Molecular genetic investigation of the Neolithic population history in the western Carpathian Basin

Molekulargenetische Untersuchungen zur Bevölkerungsgeschichte des Karpatenbeckens

Dissertation zur Erlangung des Grades

"Doktor der Naturwissenschaften"

am Fachbereich Biologie

der Johannes Gutenberg-Universität in Mainz

Anna Szécsényi-Nagy

geb. am 08. April 1984 in Budapest

Dekan:

1. Gutachter:

2. Gutachter:

Tag der mündlichen Prüfung: 3. Juni 2015

Index

1 Introduction

This dissertation focuses on the Neolithic western Carpathian Basin (including western Hungary, also called as Transdanubia), and investigates its population history with molecular genetic methods. A broad scale of multidisciplinary studies is reviewed in the introduction part of the thesis, in order to enable an adequate assessment of the ancient DNA results. Such a complex and multidisciplinary introduction is required, since the field of archaeogenetics is closely related to several co-disciplines such as archaeology, anthropology, demography and several subfields of genetics (e.g. genomics, population genetic, evolutionary genetic, forensic sciences). In the introductory chapters, I survey the major and most recent scientific results of these fields, keeping the focus on the process of Neolithisation and Neolithic period of the Carpathian Basin.

1.1 The concept of archaeological culture

The prehistoric archaeological assemblages and characteristic finds (dominated by pottery types), burial rites, settlement types are ordered into archaeological cultures. The archaeological cultures are defined mainly through the material culture, which elements correlate in time and space. However, the exact meaning of a cultural affiliation, detected mostly through pottery styles is not uniformly defined in the literature (Meyer-Arendt, 2003). It cannot be answered at the actual stage of research, whether the prehistoric archaeological cultures and groups denote ethnic groups of people, identifying each other and sharing a common ancestral, social, and cultural experience. The possibility should not be ignored that as most of these cultures were parts of larger complexes, they could also have been part of larger ethnic groups, which makes the exact time-space definition of a certain culture questionable (Meier-Arendt, 2007). A prehistoric (even Neolithic) tribe could also have such common language and culture, which are nowadays captured in the form of archaeological cultures (Zalai-Gaál, 2010, p. 242-243.). Most archaeological cultures are named after either certain artifact types or their eponym sites.

1.2 Archaeological survey of the pre-Neolithic period in Europe

The European prehistory (span of time before recorded history, or invention of writing systems) is divided into four major sections or periods: the Stone Age, Copper Age, Bronze Age and Iron Age (Milisauskas et al., 2011). The Stone Age involves three larger eras: the Palaeolithic extends from the earliest known use of stone tools by hominins (~2.6 million years ago) to the end of the Pleistocene period around 11,000 years ago. The Palaeolithic term was originally defined with reference to stone industries (chipped stoneware). The name Palaeolithic implies hunting-gathering economy and a various cultural characteristics (R. Jameson in Shaw and Jameson 2008, p. 454.). The Palaeolithic period is sectioned into Lower, Middle, and Upper Palaeolithic. During the third, latest period, the anatomically modern humans settled the European continent. The most recent glacial period, containing the so called Last Glacial Maximum [LGM between 26,500-20,000/19,000 years ago (Clark et al., 2009; Zagwijn, 1992)], occurred in this period, which caused a virtual depopulation of the northern parts of Europe, and the concentration of the Palaeolithic population into some southern refuge zones (Bocquet-Appel et al., 2005).

The Palaeolithic period was followed by the Mesolithic interim era in Europe (Bailey and Spikins, 2008; Zvelebil, 2009). The material culture of this period was characterized by new forms of chipped stone tools (microliths). The Mesolithic era ended different times in different regions of Europe, and was succeeded by the Neolithic period (R. Jameson in Shaw and Jameson 2008, p. 394.). The term Neolithic refers to a period, when subsistence was characterized by dependence on domesticated plants or animals, sedentariness in permanent villages, and the appearance of pottery and polished stone tools. The Neolithic period was followed by the Chalcolithic (also called as Copper Age) or directly by the Bronze Age in Europe. The usage of Chalcolithic as a chronological term varies due to different research traditions from country to country. The Chalcolithic or Eneolithic means a period when early copper (and gold) metallurgy appeared alongside the widespread use of stone tools. In the following Bronze Age, beginning in the third millennium BC in Europe, a more advanced metallurgy appeared in Europe, wherein tin or arsenic and copper ore were smelted into cast bronze (Kristiansen and Larsson, 2005). The last period of the human prehistory was the Iron Age, started around 1,200 BC in the Mediterranean Europe, meaning the advent of the ferrous metallurgy (P. Wells in Milisauskas 2011, p. 405-461.).

1.2.1 Introduction to the archaeological records of the Mesolithic period in Europe

It is challenging to capture the meaning of the term Mesolithic or Middle Stone Age in Europe, since the Mesolithic period had different duration in diverse geographical regions (Zvelebil, 2009). The Mesolithic in Central Europe is dated between ca. 11,000-6,000/5,500 cal BC (calibrated before Christ date). Its beginning is associated with the onset of the Early Holocene climate, and its end is signalised by the agricultural transition. Generally, hunting and gathering were characteristic for the Mesolithic way of life, which means subsistence based on exploitation of aquatic and woodland resources. Its end has been defined in relation to an economic shift, the beginning of the food-production. The hunter-gatherer economic subsistence lasted longest in northern parts of Europe (e.g. Ertebølle culture), where these forager communities lived contemporaneously with southern European Neolithic farming communities (Hartz et al., 2007).

At the beginning of the Late Mesolithic, the Boreal changed to the Atlantic biozone, as the mixed oak forests expanded and became denser. The number of known archaeological sites (in comparison with the Early Mesolithic) decreased in this period (Gamble et al., 2005; Gkiasta et al., 2003). It may reflect residential changes besides population decreasing, and a concentration of people to some larger occupied camps. It is even possible that with a restructuring of gathering and exploitation of wild resources, a small scale horticulture system developed in this late phase of the Mesolithic period (Bánffy, 2006; Gronenborn, 1999).

Mesolithic sites are clustered in few regions in southeastern Europe: along the Aegean seacoast, in Thessaly, the Dinaric Alps, the Ionian hinterland, and along the Danube in the northern Balkans. A special island of increasing forager sedentism has been observed in the Danube Gorge, at the boarder of today's Serbia and Romania. According to the excavator D. Srejovic, agriculture would have developed on a local autochthonous way in the Lepenski Vir culture, where the domestication of dog, cattle, pigs wild millet, wild einkorn would have taken place on a similar way as in the Near East. The observations of D. Serjovic in Lepenski Vir site led to the conclusion, that the neighbouring hunter-gatherers would also change over to food production in the Lepenski Vir IIIa-b phases (parallel to the Proto-Starčevo horizon), and the new subsistence would quickly expand to the other river valleys, resulting the first Neolithic Starčevo-Criş-Körös culture complex in the region (Serjovic, 1988). Another scholar, B. Jovanovic has supported an alternative explanation, based on the site Padina. The Lepenski

Vir I-II culture layers would already belong to the Starčevo period (6,500–5,700 cal BC) in his view, which re-evaluate the interpretation of the higher culture layers (Jovanovic, 1969). R. Tringham suggests that farmers did in fact effected important changes on the economy of the foragers in Lepenski Vir culture. This phase was called by M. Zvelebil as a "phase of co-operation". The contacts and interchanges might have negatively influenced the forager culture's social structure, resulting the decline of their subsistence (Tringham, 2000).

