POSSIBILITIES AND LIMITATIONS IN THE ARCHAEOGENETIC ANALYSIS OF ANCIENT HUMAN REMAINS

B. G. MENDE

Laboratory of Archaeogenetics

Archaeological Institute of the HAS

Uri u. 49, 1014, Budapest - Hungary

Tel. +36 1 224 67 80 - Fax: +39 224 6719

E-mail: mende@archeo.mta.hu

Abstract

Archaeogenetic investigations – parallel to the wide expansion of molecular genetics – have recently gained importance in archaeology and population history. This positive change in its role in historical research is based partly on the technological development of the last two decades and partly on the recognition of the fundamental conservatism of the DNA.

The field was open for a multi-respect analysis of the DNA of several thousand years old human remains. The polymorphism of the DNA, especially on certain sections of the mitochondrial DNA, offered the possibility for the most thorough examination ever in relation to the spread and genetic variability of the human species. Because of the fragmentary character of the preserved ancient DNA sections, the morphometric features of the human skeleton and the genetic haplogroups formed by the DNA-based polymorphism cannot be correlated. In fortunate cases, archaeogenetic investigations make it possible to study illnesses of genetic origin or analyse kinship relations in smaller burial groups. The examination of patrilineal and autosomatic inheritance can be of great help in answering the major questions of the population history of the Carpathian Basin.

The most recent investigations concentrate on the testing of archaeological and historical preconceptions regarding the eighth to twelfth centuries, with special emphasis on the problem of population and ethnic group. The most spectacular results, however, can be expected in relation to the population problems of the Neolithic. The archaeogenetic laboratory in the Archaeological Institute of the Hungarian Academy of Sciences was established to answer these questions.

The long-term financing of such investigations, however, has not been solved properly, not least because of the time needed and the difficulties one has to face when trying to provide an interdisciplinary interpretation. In order to decrease the possibility of modern human DNA contamination, there is a recent tendency to limit the number of research groups and focus the limited resources – among them the available grants and funds – in the major archaeogenetic research centres.

KEYWORDS: ANCIENT DNA, MITOCHONDRIAL POLYMORPHISM, CONTAMINATION, ARCHAEOGENETIC RESEARCH

KULCSSZAVAK: DNA, MITOKONDRIÁLIS POLIMORFIZMUS, SZENNYEZŐDÉS, ARCHEOGENETIKAI KUTATÁS

Introduction

During the second half of the last century it became a fundamental paradigm that answering certain questions requires a close co-operation between various branches of research. Among humanities, one of the most imminently affected discipline was archaeology, that embraced a host of scientific methods in addition to its traditional auxiliary or sub-disciplines, such as physical anthropology, pedology and geology, archaeobotany and archaeozoology. A range of material studies and various dating methods were introduced into archaeological research that have formed a new trend. By the 1960s and 1970s, „New Archaeology" gained a decisive role in archaeological research in the „Western World". The interpretation and proportion of results yielded by natural sciences, however, were often misrepresented in the conclusions drawn by archaeologists and historians due to their schooling, approach and the insufficient knowledge of the

HU ISSN 1786-271X; urn:nbn:hu-4106

faults and limitations of the methods. In hindsight, it may be considered inevitable that trust in the direct application of scientific results as panacea for archaeological problems eventually dissipated. As a predictable backlash, strong criticism and a sceptical attitude toward results gained by using new scientific methods became commonplace in many historical-archaeological treatises. The application of scientific dating methods suffered most from this criticism, although classical physical anthropology, trying to rejuvenate itself by the extensive use of biostatistics, has also been bruised.

Revolutionary changes that took place in molecular biology and biotechnology from the 1970s onwards, have lent a new momentum to interdisciplinary research in archaeology. The most evident sign of this was that archaeogenetics, that has since developed into a new, independent field of research, was included into archaeological investigations. The appearance of molecular genetic profiling may be traced back to three main reasons. These were as follows:

- the increasingly precise understanding of the function and structure of DNA, that lead to the recognition of its fundamental conservatism,
- the development and availability of the polymerase chain reaction (PCR) technology,
- the recognition of DNA preservation capacity of bones from archaeological contexts.

