abruptly burst, however, after the high-profile publication, and subsequent torrent of falsification, of DNA reportedly amplified from a dinosaur bone. The authors of that study reported that they had successfully amplified and sequenced DNA from two 80 million year old bone fragments. They dutifully described the extreme difficulty by which they

had obtained the reported sequences, concluding that because the resulting sequences differed from all other DNA sequences in the then current data base, they must be endogenous to the targeted bone fragments. A phylogenetic analysis of these sequences was not performed. Subsequently, a cadre of independent labs and investigators reported that when a phylogenetic analysis was performed, the "dinosaur" sequences were found to group with humans and other primates, rather than with birds or reptiles as might be expected of authentic dinosaur DNA. Independently, these investigators each concluded that the reported sequences were in fact derived from human mitochondrial DNA (mtDNA)-like nuclear pseudogenes. With the publication of these incriminating studies, a pall fell over the aDNA community --- a pall that was certainly not alleviated by the emerging reports that published sequences for any millions-of-year-old fossils, including all of the amber-entombed insects, were also spurious.

Why is aDNA so Difficult?

The challenges of aDNA studies can be summarized as 1) poor DNA quality and quantity, 2) contamination from exogenous DNA, and 3) PCR inhibition. DNA from extinct organisms is typically fragmented, usually in sizes from 100 - 300 bp, and in very low copy number. Once the mechanism of DNA repair in the living organism is turned off with death, DNA quickly degenerates and is modified by breakage due to loss of bases and/or is rendered inaccessible to enzymatic amplification due to crosslinkage. These characteristics conspire to introduce contamination as a continual nightmare for aDNA researchers. In any PCR amplification, the DNA polymerase will always favor intact contaminating DNA to the desired aDNA. Contamination usually stems from two possible sources: 1) human genomic DNA from handling and preparation of ancient samples and 2) PCR amplicons from labs wherein modern DNA is being investigated.

Low-quality DNA can also result in amplification of artifact sequences. If aDNA is poorly represented in the PCR reaction, modern DNA of any abundance can compete for primer annealing. Mis-priming of modern DNA,

combined with weak priming of target DNA, can result in the generation of artifact sequences early in the reaction that can then be exponentially amplified in subsequent PCR cycles. The result is a chimeric sequence that can mimic what would be expected of an authentically ancient sequence by being similar to, yet different from, DNA from extant organisms.

These problems can be compounded or (even worse) amplification made impossible by the co-purification of PCR inhibitors. Extraction modifications that are favorable to DNA retention are also favorable to the co-purification of PCR inhibitors. Soil-derived degradation products (collectively known as humus) often become associated with a subfossil specimen during the organic decay process. These products can act as strong inhibitors of *Taq* polymerase. Another cause of inhibition results when cross-links between reducing sugars and amino groups occur as a consequence of the Maillard reaction. As an end result, the investigator may not be able to amplify DNA from an ancient extract without really knowing if the lack of amplification relates a lack of target DNA, or simply to the presence of inhibitors.

Ancient DNA Solutions

Contamination control is absolutely necessary to good aDNA technique. To put it bluntly, no investigator should even attempt an aDNA study without the benefit of a specially-designed clean-room laboratory. Figure 1 illustrates an ideal clean room design. This laboratory must be physically separated from the main laboratory and independently accessible. There should be three chambers contained within: an ante-room for donning clean room garb, a small PCR preparation room, and an aDNA extraction room. All three chambers must maintain positive pressure with respect to outside air, and as an additional contamination control measure, the chambers should be routinely exposed to short-wave UV light. By isolating the various phases of aDNA preparation and amplification, the investigator can also identify the precise step wherein contamination might be introduced (as illustrated in Fig. 1).

Although lack of PCR amplification can relate to numerous characteristics of aDNA, inhibition can be unambiguously identified. If adding an aliquot of aDNA extract to a positive control (i.e., "spiking") prevents amplification, then polymerase inhibition is immediately confirmed. Furthermore, there are techniques for identifying the intensity of PCR inhibitors in a given aDNA extract (Fig. 2). By adding progressively greater amounts of an ancient extract to a positive PCR control (e.g., if one were attempting to amplify fossil human DNA, one might choose modern rabbit DNA as the positive control), the investigator can determine the precise amount of extract required to prevent amplification. Therefore, the investigator knows that he must use less than that amount in a PCR reaction to have any hope of successful aDNA amplification.

