

NEW TRENDS OF ARCHAEOGENETIC RESEARCH IN THE INSTITUTE OF ARCHAEOLOGY, RESEARCH CENTRE FOR THE HUMANITIES, HUNGARIAN ACADEMY OF SCIENCES

ANNA SZÉCSÉNYI-NAGY – BALÁZS GUSZTÁV MENDE¹

Molecular genetic investigations of human and animal remains containing genetic material from ancient times focus on questions such as genetic connections and their types between certain persons or groups, genetic origin and relatedness of populations from different geographic regions and time periods, molecular evolution of pathogens and their related epidemiological issues. Since technologies used in the study of archaic samples today are based on the highest level of molecular genetic methodology and laboratory facilities and the extracted ancient DNA data are also becoming better and more informative, their evaluation requires an increasing knowledge of bioinformatics. Interpretation in the archaeological and historical contexts can only be complete with a high-level cooperation of the research fields concerned. In the followings, we present the new research directions of the Laboratory of Archaeogenetics (Institute of Archaeology, RCH HAS), which clearly show the collaboration and cooperation scheme among the various fields of bioarchaeology and archaeology.

Archaeogenetics, also named as palaeogenetics, is a recently established discipline that evolved parallel with the discovery and spreading application of polymerase chain reaction (PCR). PCR is an enzymatical reaction that amplifies low DNA content typical to ancient remains of organisms, and enables to detect and sequence such amplicons. For twenty years, PCR-based analysis types served archaeogenetics which expanded rapidly in the 1990s. Researchers of the field, however, became divided by the quality and credibility of the mushrooming number of ancient DNA papers.² The criteria of authenticity of ancient DNA results and the possible contamination of the described sequences by modern DNA were the most debated issues.³ The methodological development of DNA extraction and the canonization of laboratory protocols had allowed by the end of the 2000s for the large-scale typing of ancient samples through PCR and Sanger-type sequencing.⁴ These mass studies, however, were only focused on e.g. the very short fragments (few times 100-200 nucleotides) of the nearly 3 billion nucleotide pairs of the human genome. The second revolution of archaeogenetics started around 2005, when next generation sequencers (NGS) were first tested for ancient DNA. The first archaic human genome was sequenced by 2010,⁵ and by now archaeogenomics or palaeogenomics has become a distinct academic area.

Even today, next generation random (shotgun) sequencing of whole ancient genomes is very expensive. Despite the fact that the constantly evolving DNA extraction and DNA library construction techniques improve a lot the expense/income margin, very few ancient human genomes (<100) have been sequenced in good coverage so far. Due to the fragmented depth of the sequence coverage and the still unknown function of genomic regions, a significant part of the total length of sequenced human genomes

¹ Hungarian Academy of Sciences, Research Centre for the Humanities, Institute of Archaeology

² Stoneking, Mark: Ancient DNA: how do you know when you have it and what can you do with it? *Am. J. Hum. Genet.* 57 (1995), 1259–1262; Cooper, Alan: Reply to Stoneking: ancient DNA – how do you really know when you have it? *Am. J. Hum. Genet.* 60 (1997)/4, 1001–3.

³ Cooper, Alan – Poinar, Henrik N.: Ancient DNA: do it right or not at all. *Science* 289 (2000), 530–531.

⁴ Brandt, Guido et al.: Ancient DNA reveals key stages in the formation of Central European mitochondrial genetic diversity. *Science* 342:6155 (2013), 257–61. Doi: 10.1126/science.1241844.