The emergence and rapid proliferation of archaeogenetic investigations in international scientific research are largely the result of the more-or-less simultaneous application of DNA based methods using archaic human and animal remains, and research in forensic medicine, criminology and various areas of biotechnology. Results of global importance have been born in the shadow of the much publicised human genome project (HUGO), stem cell research and tumour genetics. These included determining the complete base sequence of mtDNA and the description of polymorph characteristics of the non-coding region, as well as the identification of the genetic and geographical origins of the human species, which helped refute the theory of multiregional evolution. The maternal and paternal lines of inheritance could be mapped, the tree of haplogroups was drawn and the process of settling the various continents could be clarified. Meanwhile, a consensus was reached in professional circles as to the technology and laboratory environment required for the isolation and amplification of ancient DNA stock. An increasing number of research teams started publishing archaic mtDNA sequences, and experience acquired in other areas of genomics have also been successfully applied. These included the criteria for grouping, the methods of establishing chronological sequences for mutations as well as the clarification of origins of and relationships between living human populations. As of today, sufficient knowledge has been accumulated to state that the successful application of DNA studies is limited to at most ten thousand years. It seems, however, that within this time interval only the remains of people buried during the last two or three millennia can be studied on a regular basis, in spite of the fact that the DNA preservation and consequent analytical potential of bones is less dependent on the absolute time of deposition than on the taphonomic effects of micro-environmental factors in the deposit.

Studying archaic DNA samples has had its own childhood diseases. The most important of these has been the problem of contamination by modern human DNA. Strict laboratory protocols have been developed in order to exclude false positive results. These include the complete spatial isolation of the pre-PCR and post-PCR phases of processing, as well as the parallel analysis of samples. In recent years, cloning has gained increasing importance in identifying contamination by modern DNA. Just for general information: if the DNA content of an archaic bone sample recovered under average circumstances is taken as one unit (1, in a mathematical sense), the number of copies obtained after amplification may be on the order of millions or even tens of millions (10^{6-7}). This order of magnitude corresponds to the concentration found in the fresh, live tissue. However, recently unrealistic requirements have been put forward in connection with cloning to the detriment of the publication possibilities of small laboratories and research teams. This is particularly worth mentioning here, since there is a general consensus that DNA obtainable from fossil bone remains tends to be heavily fragmented. The size of such fragments does not exceed a few hundred pairs of bases. Should it be possible to amplify segments longer than this, one should always be aware of the risk of contamination.

Archaeogenetic research in the Carpathian Basin

Archaeological investigations in the Carpathian Basin have built their chronologies on theoretically assumed population changes, reconstructed on the basis of material remains and written sources. Since the precision of the latter in reconstructing events in population history is limited, the methods of archaeogenetic investigations complementing the results of traditional physical anthropological studies are of prominent importance. In order to facilitate complementary research carried out in this field, a research agreement was signed in 2001 between the Archaeological Institute of the

HU ISSN 1786-271X; urn:nbn:hu-4106

Hungarian Academy of Sciences and the Institute of Genetics, Biological Research Center (Szeged). This co-operation has been supported by state-sponsored grants. In order to increase the efficiency of research work and to create a database, a PCR laboratory was created in the Archaeological Institute of the Hungarian Academy of Sciences.

The construction of this laboratory took place between 2002 and 2004. The design of working areas was strictly defined by the protocol requirements of isolating and typing archaic DNA. The processes of isolation and the PCR phase had to be separated as much as possible. Therefore, the so-called preparation room, used in the isolation of archaic DNA and the PCR processing area are located at the opposite ends of a corridor, with a general preparation room in between. An isolated space within this latter room is maintained for polishing and pulverising ancient human bone.

Our investigations, carried out in co-operation with the research team of the Biological Research Center (BRC) have shown that the mtDNA sequences identified in the AD 10th century samples belong to Asiatic haplotype groups in a far greater proportion than those taken from modern Hungarian populations. Currently, our data base contains the sequences of 70 archaic samples drawn from 120 graves representing the 8-12th century populations of the Carpathian Basin. Recently, our research has targeted the amplification of our data base both in time and space, the mapping of the paternal lines of inheritance and the testing of the relevance of using autosomal markers in our studies. In addition to human bone samples, the analysis of animal remains has also been carried out in the BRC. To date, mtDNA sequences have been obtained for horse, cattle and sheep.