1.2.2 Traces of the Mesolithic in western Hungary

The Mesolithic archaeological records have been very sporadic until the 1990s in Hungary, represented mainly by reconnaissance collections. The hypothesis of a depopulated Transdanubia during the Mesolithic era has been reframed by intensive surveys in the last decades (Bánffy, 2006; Bánffy et al., 2007; Eichmann et al., 2010). Several traces of the Mesolithic population have been identified in western Hungary, which population probably subsisted even in the early Neolithic times. Palynological evidence show a pre-Neolithic (seventh millennium BC) human activity around the western shore of the Lake Balaton, which resulted in open vegetational areas and increase of hazel pollen grains (Juhász, 2004). Around the Balaton region, a frontier zone has been supposed between the early Neolithic Starčevo culture, and the late Mesolithic autochthon people. The latter ones might have even reached a horticulture level of agriculture (Berzsényi and Dálnoki, 2006). According to settlement patterns, the shift to the food-production in middle and northern Transdanubia took place in the Middle Neolithic period, when the settlements were relocated from the lake shores to the fertile loess areas (Bánffy, 2006).

Although there are more and more signs discovered of a Mesolithic population substrate, there has not been found any skeletal remains or burials from this interim era of Transdanubia. We can calculate with the estimation of J. Petrasch, who assessed that the late Mesolithic population density in Transdanubia was 5-10 inhabitants pro 100 km^2. Considering the possible formation region of the *Linearbandkeramik* culture, he has supposed that about 1,750 hunter-gatherers lived on a territory of 30,000 km^2 (encompassing north Transdanubia, southeast Slovakia, Burgenland, and the region south of Danube in Lower Austria) (Petrasch, 2001).

1.3 The process of the Neolithisation in Europe

Studying the process of Neolithisation, the word Neolithisation or agricultural transition should be defined at first. It means the transition from foraging to farming, a process in which the change of subsistence was linked to technological novelties. In the 1950es, the original criteria to be “Neolithic” were pottery production (after V. G. Childe), and the polished macrocrystalline stone edge tools. By the 1990es, several new conditions have been added, such as sedentary settlements, architecture, social differentiation and symbolic expression (Tringham, 2000). The transition has even been interpreted as a social-symbolic process, where the agricultural domestication would have been preceded by the social domestication of the communities (Hodder, 1990, p. 20-44.). The so called “Neolithic package" is a collection of these phenomena. It does not have to be regarded as a solid construction rather as a range of sets of packages. However, the so-called staple crops, domesticates, pottery, polished and ground stone tools are present in all sets of these packages according to M. Özdogan (Özdoğan, 2008).

1.3.1 The demic diffusion and wave-of-advance model

The godfather of the "Neolithic Revolution" and that of the migrationist theory (and diffusionist model) was V. G. Childe, who described for the first time the Neolithisation of Europe as a population replacement by the immigrant farmers. He has set up the diffusionist model, based on the assumption that all cultural innovations originated from a source region, where the civilization was the earliest (Orient), and diffused from there to Europe (Childe, 1957). Furthermore, he has also brought the concept of demographic pressure into the theories of the Neolithisation (Childe, 1942).

The gradual spread of farmers (which is the meaning of the term demic diffusion) from the Near East into Europe was firstly demonstrated by radiocarbon dates of Neolithic settlements, which dates have a changing distribution along the southeas-northwest axis of Europe (Clark, 1965). Several archaeologist scholars have supported the demic diffusion (or by some scholars also called as colonisation model) of the agricultural transition in Europe (e.g. Bogucki and Grygiel 1993; Renfrew 1987). Furthermore, some anthropologists have also taken a stand on the demic diffusion model (e.g. Lalueza-Fox, 1996, Pinhasi and von Cramon-Taubadel, 2009).

A. Ammerman and L. Cavalli-Sforza have performed a regression analysis with radiocarbon dates and supposed centres of the Neolithisation, defining the diffusion rate of the Neolithic dispersal about 1 km/year with a relatively high correlation coefficient (R=0.8) (Ammerman and Cavalli-Sforza, 1971). The rate of diffusion has been calculated several times since the first study of Ammerman and Cavalli-Sforza. For example the observed rate (0.6–1.3 km/year) calculated by R. Pinhasi has been regarded as consistent with the demic-diffusion model, hence cultural diffusion probably should go faster (Pinhasi et al., 2005). Nevertheless, this rate of dispersal was considered as too low for the *Linearbandkeramik* expansion (dissemination of the first Neolithic culture in Central Europe) by several scholars (Whittle 1996, p. 151.; Bogucki, 2000).

The diffusion theory has been further shaped, when A. Ammerman and L. Cavalli-Sforza introduced the so-called wave-of-advance model, which has combined two features: active population growth at the periphery of the farmers' territory, and an isotropic local migratory diffusion or range expansion. The result would be a population spread in all directions at a steady rate. The admixture between farmers and foragers would lead to genetic gradients, having a maximum on the oldest Neolithic sites (Ammerman and Cavalli-Sforza, 1973). The prediction was underpinned by classic genetic markers, evaluated with principal component analysis (PCA) (Menozzi et al., 1978). Using the principal components (PC), seven synthetic maps were constructed. The first component (map), showing a gradual distribution of populations from the Middle East to North/West Europe, has been explained by the wave-of-advance model (Ammerman and Cavalli-Sforza, 1984). The model of Ammermann and Cavalli-Sforza has been supported by further genetic analysis. Using spatial autocorrelation method, a similar southeast-northwest cline has been obtained for a quarter of the studied genetic markers (Sokal et al., 1991). Among archaeologists, C. Renfrew involved the wave-of-advance hypothesis in his theory about the spread of the Indo-European languages along the expansion of agriculture (Renfrew, 1987).

However the wave-of-advance model has also been criticized, e.g. by M. Zvelebil, who pointed out that it is not obviously the Neolithic migration, which is seen on the first PC, but other dispersal waves could have also caused such cline (Zvelebil, 1998). Furthermore, M. Richards et al. have stressed that the routes of the Neolithisation might have been used several times, for example during the Early Upper Palaeolithic, hence the "one PC-one migration event" theory could not be upheld (Richards et al., 1997).

The above-mentioned two expressions, migration (folk migration), and demic diffusion have been clearly distinguished by M. Zvelebil. He has described folk migration as a directional movement to a previously defined region, whereas the demic diffusion is rather a slow non-directional and sequential colonization in his view (Zvelebil, 2001). Most of the archaeologists agree nowadays that the process of Neolithisation cannot be regarded as one large-scale migration through the continent, but rather a heterogeneous event, altering on smaller scales (Gronenborn, 2007; Zvelebil, 2000).

1.3.2 The indigenous model and the integrationist approaches

The indigenous or autochthonous model says that the agricultural transition took place on a given region independently. The appearance of domestic plants or animals would rather be signs of exchange or natural spread, than proofs of migration. The key examples of this model are the cultural diffusion and frontier contacts (Zvelebil, 2000). From the two main forms of the indigenist transition, some have considered the cultural diffusion as the key spreading characteristic of farming technology (Barker, 1985; Dennel, 1992). Other scholars have pointed out the ideological and social restructuring of the communities as components of the shift to farming (Thomas, 1996; Whittle, 1996). Focusing on specific regions of Europe, the role of the hunter-gatherers has been emphasized for example in the western Balkans by M. Budja and in southern Central Europe by A. Tillmann and N. Kalicz (Budja, 2005, 1999; Kalicz, 2010; Tillmann, 1993). The indigenous model has special significance on the outlying regions of the LBK territory, where the hunter-gatherers lived parallel with the southern farming communities (Whittle and Cummings, 2007).