⁵ Rasmussen, Morten et al.: Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463 (2010), 757–762. doi:10.1038/nature08835.

remains crude in the research of population genetic events and population histories. Beside the random sequencing approach, variants of NGS have been developed that are targeted at the investigation of a mass number of nucleotides (SNP-single nucleotide positions) in the human genome.⁶ This method provides sufficient information to detect migration or admixture of populations in prehistory.⁷ However, population movements at the end of prehistory, in the Roman era and the Migration Period mixed up the population of Europe to such a degree that the differences between the gene pools of peoples living in geographical proximity decreased considerably. Therefore, archaeogenetic research focusing on the early Middle Ages seeks rare variants in completely sequenced recent and archaic genomes, which allow for, e.g., a more accurate estimate of the genetic heritage of the Anglo-Saxon migration (5th–7th centuries) in Britain today.⁸

The Institute of Archaeology of the Research Centre for the Humanities, Hungarian Academy of Sciences is involved in the above outlined research trend of archaeogenetics and archaeogenomics in several ways. In 2004 a laboratory for archaeogenetics was established in the institute, specifically for the genetic analysis of human and animal skeletal remains from archaeological contexts. The grants from the 5th NKFP Framework Program, coordinated by Professors István Raskó and Csanád Bálint, as well as other academic funding covered the expenses of the laboratory equipment which increased significantly in the 2010s through infrastructure development awards by HAS. Today, the facilities of the well-equipped, two-unit (pre- and post-PCR) laboratory allows the application of NGS technologies (*Figs 1, 2*). Using the Illumina MiSeq sequencer purchased in autumn 2015 (*Fig. 3*), several millions of nucleotides can be studied in one sequencing run with the NGS technology, instead of the previously analyzed few hundred nucleotides

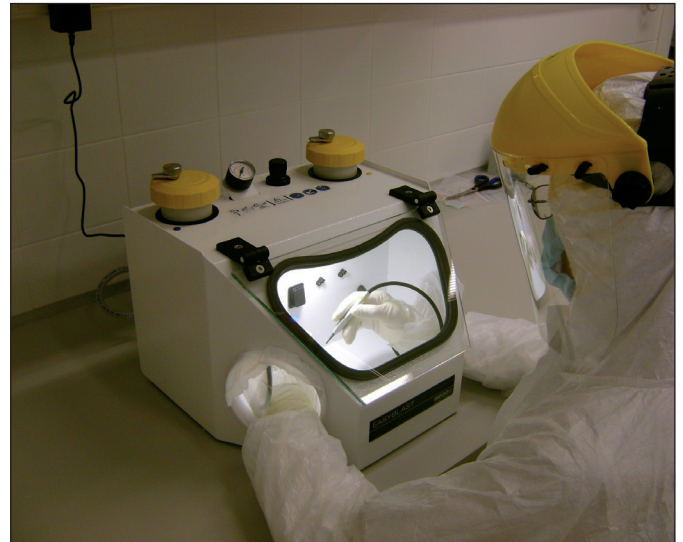


Fig. 1: *Cleaning of the ancient bone and tooth samples by sandblasting in the Laboratory of Archaeogenetics of the IA RCH HAS*



Fig. 2: *DNA library construction takes place under a PCR box in the pre-PCR laboratory*



Fig. 3: *Illumina MiSeq desktop sequencer*

⁶ Mathieson, Ian et al.: Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528 (2015), 499–503. doi:10.1038/nature16152.

⁷ Haak, Wolfgang et al.: Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522:7555 (2015), 207–11. Doi: 10.1038/nature14317.

⁸ Schiffels, Thomas et al.: Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat. Commun.* (2016)/7, 10408. Doi: 10.1038/ncomms10408.

(e.g. mitochondrial DNA control region). Due to the multiple readings of the DNA strand, it is no longer required to authenticate the results by transforming the DNA molecules into bacteria, cloning these DNA fragments, and then resequencing the colony PCR products. The duration of laboratory work dropped significantly, due to the omission of cloning; the evaluation of the data, however, requires a higher level of bioinformatic knowledge and skills than before.