Primer design is critical for addressing the challenge of poor DNA quality. Whereas modern DNA can be amplified in long sections, with primers conveniently designed to anneal to conserved regions of the genome, aDNA must be amplified in short segments due to its fragmentary nature (Fig. 3). Because one must design primers more frequently, perhaps every 50-100 bases, versus every 500-1000 bases with modern DNA, the investigator always confronts the possibility that the aDNA primers will either be too specific (if they are designed from the sequences of a single organism) or that they will be too general (amplifying numerous organisms, but poorly). Moreover, due to the demand for designing many primers, it is unavoidable that relative efficiency of individual primer pairs will vary. Whereas certain primer pairs will be capable of amplifying a few (or even single) molecules, others will be far less reliable.

Virtually every successful aDNA project to date has focused on mitochondrial DNA (mtDNA). This is the molecule of choice for a very practical reason: most cells possess multiple mitochondria but only a single nucleus. For every single-copy nuclear gene within a given cell, there will be approximately a 1000-fold excess of mitochondrial genes. Given the fact that only very small amounts of total DNA can usually be recovered from ancient specimens, one has

a far greater chance of recovering a mitochondrial gene via PCR, due to their sheer numbers, than a nuclear gene.

Ancient DNA Renaissance

The negative publicity that resulted from the flawed dinosaur study (and all of the other studies to have claimed millions-of-year-old DNA) was sufficient to cast doubt on the method itself. After a period of introspection, a number of aDNA investigators proposed explicit measures for assuring the repeatability of their results (summarized in Table 1). As the ultimate test of authenticity, aDNA researchers agreed that sequences (especially ancient human sequences) should be confirmed via independent procedures in two laboratories that are geographically-separated.

Using these new guidelines, investigators have begun to examine a controversial theory relating to the origins of anatomically-modern humans. It is well-established that *Homo erectus* emerged from Africa, approximately 1.5 million years ago, to establish populations in virtually every corner of the globe. The controversy relates to subsequent events. In one formulation, it is hypothesized that modern human features evolved by slow parallel evolution in *Homo erectus* populations, with genetic continuity maintained by individual migration and occasional interbreeding. This model of human origins is generally referred to as the "multiregional hypothesis". In the opposing, "out of Africa" hypothesis, geneticists have concluded that a wave of anatomically-modern humans emerged from Africa, sometime between 200,000 and 100,000 years ago, to replace their non-modern predecessors. The latter model implies that Neanderthals, although temporarily coincident in time and space with these modern humans, would not (and probably could not) have interbred with them.

Until recently, the choice between these two models depended entirely upon interpretations of the fossil record and on DNA comparisons from living humans. In what was described as a methodological "tour de force", however, Svante Pääbo, Mark Stoneking, and their colleagues characterized a small region

of the mitochondrion directly from the Neanderthal type specimen, originally recovered in western Germany in 1856. The study found that the Neanderthal sequence falls well outside of the range of modern-human mtDNA variation. Moreover, the results were confirmed with phylogenetic analysis, revealing that the Neanderthal sequence is the clear evolutionary outlier to the clade containing modern human mtDNA sequences from all reaches of the planet. This draws the conclusion that Neanderthals were not ancestral too, nor interbreeding with, modern humans.

Another recent study of human fossil DNA has reached an opposing conclusion. Peacock and colleagues determined that mtDNA recovered from a 60,000-year-old anatomically-modern human found in Australia might tell a different story of human origins. These investigators found that DNA from this specimen is quite distinct from extant (and a few fossil) mtDNA lineages found both in Australia and in other parts of the world, including Africa. Rather, mtDNA from this fossil (called LM3) was found to group with a nuclear insert sequence that is generally believed to have diverged prior to the diversification of modern-human mtDNAs. The authors therefore reasoned that their results cast doubt on both the geographic and temporal aspects of the out-of-Africa hypothesis, concluding that if replacement did occur, part of the replacement must have occurred in Australia, and that some of those replaced must have been anatomically modern.