The integrationist position, focusing on smaller scale movements and considering the social aspect of the Neolithisation phenomenon, regards the following processes: leapfrog colonization, frontier mobility and contact (Thorpe, 1996; Zvelebil, 2000, 1989). Leapfrog colonization means a selective colonization by small groups, which form colonies surrounded by native forager inhabitants. M. Zvelebil considered it as a possible scenario for the Neolithic transition in southeast and Central Europe, where the farming groups would have targeted the fertile loess regions, building enclaves in hunter-gatherer milieu (Zvelebil, 2000). Frontier mobility could have taken place in contact zones between foragers and farmer communities, where social networks modulate its way and intensity. The last expression, contact means a

channel of communication and innovation along trading networks according to M. Zvelebil (Zvelebil and Lillie, 2000; Zvelebil, 2000).

None of the above presented mechanisms should be considered as the general explanation for the Neolithic transition, all of them could have had a role in the agricultural transition, in various range and combination from region to region.

After surveying the most important theories and terms of the European Neolithic transition, the Neolithic archaeological records ranging from the core area to Central Europe are summarised in the following chapters.

1.4 Archaeological records of the Neolithic period of Europe

1.4.1 The Primary Zones of Neolithisation: Levant and Anatolia

The Neolithic way of life begun to develop on the area of the "Fertile Crescent" around the tenth millennium BC. The original definition of the term "Fertile Crescent" included the geographical extent stretching from the Levant to the Mesopotamian river basin (Breasted, 1916, 100-101.). This core area of the Neolithisation in the Near East has been extended by M. Özdoğan, to cover all of Levant, northern Syria, Iraq, south-eastern Turkey, central Anatolia and Cyprus (Özdoğan, 2008). Congruently, R. Pinhasi and his colleagues, provided quantitative support (based on radiocarbon data) to the origin of agriculture in the area north-east Syria, northern Mesopotamia, and part of south-east Turkey (Pinhasi et al., 2005).

The first sedentary culture with harvesting farming cereals were probably the people of the Natufian culture in the Late Epipalaeolithic (~10,000-9,300 cal BC) period (Bar-Yosef, 1998). Although in other opinions, this transition happened only in the Pre-Pottery Neolithic A (PPNA, dated from ~9,600 cal BC). The first domestic plants are unequivocally attested in the early Pre-Pottery Neolithic B (PPNB, from ~8,500 cal BC), with the first domestic animals documented in southeastern Anatolia (Özdoğan, 2008; Vigne, 2008). At the time of the transition from Pre-Pottery to Pottery Neolithic period in the core area, the Neolithic appeared in the neighbouring regions ad well. This transition in the primary zone went along with depopulation and major social changes. The new settlements of the Pottery Neolithic period throughout southeastern Anatolia, Iraq, Syria, and in the Levant were "simple" villages without all complexity and monumentality of the preceding PPN culture of the same regions (Özdoğan, 2011, 2008).

The movement of the Neolithic communities from the core area took place through two distinct routes according to C. Perlès: one through the Anatolian plateau and another, reaching the Aegean bypassing the Anatolian plateau on a maritime way (Perlès, 2005). The Western parts of Turkey on the other hand, deducting from the distribution of the stamp seals, could have played an important mediating role, not only at the beginning of the Neolithic in Greece around 6,500/6,400 cal BC, but also at the beginning of the Neolithic in Southeast Europe at around 6,000 cal BC (Lichter, 2005).

M. Özdoğan has separated two migration waves or movements of the Neolithic communities from western part of Turkey to the Balkans. The first was signalised by monochrome pottery, around the mid-seventh millennium BC, spreading to eastern Thrace and perhaps further into southeastern Europe (Özdoğan, 2011). During the second wave of spread at the end of the seventh millennium BC, several new settlements appeared in northwestern Turkey and Balkan Peninsula, without detectable predecessors. This more rapid and massive movement can be well characterised by red-slipped and burnished pottery, tubular lugs, plastic decoration in relief, anthropomorphic or zoomorphic vessels, steatopygic figurines and pintaderas (Özdoğan, 2011).

N. Özdoğan defines a so-called interim zone, which covers all western parts of Anatolia, the Aegean, Marmara region and most of the Balkans. In this interim zone, the substratum of the Final Palaeolithic/ Mesolithic periods is either missing or insignificant (except the coastal areas and the northern part of the Balkans). The Neolithic package arrived into this zone fully developed, with symbolic and prestige objects, indicating that the specialised craftsmen had also moved in from the core area (Özdoğan, 2008). Northwest Turkey, in contrast to the interim zone, seems to have been only on the periphery of the main dissemination route of the Neolithic to Europe, where the Sea of Marmara was rather a barrier than a bridge in the dispersal of Neolithisation (Lichter, 2005).

1.4.2 Early Neolithic in southeastern Europe

Neolithisation of Greece

The earliest European traces of Neolithic subsistence have been found in Thessaly (Greece) from the mid seventh millennium BC. This transitory phenomenon of Pre-Pottery (or aceramic) Neolithic horizon, characteristic for domesticated animals and plants with lithic assemblage of geometric microliths, is only conceivable in Crete and Greece in Europe. The radiocarbon dates for the Greek Pre-Pottery Neolithic sites (Argissa, Knossos, Franchthi) are all at around 6,800 BC (Tringham, 2000). Nevertheless, some scholars have doubted the existence of a real aceramic horizon, and interpret this period, as ceramic were too rare and prestigious to discard (e.g. Bailey, 2002, p. 77.; Budja in Bogucki and Crabtree, 2004, p. 235.).

The radiocarbon dates of the subsequent Ceramic Early Neolithic in the south Balkan Peninsula fall between 6,600–5,800 cal BC (Tringham, 2000).

The precise point of origin of the first farmers in Greece, and the motivation behind their migration remain among the most controversial chapter of the Aegean prehistory. The intensively debated (e.g. by Zvelebil and Lillie, 2000) demic diffusion model for southeastern Europe has been introduced by T. van Andel and C. Runnels. They have hypothesised that Anatolian farmers had settled first in Thessaly, from where, after reaching a demographic saturation, they moved further towards the Danube and the Carpathian Basin (van Andel and Runnels, 1995). On the other hand, M. Budja and other scholars have suggested that the local foragers in Greece had the potential for an indigenous adoption of agriculture (Budja, 1999; Tringham, 2000; Zvelebil and Lillie, 2000). C. Perlès has argued for a Mesolithic-Neolithic discontinuity in Greece, because the Early Neolithic lithic technology differs completely from the Mesolithic ones (e.g. Francthi cave comparing with Argissa or Slesko) (Perlès, 2001). However, in K. Kotsakis' opinion, different habitational environments could cause the observed lithic technological differences. Although, there is a Mesolithic "gap" in Greece, the missing archaeological sites can also be a result of research history (alluvial deposits may cover unobserved Mesolithic sites, or postglacial rise in the sea level) in his and other scholar's opinion (Chapman, 1994; Kotsakis, 2001).

The Neolithisation of Greece has been interpreted by C. Perlès as a maritime colonisation, performed by small groups, arriving mostly from Levant. This conclusion was drawn through the observed stylistic and technical parallels between Levant and Greece. In contrast, Balkans and northwest Anatolia show another cultural circle. Common features between Greece and Anatolia are mainly general background phenomena, probably originated from Levant according to C. Perlès (Perlès, 2001). Congruently, S. College and his colleges have found archaeobotanical similarities between southern Levantine, Cypriot, and Aegean sites, standing in contrast to Anatolia and Euphrates Valley/central Steppe region of Syria. They have also suggested two ways (maritime to Greece and land route through the Anatolian plateau to Thrace) for the Neolithic dispersal (College et al., 2004). Nevertheless, the issue of the origin of the first farmers is still controversial. C. Runnels e.g. has explained the Early Neolithic in Greece, as *"a peripheral extension of the Anatolian cultural core"* (C. Runnels in Bogucki and Crabtree, 2004, p. 223.)