The archaeogenetic part of the German-Hungarian bioarchaeological research project, funded by the German Research Foundation between 2010–2014, focused on a section of the mitochondrial DNA that is ca. 400 nucleotide pairs in length (Fig. 4).⁹ Out of the nearly 600 Neolithic samples pre-filtered and processed under this program, samples with the best DNA conditions shall continue to be examined in our laboratory, in cooperation with the Department of Genetics, Harvard Medical School, Boston and the Max Planck Institute for the Science of Human History, Jena. The Hungarian lab performs sample preparation, DNA extraction and DNA library preparation; that is, all the steps that require a pre-PCR laboratory. In “DNA libraries” we archive the DNA, providing DNA strands with special molecular barcodes.¹⁰ These libraries can serve as a basis for repeated sequencing analyses or sequencing of different DNA components of the samples (e.g. bacterial or human DNA). We extended the new genomic research to an additional 150 Copper Age and Bronze Age human bone remains, thus covering 4,000 years of prehistory in the central part of the Carpathian Basin. Our aim, on the one hand, is the complete mapping of the mitochondrial genome (16,569 instead of 400 nucleotides pair long), in order to provide far more accurate mitochondrial haplogroup definitions. In addition to the maternally inherited mitochondrial lineages, we also map the paternal Y chromosome lines in high resolution. DNA samples passed through the multi-stage filter are selectively tested for the so-called hybridization capture of 1,240 thousand SNP at the Harvard Medical School. In the last phase of the research, we plan to

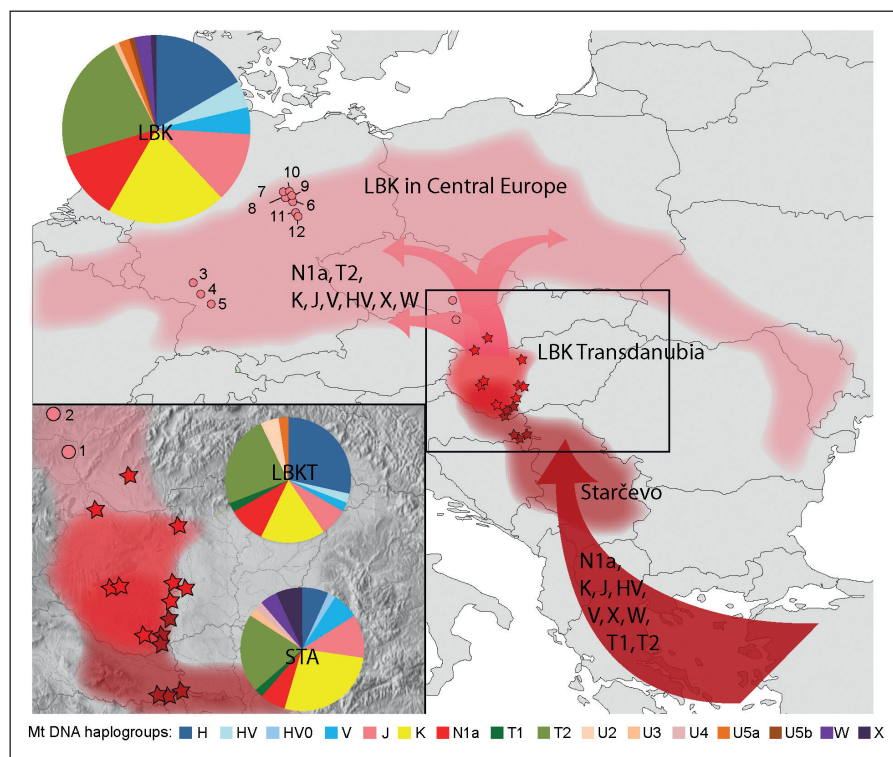


Fig. 4: The mitochondrial “Neolithic package” and the Continental Route of the Neolithic dispersal. The shaded areas of the maps show the spread of the Starčevo Culture, the LBK (Linearbandkeramik) in Transdanubia and in Central Europe. The arrows show the direction of the farmers’ migration into Central Europe, indicated by archaeological and genetic records. Stars mark the archaeogenetically studied LBK and Starčevo sites (after Fig. 12 in Szécsényi-Nagy 2015).*

⁹ Bánffy, Eszter: [German-Hungarian bioarchaeological research project in the Archaeological Institute of the Research Centre for the Humanities, Hungarian Academy of Sciences](#). *Hungarian Archaeology* (2013)/Summer; Szécsényi-Nagy, Anna et al.: Tracing the genetic origin of Europe’s first farmers reveals insights into their social organization. *Proc. Roy. Soc. B*, 282. (2015), 20150339. DOI: 10.1098/rspb.2015.0339.