There is, however, a period in the prehistory of the Carpathian Basin which is of essential significance from the viewpoint of the population history of the entire continent of Europe: the period of the Early and Middle Neolithic. Investigating the spread and directions of Neolithization using methods of archaeogenetics is one of the hottest research subjects in European archaeology. From this point of view, it is of special interest that, on the maternal side, the overwhelming majority of populations inhabiting modern Europe originates from people who lived in this area already in Palaeolithic times. Only one of the seven main haplogroups characteristic of modern European populations is of Neolithic origins. The basic question of how this Neolithic group spread and exerted its population genetic effects can and should be studied in the area bordered by the northern Balkans, the Alps and the Carpathian Mountains.

Limitations of archaeogenetic analysis

As mentioned before, the application of scientific results in the interpretation of historical processes is hampered by major difficulties. It is therefore necessary to face the limitations which *ab ovo* determine the direct applicability of results obtained by archaeogenetic research. These difficulties may be summarised as follows:

- limitations on conclusions and difficulties of interpretation and chronology;
- limitations posed by the "inaccuracy" of databases;
- limitations diachronic and taphonomic of DNA preservation, depending on the microenvironment;
- limitations of technology and financing research.

The difficulties of interpretation are mainly due to the fact that samples subjected to genetic analysis already represent an archaeological-historical, consequently also chronological, preconception. Deciding, who among the AD 10th century population should be considered first generation, "conquering" Hungarian, is currently determined by social scientific methods. However, bone samples taken from graves or sets of burials selected and characterised from an archaeological point of view, by definition "absorb" uncertainties of these selection criteria. Experts in genetics and biostatistics will treat them as unambiguous raw data.

When evaluating our results, it is also important to consider the relationships between the databases at our disposal. Such databases shall be evaluated according to geographical and chronological aspects. One may say in general terms that most reference databases are built on "modern" samples. This means that the information concerning the population of a given area either has little or no time depth, or offers possibilities of interpretation of extremely long time spans, as is the case with mtDNA haplogroups. This is clearly exemplified by the comparison between modern Hungarian "samples" and populations inhabiting geographical areas considered to be of outstanding importance from the viewpoint of ancient Hungarian history (regardless of blood typing or other genetic characters). This solution is understandable, considering the scarcity of relevant data representing this historical and archaeological period. On the other hand, one cannot ignore the fact that populations included in this comparison also have their own "prehistories" with their own chronological and spatial dimensions. In other words, the sample used in comparisons can be considered "constant" only in exceptional cases.

HU ISSN 1786-271X; urn:nbn:hu-4106

Even using the ample evidence of historical, archaeological and linguistic data, populations cannot be characterised as precisely as would be necessary to support the conclusions drawn from them.

Similarities and differences between the patterns identified on the basis of the studied genetic traits may be analysed descriptively adhering to certain rules. It is a fundamental problem, however, that our possibilities are extremely limited in answering basic questions such as the sources and especially the timing of the differences between populations. The starting point of our investigations has been that the differences between the Asiatic haplotype distributions in AD 10th century populations in the Carpathian Basin and modern Hungarian populations was the effect of the 10th century conquering Hungarians of Asiatic origin. However, in order to correctly evaluate the changes of the haplotype patterning of the populations within this geographical area, one must also consider the possibility of earlier and later occurrences of Asiatic genetic elements. We should not forget that our present investigations have not been able to prove the exclusively AD 10th century origin of the Asiatic character of the maternal line of modern Hungarians. Data representing several consecutive phases are required to eventually understand the processes of change and draw realistic conclusions concerning the processes of population history.