Early Neolithic in Bulgaria

In northern and western Bulgaria, the horizon of the earliest pottery inventory was dominated by monochrome ware, expressed with absolute dates between 6,400 and 6,200 cal BC from the site Polyanitsa-platoto. White- and then red-painted ceramics defined the Bulgarian early Neolithic A (6,200–5,700 cal C), which were especially typical in western Bulgaria (Bailey 2002, p. 89.).

Different names such as Karanovo I-II, Kremikovci or Gradisnitsa have been attributed to the earliest Neolithic phase in central Bulgaria. To date, the surveyed Early Neolithic settlements have not shown detectable Mesolithic predecessors. The emergence of the earliest farmers in the region indicates a clear northern movement, which probably took place from western Anatolia at around ~6,100/6,000 cal BC, about half a millennium later than the "colonisation" event in today's Greece. M. Özdogan stressed that certain cultural elements were common features, observable in the Neolithic of western Anatolia, Marmara region and in Bulgaria, connecting these regions together (Özdoğan, 2008). On the other hand, different models have been presented by L. Nikolova for the origin of the Karanovo I culture. It can either be a succession of a monochrome phase Neolithic, which has not been detected yet in Bulgarian Thrace, or developed after the colonisation of the region (Nikolova, 1998). Anyhow, due to the absolute date evidence, the Bulgarian Early Neolithic could not be the source of the Early Neolithic horizon in the Carpathian Basin.

Early Neolithic of the north Balkans

Similarly to the southern regions, the way of Neolithisation of the north Balkans has been debated intensively too. Some scholars consider the local adoption of farming in almost the entire Balkan Peninsula (Budja in Bogucki and Crabtree, 2004; Chapman, 1993). Others emphasize the rapid way of Neolithisation of the region (Biagi et al., 2005). Drawing its time limits in absolute terms, the Early Neolithic in the north Balkan Peninsula is counted between 6,500–5,200 cal BC (Tringham, 2000).

According to the archaeological records, the northwards wave of migrations along the Vardar, Morava and Strimon rivers are marked by the spread of white and red painted pottery,

attributed to the Starčevo culture (Kalicz et al., 2012). However, the Neolithic package did not appear in the Balkans everywhere on the same manner. Whilst there were areas such as the Vardar–Morava corridor, the Maritsa basin, and the middle and lower Danube basins, bearing coeval introduction of Early Neolithic material culture and domesticates, in other areas these features were not introduced parallel. The Balkans rather shows a mosaic of different Neolithisation patterns (Zvelebil and Lillie, 2000). For example, in the Adriatic and the Dinaric Alps the Neolithisation was a slow process according to, M. Budja. The interregional networks of communication avoided the Adriatic and the Dinarides, leaving more space for hunting and gathering. The Neolithisation depended here more on social networks than on population transformation in the view of M. Budja (Budja, 2005). The Starčevo type painted ceramic disseminated in the central Balkans whereas the first farmers in the coastal areas used the Impresso type ceramic.

Starčevo culture is named after the eponym site Starčevo-"Grad", located eight kilometres southeast from Belgrade in Serbia. The culture was disseminated in the territory of the present-day Macedonia, Kosovo, Serbia, Bosnia, north of Croatia (Syrmium) and the south part of western Hungary. Its adjacent cultures were the Körös culture on the Great Hungarian Plain (Alföld) and the Criş culture in Romania (Anders and Siklósi, 2012; Biagi et al., 2005; Kalicz, 2010). Radiocarbon dates have shown that the Starčevo-Criş-Körös complex appeared around the turn of the sixth millennium BC, and lasted for almost 800 years. The oldest ^{14}C dates (~6,200 cal BC) are known from settlement Blagotin-Poljna in the Šumadija region of central Serbia (Whittle et al., 2002). The radiocarbon dates indicate a very rapid expansion of the Starčevo-Criş complex from its source in the Balkans (Biagi et al., 2005). This culture was contemporaneous with the western Bulgarian Karanovo-I complex and the Greek Proto-Slesko and Slesko horizons. They connecting element was the painted pottery, which made up however just a small part of the ceramic spectrum.

The Starčevo type settlements have typically thin occupation layers. Emmer and einkorn wheat, six-row barley, and peas have been found at Starčevo settlements. Domesticated sheep and goats prevailed in stockbreeding, but cattle and pigs did not play a significant role in the subsistence patterns of the Starčevo culture in the north Balkans (M. Budja in Bogucki and Crabtree, 2004, p. 139).

1.4.3 Neolithic in the western Carpathian Basin

The Early Neolithic began in Transdanubia at around 5,800 cal BC, when the first farmers of the Starčevo culture settled on the region south of the Lake Balaton (earliest known Starčevo radiocarbon dates are in this study from Alsónyék-Bátaszék Mérnöki telep, see Supplementary Table 1). The Starčevo culture was succeeded by the *Linearbandkeramik* or *Linienbandkeramik* culture or in English Linear Pottery culture (hereafter referred as LBK) in the Middle Neolithic (Hungarian terminology), which transmitted the knowledge of farming and sedentary subsistence to the major part of Central Europe. Parallel to the LBK, a second wave reached southern Transdanubia from the Balkans, introducing the Vinča culture to the region of western Hungary. Vinča and the following Slavonian Sopot cultures were closely related in Transdanubia, and the formation of this latter culture signalized the beginning of the Late Neolithic (local terminology) in Hungary. Sopot formed the basis of the large (both in space and time) Lengyel culture complex, which stretched in time as far as the Middle Chalcolithic. Note that in the neighbouring countries this phase is still called as Eneolithic or Late Neolithic. Since these five cultures, cultural periods were the focuses of my study, they are briefly introduced in the following chapters. Their supposed dissemination territories are represented in Figure 1.