In fact, it is possible to envisage a model whereby the results of these seemingly opposing studies are compatible with a unified replacement theory. In Figure 4, the 1856 Neanderthal specimen and the Australian LM3 specimen are roughly contemporaneous but are drawn from two different human lineages. The 1856 specimen is drawn from the terminal phase of Neanderthal evolution, shortly before replacement by early-modern humans, and the LM3 specimen is drawn from the anatomically-modern lineage, but by chance, represents a mtDNA lineage that is now extinct via the process of lineage sorting. This figure is also meant to suggest that one of the inescapable handicaps of aDNA studies is

that the investigator is entirely the slave of happenstance with regard to choice of samples for analysis. A substantial literature exists that demonstrates the many ways wherein a gene tree (the small lines within the cylinders) and a species tree (the cylinders) may or may not agree. This can be particularly problematic for aDNA researchers because they are constrained to examine relatively recent events wherein lineage sorting is still occurring. Due to the effects of differential mtDNA lineage extinction, the investigator may or may not happen across the mtDNA alleles that precisely track the history of the species or population under investigation.

Despite the manifold methodological and theoretical limitations to aDNA studies, the technique will continue to reveal historical events that are beyond the scope of contemporary comparative studies. Ancient DNA techniques are currently being applied successfully for understanding everything from ancient pathologies (e.g., malaria's role in the fall of Rome; the existence of tuberculosis in the New World prior to the arrival of Europeans; the precise culprit in the Black Death) to the reconstruction of population origins and subsequent migrations, to the determination of relationships among extinct and extant organisms. These studies will continue to amaze us, both for their technical wizardry and for the secrets that they reveal.

Bibliography

Core Texts:

Adcock, G. J., E. S. Dennis, S. Easteal, G. A. Huttley, L. S. Jermiin, W. J. Peacock, and A. Thorne. 2001. Mitochondrial DNA sequences in ancient Australians: Implications for modern human origins. Proc Natl Acad Sci U S A 98:537-542.

Krings, M., A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking, and S. Pääbo. 1997. Neanderthal DNA sequences and the origin of modern humans. Cell 90:19-30.

Woodward, S. R., N. J. Weyand, and M. Bunnel. 1994. DNA sequence from Cretaceous period bone fragments. Science 266:1229-1232.

Zischler, H., M. Hoss, O. Handt, A. von Haeseler, A. C. van der Kuyl, J. Goudsmit, and S. Paabo. 1995. Detecting dinosaur DNA. Science 268:1191-1193.

Additional Reading:

Austin, J. J., A. B. Smith, and R. H. Thomas. 1997. Palaeontology in a molecular world: the search for authentic ancient DNA. Trends Ecol. Evol. 12:303-306.

Maddison, W. P. 1997. Gene trees in species trees. Syst. Biol. 46:523-536.

O'Rourke, D. H., M. G. Hayes, and S. W. Carlyle. 2000. Ancient DNA studies in physical anthropology. Ann. Rev. Anthrop. 29:217-242.

Pääbo, S. 1993. Ancient DNA. Sci. Amer. 5:86-92.

Stoneking, M. 1995. Ancient DNA: how do you know when you have it and what can you do with it? American Journal of Human Gentics 57:1259-1262.

Wayne, R. K., J. A. Leonard, and A. Cooper. 1999. Full of sound and fury: the recent history of ancient DNA. Annu. Rev. Ecol. Syst. 30:457-477.

Table 1. Methods for verifying aDNA authenticity

• Prepare independent extracts from the same sample in two labs, geographically separated
• Prepare multiple PCR reactions for each extract
• Require several positive tubes for each PCR reaction
• Test multiple clones per positive tube
• Employ species-specific primers, if possible
• Design amplification and sequencing strategy that allows for contig assembly from multiple overlapping fragments
• Always include negative extraction and PCR controls
• Continuously apply painstaking contamination control methods

Figure Legends

Figure 1

Ideal clean room design (see text for discussion). Figure also illustrates use of multiple negative controls, exposed to different stages of ancient DNA preparation, for monitoring contamination.