¹⁰ Rohland, Nadin et al.: [Partial uracil–DNA–glycosylase treatment for screening of ancient DNA](#). *Phil. Trans. R. Soc. B*. 370. (2015).

* Szécsényi-Nagy, Anna: [Molecular genetic investigation of the Neolithic population history in the western Carpathian Basin/Molekulargenetische Untersuchungen zur Bevölkerungsgeschichte des Karpatenbeckens](#). Dissertation uploaded to ArchiMeD. (2015) Mainz, Germany.

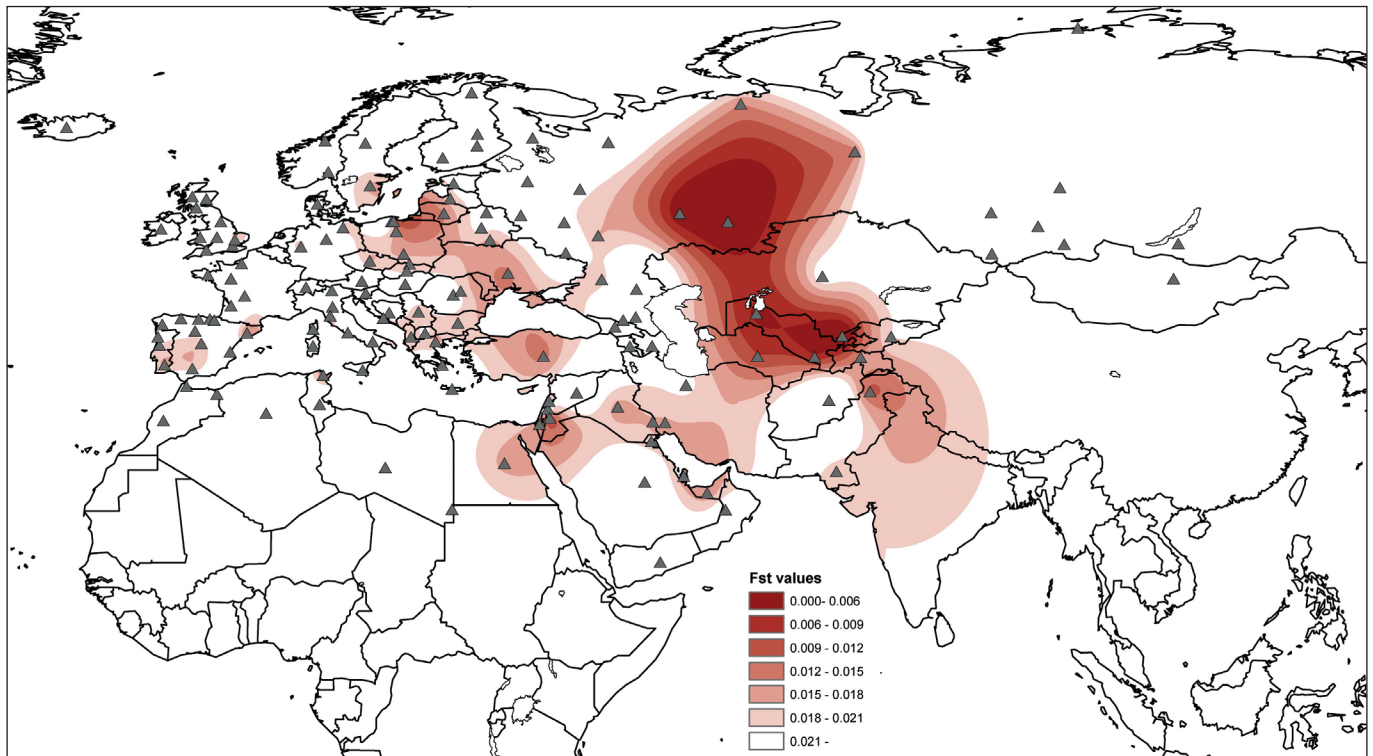


Fig. 5: Genetic distances between the 10th-century Hungarians (n=102) and 144 modern populations, calculated from mitochondrial HVR-I sequences (after Csőszy et al. 2016)**

perform full genome sequencing of the samples containing endogenous DNA in a good state.