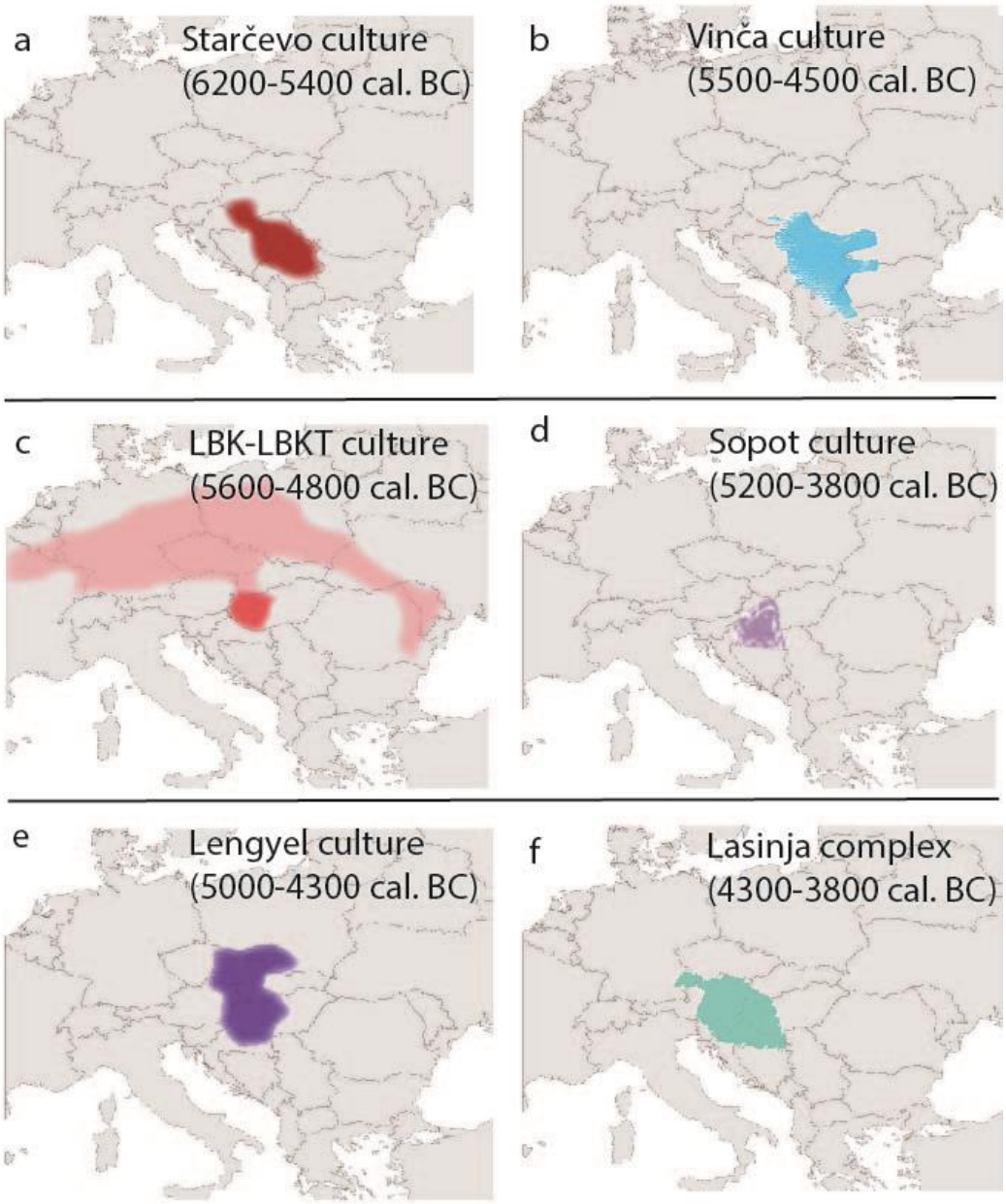


Figure 1. Dissemination of the studied six prehistoric cultures in Europe.

Fig. 1a: Starčevo culture, 1b: Vinča culture, 1c: LBK in Central Europe and in Transdanubia, 1d: Sopot culture, 1e: Lengyel culture, 1f: Lasinja complex. Note, that the approximate BC dates indicate the existence of the culture considering the whole dissemination area. References for the dates are also mentioned in the text (Borić, 2009; Brandt et al., 2013; Minichreiter and Bronic, 2006; Oross and Bánffy, 2009; Stadler et al., 2006; Whittle et al., 2002). The maps were drawn using several sources for each map: (Bánffy, 2002, 1996b; Horváth and H. Simon, 2003; Kalicz, 2010; Oross, 2013; Regenye, 2011, 1996a; Zvelebil, 2000). Dates that are more precise to the specific study region of Transdanubia are given in Figure 6.

1.4.3.1 The Starčevo culture

Starčevo was present in Hungary from the Linear B phase of the culture [relative chronology follows the study of S. Dimitrijević (Dimitrijević, 1974)]. The Starčevo culture populated both valleys and the hills of southern Transdanubia, a more varied environment than the Alföld region, settled by people of the Körös culture.

The described dissemination territory of the Starčevo culture in Hungary was enlarged little by little during the last decades. It advanced as far as the Lake of Balaton in the 1990es with the discovery of sites Gellénháza-Városrét and Vörs-Máriaasszony-sziget (H. Simon, 1994; Kalicz et al., 1998). Until 2007, 26 Starčevo sites were known from Transdanubia (Kalicz et al., 2007). Its northernmost appearance is at Tihany-Apáti, situated north of the Lake Balaton (Regenye, 2010). In the last decade, new large-scale excavations in southeast Transdanubia gave a fresh insight into the settlement history of the culture. One of those took place in Alsónyék-Bátaszék, where an enormous quantity of pottery and Starčevo features came to light. It became the largest Starčevo settlement in Hungary, and opened a new era of the Early Neolithic archaeological studies in Hungary (Bánffy et al., 2010).

Typical Starčevo type ceramic is the painted, often polychrome fine pottery, decorated with linear lines or spirals. These two types of patterns also characterise the relative chronological phases of the culture. According to the most recent records, the Starčevo culture in Transdanubia seems to be a single cultural unit with the eastern Slavonian distribution of the culture, where it was settled from the Linear A to the Spiraloid B phase. The end of the Starčevo culture in Transdanubia has been estimated to 5,450/5,350 cal BC (Lenneis and Stadler, 1995; Oross and Bánffy, 2009).

1.4.3.2 Vinča culture

Vinča culture (VIN) developed from the Starčevo culture according to J. Chapman's model, in a gradual shift, which involved a transitional phase of the late Starčevo and early Vinča culture in the latter's core region (Chapman, 1981). The Lepenski Vir culture might have contributed to its formation according to R. Tringham. As an argument for this hypothesis, she has noted that the Vinča culture shows subsistence based partially on local sources, breeding with local animals, and exploitation of local wild plants (Tringham, 2000). The transition of Starčevo to Vinča in former Yugoslavia was dated to 5,500/5,400 cal BC by E. Hertelendi and

his colleagues (Hertelendi et al., 1995), and the Vinča culture lasted ~4500 cal BC on its core region (Gläser, 1996). The absolute chronological dates still increase in number, but the exact dating of the Vinča A-B-C-D phases remains unclear.

During the long lasting Vinča period, the north Balkans witnessed an increasing cultural complexity and population growth. The Vinča culture disseminated to the large part of the mid-Balkans (Gronenborn, 1999) (Figure 1b), and it also reached the southern region of today's Hungary. It culturally affected and determined several archaeological formations in the Carpathian Basin for a millennium time span (Bánffy, 1996a; Horváth, 2006).

The presence of the Vinča culture in western Hungary has been supposed by J. Makkay at first, apropos of the pottery materials found in Fajsz-Garadomb and Bicske-Galagonyás (Makkay, 1982). N. Kalicz has also pointed out the Vinča pottery style traits (such as solid pedestals, burnished pattern designs and the row of impressions under the rim) found in LBK materials in Transdanubia (Kalicz, 1994a).

The excavations of the last decades, especially along the motorway M6 in southeast part of Transdanubia have broadened our knowledge about the impact of the Vinča material culture on the LBK formation. The sites Tolna-Mözs in Transdanubia and Fajsz-Garadomb, Baja-Bajaszentiván-Szlatina on the Alföld site of the Danube show typical Vinča A pottery characteristic, which was the basis of a presumption of the Vinča culture's presence in south Transdanubia or in the Drava - Sava interfluve (Marton and Oross, 2010).

The recently excavated first two Vinča sites from south Transdanubia are being processed by J. Jakucs. The material on the site Szederkény shows typical early Vinča (Vinča A1-2) and early Sopot (Sopot Ib) pottery characteristics in connection with LBK longhouses (Jakucs and Voicsek, 2014). The radiocarbon measurements in our project date the graves (and possibly the sites) to ca. 5,400-4,900 cal BC. Further ongoing radiocarbon dating will clarify the chronology of these Vinča settlements in southwest Hungary.

1.4.3.3 Linearbandkeramik culture in Transdanubia

The LBK was one of the major Neolithic cultures of Europe, spanning across a vast geographic region from the Paris Basin to Ukraine (see Figure 1c), determining the material culture of Central Europe for five centuries. It originated from the study region of western Carpathian Basin, where this study abbreviates the LBK culture as LBKT (T stands for

Transdanubia). The expression LBKT should not mean any cultural difference or division of the LBKT from the Central European distribution of the LBK, rather a geographic detachment, important to characterize for the comparative ancient DNA analyses.