The scientific success of this research is strongly influenced by a special contradiction. Natural scientists are under constant pressure to produce and publish new ideas, while historical studies require the accumulation of databases of representative sizes, made possible only by long term investigations. Frequently, this contradiction is not as much the product of differing disciplines and their differing paradigms, but rather of incompatible financing strategies and often the lack of appropriate research plans. Molecular genetics undoubtedly needs the freedom to interpret its results within its own frame of reference. This can be achieved by analysing a relatively small number of samples. Drawing historical conclusions and reconstructing the underlying processes, however, require larger series. Should an outlier occur within the mtDNS patterning of a certain time period, it is not sufficient to interpret it as a consequence of a genetic anomaly. It is for this reason that the division of research between our archaeogenetic laboratory and the research team based in Szeged has resulted in a fruitful co-operation. The development of methods and study of primers has been taking place in the Institute of Genetics, while the Archaeological Institute is in charge of studies of historical relevance that require the time consuming creation of major databases.

One should not underestimate the significance of sampling problems either. This is especially important when we do not have an opportunity to select the best of several samples. To date, research has shown that the relationship between the DNA preservation potential of bones (and the related success of analysis) and time of deposition is not as important as the effects of micro- and macro-environmental factors, burial rite and isolation that all influence the success of the PCR reaction. For example, bone remains of the Sarmatian population are notoriously poorly preserved in the Carpathian Basin, therefore little is known of the morphological traits of these people. It is possible that the molecular genetic profile of this population will be similarly difficult to establish to an extent that could meaningfully contribute to answering questions of population history. It is especially frustrating when sampling bones from burials of outstanding chronological importance turns out to be unsuccessful. It is also problematic to evaluate the remains of children. In the case of juvenile skeletons, which already contain only a negligible amount of cortical bone, it is more difficult to take non-contaminated samples of reasonable quantity to begin with. This has a fundamental bearing upon the success of isolation and amplification. Meanwhile, in order to clarify internal kinship relations and chronological sequences within groups of burials, one should know their genetic affiliations as well. An additional difficulty is posed by the unfortunate fact that many physical anthropologists are hesitant when destructive sampling techniques must be used, even on a relatively small scale. This is in spite of the fact that a consensus has been developing in the international literature concerning the quantitative and technical standards of proper sampling.

Hopefully, in addition to the critical comments put forward in this presentation, I could also direct your attention to the actual importance of archaeogenetic research itself. As far as I am concerned, I consider it a special honour and great luck that, as an historian trained in biological anthropology, I have had the opportunity to witness and actively promote the establishment of this typically interdisciplinary project in an archaeological institution.

Selected literature

Anderson S, Bankier AT, Barell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG, (1981): Sequence and organization of the human mitochondrial genome, *Nature* 290(5806):457-465.

Bogacsi-Szabó E, Kamlár T, Csányi B, Tömöry Gy, Czibula Á, Priskin K, Horváth F, Downes CS,

Raskó I (2005): Mitochondrial DNA of ancient Cumanians: Culturally Asian Steppe Nomadic inmigrants with substantially more western Eurasian mitochondrial DNA lineages. *Hum Biol* 77 (59), 629-652.

Meyer E. (2000): Extraction and amplification of authentic DNA from ancient human remains, *Forensic Science International* 113, 87–90.

Paabo S. (1989): Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification, *Proc Natl Acad Sci USA*. 86(6), 1939-1943.

Rebecca L., Brown WM, Wilson AC (1984): Polymorphic sites and the mechanism of evolution in human mitochondrial DNA, *Genetics* 106, 79-499.

Renfrew, C., Boyle, K. ed. (2000): Archaeogenetics: DNA and the population prehistory of Europe.

Richards, M. et al.(1996) : Paleolithic and neolithic lineages in the European mitochondrial gene pool, *Am J Hum Genet* 59(1), 185-203.

Stone AC, Milne GR, Paabo S, Stoneking M (1996): Sex determination of ancient human skeletons using DNA., *Am J Phys Anthrop* 99(2):231-8.

Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tänzer M, Villems R, Renfrew C, Gronenborg D, Alt KW, Burger J (2005): Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites, *Science 310* (5750), 1016-1018.

Thomas M, Gilbert P, Bandelt HJ, Hofreiter M, Barnes I (2005): Assessing ancient DNA studies, *Trends in Ecology & Evolution* 20, 541-544.

Salas A, Carracedo A, Macaulay V, Richards M, Bandelt HJ (2005): A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics. *Biochem Biophys Res Commun* 2005; 335(3): 891-899.

HU ISSN 1786-271X; urn:nbn:hu-4106