Figure 2

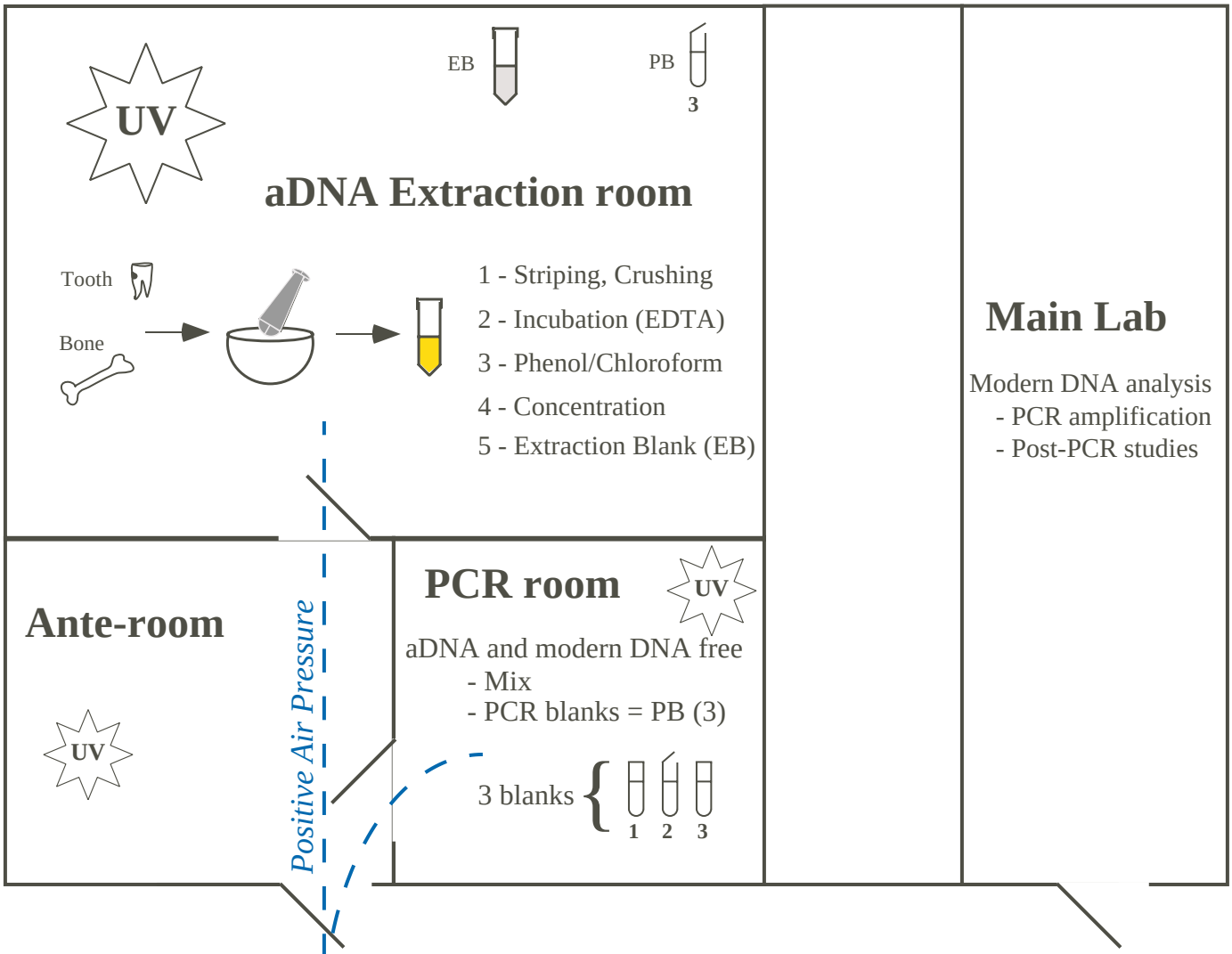
Illustration of progressive spiking of a positive PCR control for assessing the degree of PCR inhibition. At 10 ul., amplification of control is completely inhibited.

Figure 3

Comparison of PCR amplification strategy for contemporary versus ancient DNA. Drawing on top represents long strand of contemporary DNA wherein only two primers (shown as bent arrows) are necessary to amplify entire region of interest. Drawing below indicates that because ancient DNA is so fragmented, many more primers are required to amplify equivalent region.