A formative phase of the LBK (and LBKT) was defined in Transdanubia, based on the observations on the excavation of E. Bánffy in Szentgyörgyvölgy-Pityerdomb with an absolute dates of 5,480-5,340 cal BC (Bánffy, 2004). The two Central European LBK-similar houses were associated with Starčevo pottery and lithic material of Mesolithic characteristics. The initial territory of the LBK (and LBKT) would be the Balaton region and a small region west of the lake. This formative LBKT phase was likely contemporaneous with the latest Starčevo and forager sites from south and north of the Lake Balaton respectively, and could have begun at around 5,600/5,500 cal BC. The sedentary lifestyle still evolved at this time, and reached its mature complexity in the late LBKT period (Oross and Bánffy, 2009). M. Zvelebil has supposed as well that the people of the Central European LBK originated from Hungary and not from southeast Europe. An integration of the local foragers took place in the place of origin, resulting a shift is settlement pattern and adopting the agriculture from the early farmers in his view (Zvelebil, 2000).

The most typical LBKT vessels are the conical and biconical bowls with small rims. Decorative motifs are broad incised lines, which often form bands. The earlier LBKT phase is divided into two ceramic typological periods. The first one got its name after the site Bicske-Galagonyás and Becsehely-II-Homokos with the best parallels of Bíňa in south Slovakia. The next ceramic typological phase is the Milanovce-phase, named after an early LBK site in Slovakia. The Bíňa-Bicske and Milanovce phase can be dated to around 5,450-5,250 cal BC (Marton and Oross, 2010; Oross and Bánffy, 2009; Oross and Marton, 2012; Stadler and Kotova, 2010).

A recently identified strong Vinča A influence in southeast Transdanubia distinguishes this region from the northern part, which is tightly connected to the territory of south Slovakia. Southern Transdanubia became in the Late LBKT phase a periphery within the LBK distribution (Marton and Oross, 2010). The culture was divided in two groups at this time: the Notenkopf ornamented type followed by the Zseliz/Želiezovce type in north part of western Hungary and south Slovakia, and the Keszthely group south of the Lake Balaton (Kalicz, 1990). The late LBKT (especially the northern Zseliz part) lasted until 5,000/4,900 cal BC in Transdanubia, when it transformed into the Sopot and Lengyel cultures (Oross and Bánffy, 2009).

1.4.3.4 Sopot culture

The Croatian (or Slavonian) Sopot culture (referred as SOP) was named after the eponymous site Sopot near Vinkoci. It had several development phases in east Slavonia (Sopot I-IV), starting contemporaneous with the B phase of the Vinča culture. Characteristic feature of this culture is the dark monochrome ceramic and carving and tally ornament. Other forms of the Sopot culture are cannelling, pressing, and ribbon, decoration by incision and pricking. The emergence of the Sopot culture in Slavonia has been explained with Vinča influence on Starčevo substrate (Obelić et al., 2004).

The first description of the Sopot culture in the territory of Hungary was done by N. Kalicz and J. Makkay, describing the Sopot site Bicske-Galagonyás (Kalicz and Makkay, 1972). The material characteristics and chronological phases of the Transdanubian Sopot culture have been further differentiated after the discovery of the site Becsehely (Kalicz, 1988). In the traditional research view, the Sopot reached Transdanubia in its second phase (so called Sopot II-Bicske culture), which was parallel to the Vinča B/C transition period, the Zselíz III phase of the LBK (Horváth and H. Simon, 2003). On the Alföld region, the classic Bükk, Esztár and Szakálhát groups of the Alföld LBK lived at the time of the Sopot II phase. After several hints of information, indicating that the Sopot type material appeared as early as the Sopot Ib phase in Hungary (Horváth, 2006), J. Jakucs, T. Marton have managed to confirm this theory (Jakucs and Voicsek, 2014).

The early Sopot pottery type (Sopot IA) in middle Slavonia is called as Ražište type, which shows characteristic Vinča and LBK cultural influences. In north-western Croatia, the Brezovljan type developed in the Sopot IB, disseminated from there to the western part of Transdanubia (Kalicz and Makkay, 1972; Regenye, 1996a). Whereas the Brezovljan type shows more LBK similarity, the classic Slavonian Sopot type contains more Vinča elements. This latter type is also found in east Transdanubia (e.g. Bicske) (Regenye, 2002, 1996b). The Sopot culture was successively replaced by the Lengyel culture in Transdanubia, but in Slavonia its III phase further flourished until ca. 3,800 cal BC (Obelić et al., 2004).

The recently excavated, still unpublished Alsónyék-elkerülő Sopot site is the largest Sopot graveyard in Transdanubia to date (A. Osztás, T. Marton pers. comm.). It will certainly reform our knowledge about the Sopot II culture in Transdanubia.

1.4.3.5 Lengyel culture

The Lengyel culture (referred as LGY), named after the eponym site Lengyel-Sánc in south Transdanubia, emerged from an interim proto-Lengyel phase at around 5,000/4,900 cal BC. Its formation area was probably the Aszód-Svodin region in middle-north Hungary (Kalicz, 1994b). The younger phase of the Sopot culture persisted in the neighbourhood of the formative phase of the Lengyel culture, before the latter one replaced the late Sopot culture in Transdanubia. The early period of the Lengyel culture (Lengyel I phase) was characteristic for the polychromatic painted pottery. The yellow, red, and brown coloured ceramic material was produced with inclusion of sand and metal. The first phase of the Lengyel culture in Transdanubia was parallel to the C and the early D phase of the Vinča culture in the central Balkans, and to the Sopot III in Slavonia. On the Alföld region the Tisza II culture, the Herpály I-II and the early Csőszhalom were parallel with the Lengyel I period. This contemporaneity is observable on the imported Tisza objects, which appeared on the Lengyel territory (Horváth and H. Simon, 2003), and on the Lengyel imports reaching the tell settlements of the Tisza region (Horváth, 2005).

The Lengyel II phase was parallel to the Tisza III phase and the Herpály III, Csőszhalom groups on the Alföld. During the Lengyel II period, the culture disseminated to the eastern part of Hungary, it reached the Tokaj Mountains in northeast of the Alföld. The numerous Lengyel settlements, concentrating especially in southern Transdanubia, indicate an increased population size in comparison with the previous periods. The first copper objects appeared in this phase of the Late Neolithic, which were mainly jewels, found as grave objects (e.g. in graveyards of Zengővárkony and Mórágy (Dombay, 1960; Zalai-Gaál, 1988). At Alsónyék-Bátaszék, the largest Neolithic graveyard known today was uncovered between 2006 and 2009. With 2,359 graves, it opens new dimensions and interpreting possibilities in the research of the Neolithic Carpathian Basin, rewriting our knowledge about Late Neolithic societies (Osztás et al., 2012).

J. Regenye pointed out that east-west differences existed in the material culture (idol plastics and pottery types) and burial customs of the Lengyel culture. The eastern Lengyel group (located in eastern Transdanubia and middle Hungary) shows intensive influences of the Tisza culture, whereas the western group (in western boundary of Transdanubia) was on

the route between the Balkans and Central Europe, having different cultural influences from the eastern variant (Bánffy, 2002; Regenye, 2011).

Lengyel culture is often named as “painted pottery culture” in the Neolithic research of the adjacent countries. It built a large culture complex (Figure 1e), stretching out from Transdanubia to Croatia and Slovenia, west Slovakia, south Moravia, Austria and even to Lesser Poland (Bánffy, 2002; Regenye, 2011; Stadler et al., 2006). E. Ruttkay has designated the Lengyel groups of Moravia and eastern Austria as the “Moravian-Eastern-Austrian group of painted pottery”, identifying an internal cultural homogeneity but also a clear differentiation from the Transdanubian Lengyel culture.