From the beginning, the Institute of Archaeology of the RCH HAS lays emphasis on the research of Hungarian ethnogenesis. Mitochondrial DNA typing was carried out primarily on 6th–11th century anthropological findings from the Carpathian Basin (Fig. 5).¹¹ We examined the eastern relations of the Conquest period population through mitochondrial DNA analyses of the 8th-century Saltovo culture's tombs.¹² Our research involved cemeteries of the Slavic-Hungarian population that lived at the fringes of the conquering Hungarians' settlement areas. The 10th–12th century cemetery at Nitra-Šindolka was analyzed in collaboration with the Constantine the Philosopher University, Nitra and the Selye Janos University, Komarno. The fourteen successfully typed individuals buried at Nitra-Šindolka, completed with eight persons excavated from the 9th–12th century cemetery at Čakajovce, testify to a typically medieval mixed type of mitochondrial gene pool, mostly close to the maternal genetic composition of today's Europeans.¹³ Instead of sequencing a section of the mitochondrial DNA, our research in Hungarian prehistory has a new emphasis on sequencing complete mitochondrial genomes. This detailed mapping allows us to define mitochondrial groups widely spread in different geographical regions, such as the most common H haplogroup in today's Europe that also

¹¹ Csőszy Aranka – Mende Balázs Gusztáv: Beszámoló a Kárpát-medence 7–10. századi népességén végzett archeogenetikai vizsgálatok eredményeiről (7th-10th century population of the Carpathian Basin – reviewing from the perspective of the archaeogenetic results). In: *Magyar Őstörténet. Tudomány és hagyományörzés* (Hungarian prehistory, science and traditions), szerk. Sudár Balázs – Szentpéteri József – Petkes Zsolt – Lezsák Gabriella – Zsidai Zsuzsanna (Budapest: MTA BTK, 2014), 173–179.

¹² Csőszy, Aranka – Mende, Balázs Gusztáv – Langó, Péter: Arheogenetičeskie issledovani â materialov saltovskoj i drevne-vengrskoj kul'tur. In: *Centr pam'jatkoznavstva NANU Starožitnosti Livoberežnogo Podniprov'â/Antiquities of the Dnieper Left Bank Regions* (Kiev: Institut Arheologii Nacional'noi Akademii nauk Ukraini, 2012), 94–101.

¹³ Csáky Veronika et al.: Maternal Genetic Composition of a Medieval Population from a Hungarian-Slavic Contact Zone in Central Europe. *PLoS ONE* 11(3): e0151206. doi:10.1371/journal.pone.0151206. (2016)

** Csőszy, Aranka – Szécsényi-Nagy, Anna et al.: Maternal genetic ancestry and legacy of 10th century AD Hungarians. In preparation (2016)