Figure 4

Possible model wherein the ancient DNA results from a Neanderthal (1856) and a 60,000-year-old modern human from Australia (LM3) are both consistent with a replacement hypothesis (see text for discussion). "X" represents mtDNA lineage extinction, either through natural process of lineage sorting or through species extinction. Lineage sorting occurs because mtDNA is maternally inherited. When a woman has only sons, she does not pass her mtDNA on to subsequent generations; if she has no maternal relatives, or those who also have only sons, the entire mtDNA lineage can become extinct though her descendants live on.



EB = Extraction Blank

Extraction Blank → test extract reagents (EDTA, PK, Sarcosyl)

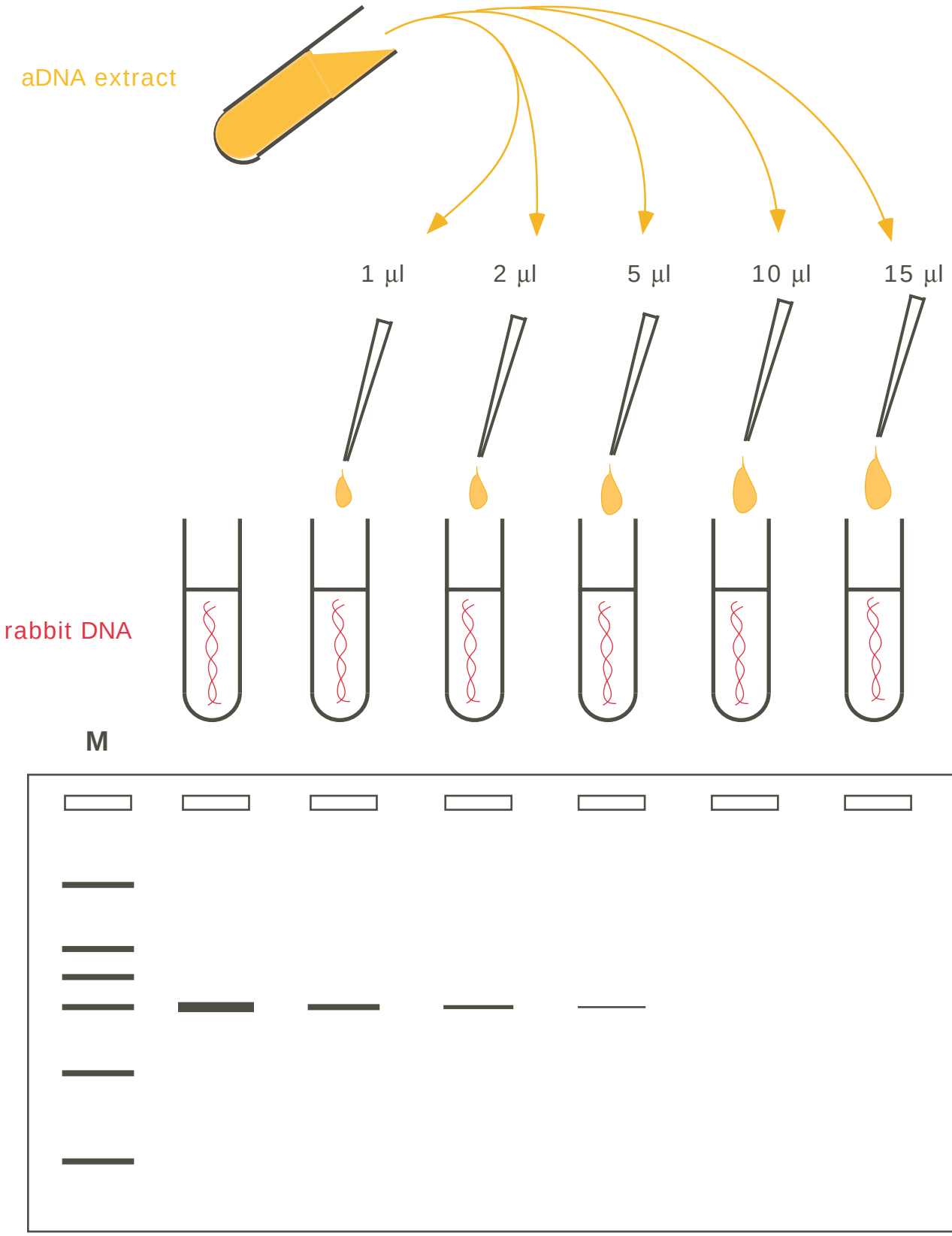
3 PB = 3 PCR blanks

1 PCR reagents only → test PCR reagents

2 PCR reagents only, open in PCR room → PCR room air test (aerosol)

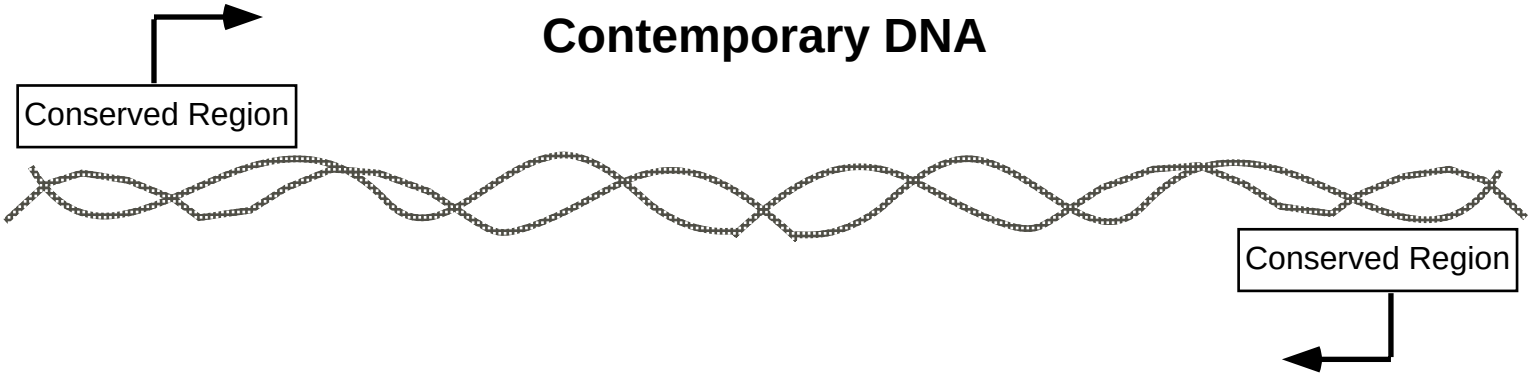
3 PCR reagents only, open in extraction room → Extraction room air test (aerosol)