At the end of the Lengyel II phase, the Late Neolithic came to its end in Transdanubia. Nevertheless, the division of the Neolithic and Chalcolithic epochs of the region is an artificial approach. The two were connected by continuous development of the discussed culture (Regenye 2011, 53).

1.4.4 The Early and Middle Chalcolithic in Transdanubia

The Lengyel III period represents the first Chalcolithic culture in Transdanubia, which was parallel to the Tiszapolgár culture on the Alföld region. This latest phase of the Lengyel culture was differentiated by P. Raczky in 1974, based on the material of the Veszprém site. The ceramic of this late Lengyel period was typically unstained, rather the plastic ornaments (knots) were characteristic for this phase instead (Raczky, 1974). The Lengyel III period was followed by the Balaton-Lasinja culture in Transdanubia.

The Lasinja culture got its name after the eponym site Lasinja, found south of the city of Zagreb (Dimitrijević, 1979). Several hypotheses have been described about the emergence of the Lasinja culture. Some scholars have emphasized the continuity and peaceful development of the culture (Minichreiter and Markovic, 2011), whereas other have pointed out the role of migration in its development (Horváth and H. Simon, 2003). Since the Lasinja culture (also called as Lasinja complex or circle) was distributed on a large area (see Figure 1f), probably different formation scenarios could have taken place on various cultural substrates.

Climatic changes could also have had a role in its emergence. A drier, warmer climate began in Eastern Europe at the time of the late Lengyel/Proto-Tiszapolgár period, which might have resulted in a migration wave of people, from the north Pontic region toward

southeastern Europe. This wave could have reached the northwestern Balkans, where they probably influenced the emergence of the Balaton Lasinja culture (Horváth and H. Simon 2003, p. 119.). The southern elements in ceramic typology of the Balaton-Lasinja have been revealed in an early stage of the research (Kalicz, 1973, 1969), and it has also been verified in the western and northern parts of the cultural complex (Bánffy, 2002). N. Kalicz found Balkan effects on some pottery forms and on the metallurgy, and therefore he has connected the emergence of the culture to infiltration of new groups from the south, possibly in relation with the termination of the Vinča culture (Kalicz, 1983, 1973, 1969). However, this influences might have reached Transdanubia one period earlier, in the late phase of the Lengyel III, as it was shown on ceramic sequences (Bánffy, 2002, 1994). On the other hand, BL also had common cultural elements with the Lengyel culture (Kalicz, 1995), and these characteristics have often been interpreted as evidence for an indigenous development of the Balaton-Lasinja.

The Lasinja complex in Slovenia and Croatia (Jaksić-Čaire group) and the Balaton-Lasinja culture in Transdanubia (referred as BL) were culturally connected to each other, and they were parallel to the Bisamberg/Oberpullendorf in Lower-Austria and its successor, the Baalberg culture. In Moravia, the Jordansmühl (Jordanów) group belonged to this circle, and the Lasinja horizon reached even the southeastern part of Bavaria (Münchshöfener group). In the Alföld region, the Bodrogkeresztúr culture, in Slovakia the late Ludanice and early Furchenstich ceramic culture were contemporaneous with the Lasinja complex (Horváth and H. Simon 2003, p. 107.).

The inhumation graves were very sporadic in the whole Lasinja circle. This lack of skeletal remains stands in contrast with the richness in graves of the contemporaneous Bodrogkeresztúr culture on the Alföld. Cremation also occurred (Horváth and H. Simon, 2003; Kalicz, 1973, 1969), which had a cultural tradition from the Lengyel culture onward (Zalai-Gaál, 1988, p. 70-71.).

The Lasinja circle has a new characteristic metallurgy, meaning the peak of the Central European Neolithic/Chalcolithic metallurgy. It is represented by the horizon of metal hoards, containing the Stollhof-Csáford type gold disks. The long distance routes across southeast and Central Europe and intensive contacts between the Alföld region, Transdanubia, and the Vinča culture were consequences of the new inventions. The Lasinja complex had a mediator role between the Balkans and Central Europe. The trade channel along the Drava river, then north

to the Wien basin and further to west, following the Danube reached even Bavaria (Bánffy, 2002).

The Balaton-Lasinja period (~4,300-3,900 cal BC in Transdanubia) was eventful in Central-Europe as well. The Funnel Beaker or *Trichterbecher* culture emerged in north Europe, and it dispersed toward the south, building the Baalberge culture/group in Central Germany, Moravia and Lower-Austria (Ruttkay, 1995). From the southwestern direction, the Vasi a bocca quadrata culture from Italy reached Slovenia, and it could also have a contact with the Lasinja circle (Bagolini and Barfield, 1991).

The BL was followed by the *Furchenstich* culture in Hungary, at around 3,800 cal BC. In the meanwhile, the Lasinja tradition maintained further on other parts of the complex [e.g. until 3,300 cal BC in Slavonia (Minichreiter and Markovic, 2011)].

1.4.5 Neolithic in Central Europe: spread of the LBK and its successor cultures

The LBK spread from western Hungary on a remarkably rapid way to Lower Austria, west Slovakia, and further into Czech Republic, Germany, up to the Rhine valley. J Petrasch estimated that the dispersal of this Earliest LBK took 50-100 years, based on the homogeneity and the closeness of the published radiocarbon dates from the early phase the culture. This vast region encompassed the dissemination territory of the "Earliest/*Älteste* LBK", which showed homogenous ceramic characteristics on this region of ~135,000 km^2 (Petrasch, 2001). The most accepted hypothesis is nowadays that the LBK spread by colonization in Central Europe, although some scholars advocate indigenism (e.g. Whittle, 1996; Price in Price, 2000, p. 1-19.). Arguments for the LBK colonization are the suddenly appeared new house types, from the Mesolithic different stone tool compositions (mainly blade tools) and the new ceramic technology (Bogucki and Grygiel, 1993). P. Rowely-Conwy summarizing the current knowledge about the LBK migration in Central Europe has noted that the boats could be important factors, accelerating the LBK migration along the rivers. He has initialised a new expression “lurches of advance” instead of “wave of advance” of Cavalli-Sforza (Ammerman and Cavalli-Sforza, 1984), hence the spread was punctuated in his view. The people of the LBK culture followed loess and fertile soils, and they built “settlement cells” on the prolific lands (Rowley-Conwy, 2011). D. Gronenborn has argued as well for a rapid migration that involved

a “*multi-faceted combination of migrations, adaptations and acculturations*” (Gronenborn, 2007, p. 73).

Earliest LBK began around 5,600/5,500 cal BC in Transdanubia (Oross and Bánffy, 2009), or even earlier, if we consider P. Stadler's radiocarbon results (5,700/5,650 cal BC). According to P. Stadler and his ^{14}C dates from site Brunn-II, the Earliest LBK reached or developed in Austria probably around 5,650 cal BC (Stadler and Kotova, 2010). However, the so-called old wood effect could cause some misdating in his analyses (interpreted by Gronenborn, 1999, who referred to personal communication with P. Stadler).

LBK reached Bavaria around 5,500 cal BC. The expansion over a distance of 650 km took place during the first 200 years of the culture. From Bavaria, a second advance moved toward the Rhine, probably around 5,400 cal BC. Finally, a third advance started with the onset of the early Flomborn phase between 5,300 and 5,250 cal BC, settling the Rhineland and Alsace (Gronenborn, 1999). This Flomborn phase is also called as the "*ältere* LBK". The pottery of this period is not only found in the original distribution area, but also beyond it, in Netherlands, southern Poland (upper courses of the Vistula and Oder) and western part of Ukraine. The settlement structure changed in the Flomborn phase, special major settlements emerged, and the settlement density increased in such a trend that a population growth of minimum 1.4% pro year has been estimated by J. Petrasch (Petrasch, 2001).