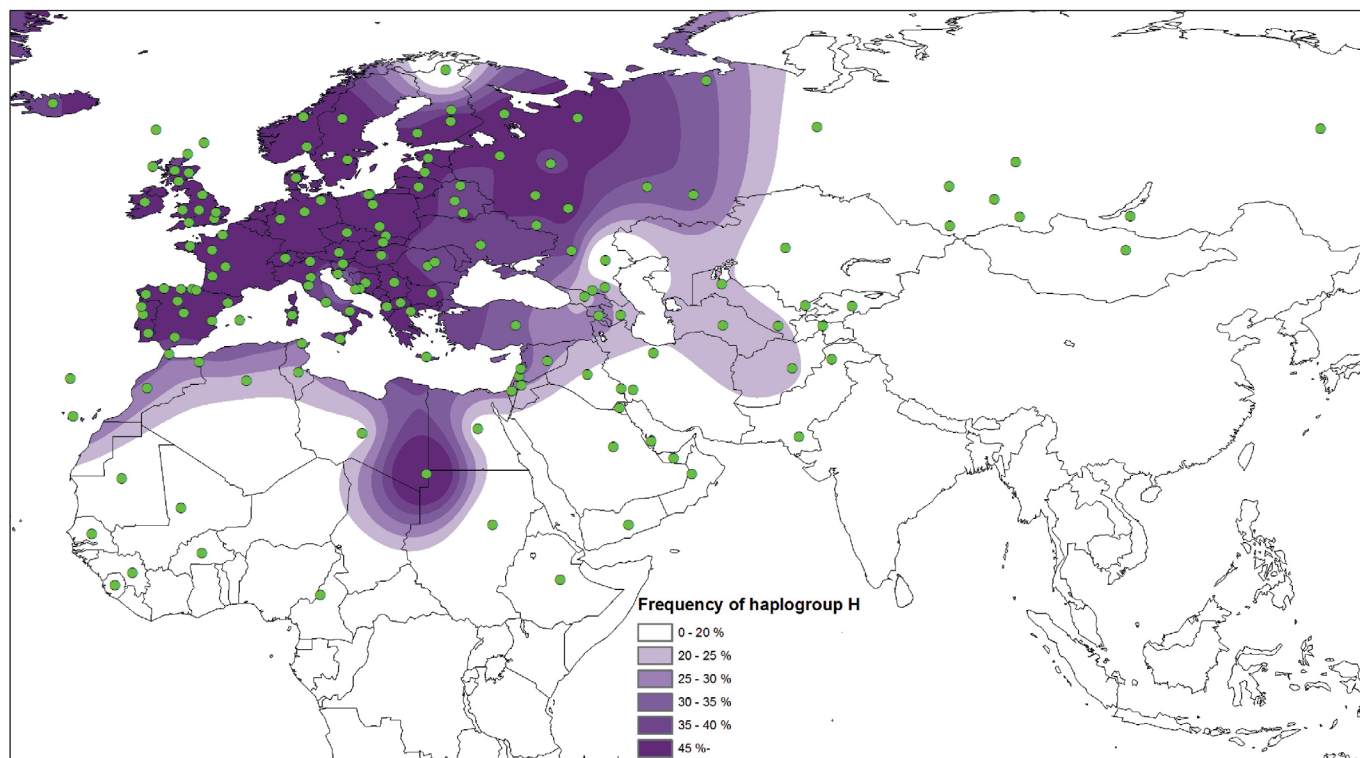


Fig. 6: Frequencies of mitochondrial haplogroup H in modern populations. Green dots indicate the location of the considered populations, the frequency of haplogroup H higher than 45% is indicated by purple colour

appears in 22% in conquering Hungarians (Fig. 6).¹⁴ Expeditions in Russia and Ukraine led by Attila Türk in the past few years collected known and currently available human skeletal samples related to ancient Hungarians. This series of bioarchaeological samples is the basis of a new research program that we have been launched in cooperation with the Doctoral School of Biology at the Eötvös Loránd University.

We extended the NGS-type investigation line to the paternal Y chromosome, X chromosome and autosomes as well. The variability types of Y chromosomes extracted from human remains of the Hungarian prehistory, compared to entire mitochondrial genomes, might reveal different migration patterns or might highlight genetic information on social practices such as patrilocality and certain forms of exogamy. The nuclear autosomes, due to their diploid nature, recombination and way of inheritance, bear the genetic heritage of several ancestors of the investigated person. Conclusions can be drawn to the complex genetic heritage of entire populations from mass SNP typing of 5-10 individuals per group. Collating our results with the increasing number of ancient genomic data, information on the genetic origin and composition of a certain population as well as the degree of its mixing with other populations is gained. Under a new Hungarian ethnogenesis project,¹⁵ we also plan full genome sequencing on the best preserved human samples, in cooperation with the MPI in Jena and the Harvard Medical School.