Yoder and Delefosse, Fig. 1

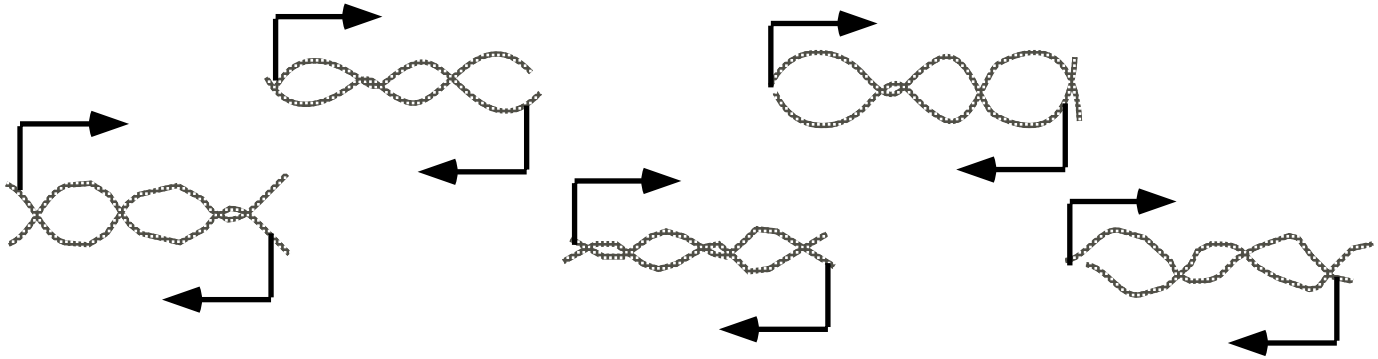


Yoder and Delefosse, Fig. 2

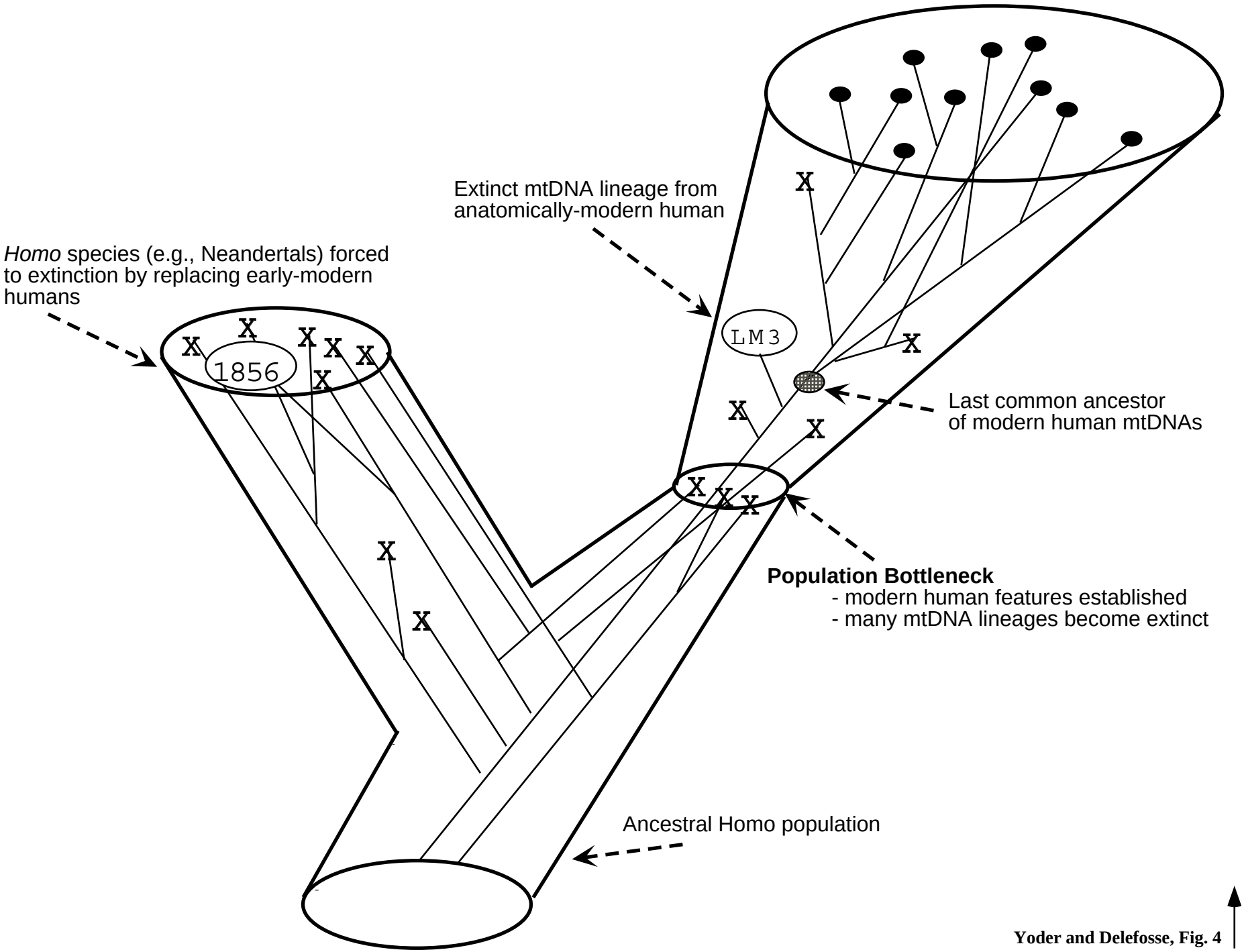
Contemporary DNA



Ancient DNA



Yoder and Delefosse, Fig. 3



Yoder and Delefosse, Fig. 4