Archaic remains yield more and more information, and not just on the human endogenous DNA content. Several pathogens of the human body left DNA traces in these ancient specimens and can these be extracted along with the investigated individual's own DNA.¹⁶ Research in the pathogens not only focuses on their evolution and differentiation, but also on the pathogen-human and pathogen- wild and domesticated animal interactions and coevolution. While the classical methods were usually only suitable for the

¹⁴ Csősz, Aranka – Szécsényi-Nagy, Anna et al.: Maternal genetic ancestry and legacy of 10th century AD Hungarians. In preparation (2016)

¹⁵ Moments project plan "Early History of the Hungarian People – New Perspectives (Archaeology – Bioarchaeology – Written Sources)", submitted by Attila Türk, is currently under review.

¹⁶ Rasmussen, Mortan et al.: Early Divergent Strains of *Yersinia Pestis* in Eurasia 5,000 Years Ago. *Cell* 163 (2015), 1–12.

detection of pathogens,¹⁷ methods with deeper resolution and simultaneous multi-pathogen screening with DNA hybridization captures enable us to correlate the appearance of certain diseases with the genetic variability of the human population, geographical characteristics of its environment, and its movement. This is a particularly relevant question in the major target period of research in our laboratory, that is, in Neolithic and Bronze Age population history. Therefore, in cooperation with the Christian-Albrechts-University of Kiel (Institute of Clinical Molecular Biology), we started the molecular screening of Neolithic human and animal remains showing osteological symptoms of tuberculosis. The research of contagious diseases may also help archaeological research in answering questions such as the yet unknown background of Copper Age multiple burials and mass graves. Although traditionally explained by ritual reasons, these phenomena may as well refer to infectious diseases with sudden outbreaks, having significant demographic impact on the population within a short period of time. This research is planned together with the Jena MPI laboratory.

Several factors play a role in the efficient operation of the laboratory; one is the acquisition of projects which requires regular contact and project-oriented cooperation with international and domestic partners. At the same time, it is necessary to ensure the conditions of permanent laboratory work, or at least the continuity of projects. The operation of an archaeogenetic laboratory does not only depend on the mastering of developments and analyses of molecular genetic methods. The researched archaeological period, phenomenon and genetic data of a set of samples only becomes meaningful if collated with the archaeological context and data from other bioarchaeological fields. Therefore, this work is typically teamwork, where from sampling through the interpretation of bioinformatic and biostatistical data to archaeological and historical interpretation the expertise of all researchers is needed. Thus, in the case of a favorable outcome, the international and domestic success of the above outlined investigations will be the common achievement of not only the laboratory staff, but the Archaeological Institute and the representatives of cooperating scientific fields as well.

RECOMMENDED LITERATURE:

BRANDT, GUIDO – SZÉCSÉNYI-NAGY, ANNA – ROTH, CHRISTINA – ALT, KURT W. – HAAK, WOLFGANG
Palaeogenetics of Europe – the known knowns and the known unknowns. *Journal of Human Evolution* 79 (2015), 73–92.

BOLDOG, ZOLTÁN – PETKES, ZSOLT – SUDÁR, BALÁZS – TÜRK, ATTILA

[From the ancient homelands to the Carpathians](#) – From the finds to the apparel. Travelling Exhibition of Posters on the Most Recent Results from Archaeological Research on Ancient Hungarian History and the Conquest Period (The research and the article have been created within the framework of the MÓT 28.317/2012 program). *Hungarian Archaeology* (2015) Autumn.

HAAK, WOLFGANG ET AL.

Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522:7555 (2015), 207–11. doi: 10.1038/nature14317

HAGELBERG, ERIKA – HOFREITER, MICHAEL – KEYSER, CHRISTINE

Ancient DNA: the first three decades. *Phil. Trans. R. Soc. B* 370 (2015), 20130371; Published 8 December 2014. DOI: 10.1098/rstb.2013.0371.

¹⁷ Pósa, Annamária et al.: Tuberculosis in Late Neolithic-Early Copper Age human skeletal remains from Hungary. *Tuberculosis (Edinb)*. 95 Suppl 1, (2015) S18-22. Doi: 10.1016/j.tube.2015.02.011.