The next phase was characterized by the "music note" or *Notenkopf* style in the eastern part of the LBK distribution, while in the western part incised decoration appeared without punctates. The people of the LBK settled in this period in the lower Vistula and Oder Rivers, and in Uckermark in northeast Germany.

In the Late LBK period, the expansion continued toward the Paris Basin and Belgium. The pottery decoration became further regionalised, establishing styles as Želiezovce/Zseliz in Moravia, southern Poland, Transdanubia, Šarka in Silesia and Bohemia (Bogucki and Grygiel, 1993).

During the last phase of the LBK, a trend in development of regional styles and practices appeared, and several “daughter cultures” emerged. The LBK was succeeded by Rössen and Stroke-Ornamented Pottery (Stichbandkeramik) culture in the central and western part of its distribution, and by the Lengyel culture in its eastern province.

1.4.6 The maritime route: spread of the agriculture in the Mediterranean

A second movement beside that of the Starčevo culture started at around 6,000 cal BC along the northeastern Mediterranean coastline. Two major cultures emerged, the Impressed ware culture in the eastern Mediterranean (eastern Adriatic and Ionian coastal area) and the Cardial culture in the west Mediterranean. They both bore the inventory of the "Neolithic package", originated from the Near East (Zilhão, 2001). The earliest impressed type pottery appeared on Sidari site on Corfu (Layer C) at about 6,200 cal BC (Forenbaher and Miracle, 2006). The dissemination of Impressed or Impresso wares in the eastern Adriatic began with a rapid 'leap frog' colonisation of Italian Peninsula and southern Dalmatia, followed by a slower expansion [with the full suite of domestic animals and new lithic technology (Perlès 2001, 49-50.)] northwards along the coastal lines of Dalmatia. Colonisation happened clearly through seafaring (Forenbaher and Miracle, 2006). On the other hand, contacts and frontier zones were also established between the coastal regions and the hinterland areas, which latter were populated by indigenous foragers even at this time. However the sets of dated archaeological sites around the Mesolithic-Neolithic transition are biased by see level rise during the early- to mid-Holocene, which remind us to be very cautious with the interpretation of radiocarbon gaps between Mesolithic and Neolithic settlement traces in the Mediterranean (Mlekuz et al., 2008).

1.5 Genetic systems with uniparental inheritance in the research of the human population history

1.5.1 The mitochondrial DNA

1.5.1.1 Characteristics of the mitochondrial DNA

The mitochondrion cell organelle has an energetic function in the eukaryotic cells, as it produces most of the cell's supply of adenosine triphosphate. Additionally, it participates in calcium signalling, intermediary metabolism, and apoptosis. Mitochondrion has an origin of an endosymbiotic bacteria in the eukaryotic cells (Thrash et al., 2011). This putative source is reflected in its genome: each mitochondrion may contain several identical mitochondrial genomes, which are circulars and composed of double-stranded DNA molecule. The 16 kilobases (16,569 bp [base pair]) long mitochondrial genome (mtDNA) contains 37 genes and a so called control region or D-Loop region (stretching between nucleotide positions 16,024-576) (Anderson et al., 1981; Chan, 2006). Certain bases within the D-loop region are conserved, but the larger part is highly variable. They are divided into three segments in the literature, the Hyper Variable Region or Segment I (HVS-I) II (HVS-II) and III. These polymorphic segments have an estimated mutation rate of $1.64*10^{-7}$ for the HVS-I and $2.297*10^{-7}$ for the HVS-II, which are about 10 times higher than that of the coding region's (Soares et al., 2009). The reason for this difference is that the control region is non-coding and thus accumulates more mutations than the coding part of the genome. The exact ranges of the sequenced HVS-I and HVS-II segments vary form study to study, but most of the scholars consider the following division as to be canonical: HVS I (nucleotide position [np] 16,024-16,365); HVS II (np 73-340) and HVS III (np 438-574) (Malyarchuk et al., 2002). The D-Loop or control region does include some of the fastest sites in the molecule (e.g., 16,311, 16,189, 16,129, 16,093, and 16,362 in HVS-I and 152, 146, 195, and 150 in HVS-II), but such mutation hot-spots occur in the coding region as well.

The mitochondrial genome is exclusively maternally inherited in humans (Giles et al., 1980), which enables genealogical researchers to trace maternal lineages far back in time. There is usually no change in mtDNA from parent to offspring, since mtDNA recombines only with copies of itself. Phylogenetic and population genetic analyses are therefore free of those complexities that are imposed by biparental recombination of the autosomes.

On the other hand, mtDNA has a much higher mutation rate than nuclear DNA. The whole molecule has a high mutation rate as well [latest estimate is $2.67*10^{-8}$ /site/generation (Fu et al., 2013)]. The molecular clock hypothesis, which postulates that DNA sequence evolution is roughly constant over time in all evolutionary lineages, makes the calculation of divergence times and certain nodes of the phylogenetic tree possible. The human mtDNA sequences, excluding the D-loop, evolves at roughly constant rates (Ingman et al., 2000). The estimated most recent common ancestor (MRCA), the so called "mitochondrial Eve" lived around 157,000 years ago (Fu et al., 2013) or 124,000 BP (BP stands for years Before Present, given in an interval of 99,000–148,000 BP) (Poznik et al., 2013).

The polymorphic nature of the mtDNA makes it useful in assessing genetic relationships of individuals or groups. At the advent of the mtDNA research, HVS-I and HVS-II were the focuses of the population genetic investigations (Richards et al., 1996), besides detection of selected coding region positions by RFLP (Restriction Fragment Length Polymorphism) techniques (Torroni et al., 1993). As the sequencing methods advance, more and more studies focus on the whole mitochondrial genome, gaining fine resolution of the phylogenetic structure of the mtDNA tree, and more precise mutation rate of the mitochondrial genome (e.g. Ingman et al., 2000, Fu et al., 2013).

The human mitochondrial genome has a low effective population size (because of its uniparental inheritance), which leads to increased genetic drift. This phenomenon causes geographic structures of its variability, such as continent-specific mtDNA lineages. The related (usually monophyletic) lineages have been assembled to haplogroups, which contain lineages or haplotypes sharing common ancestors.

1.5.1.2 Haplogroups, lineages and phylogeny of the mitochondrial DNA

The mtDNA haplogroups are signed by letters of the Latin alphabet. The first mtDNA haplogroups were described with the letters A, B, C in Native Americans (Torroni et al., 1993). By now, all letters of the Latin alphabet have been used, except the O (van Oven and Kayser, 2009). Haplogroups are assigned for related lineages that bear similar single nucleotide polymorphisms (SNP) structure or share "haplogroup defining" alleles. By convention, the nucleotide positions of the human mtDNA genome are numbered from 1 to 16,569 according to the revised Cambridge Reference Sequence (rCRS) (Anderson et al., 1981; Andrews et al.,

1999). Detected SNPs and insertions/deletion polymorphisms are signed as differences to the rCRS. A recent approach has rooted the nomenclature of the mtDNA tree, and suggested the switch of the reference genome to a so called “Reconstructed Sapiens Reference Sequence” (RSRS) (Behar et al., 2012b). This ancestral sequence has been defined by considering all available mitochondrial genomes from Homo sapiens (n=18,000) and all Homo neanderthalensis genomes as out-groups on the mtDNA tree. The ancestral mtDNA sequence of extant humans should correspond to the bifurcation of L0 (sub-Saharan) haplogroup and L1’2’3’4’5’6 cluster. That means a recent African origin for all mtDNAs, which is a subsequently supported, but rather old theory (Cann et al., 1987).

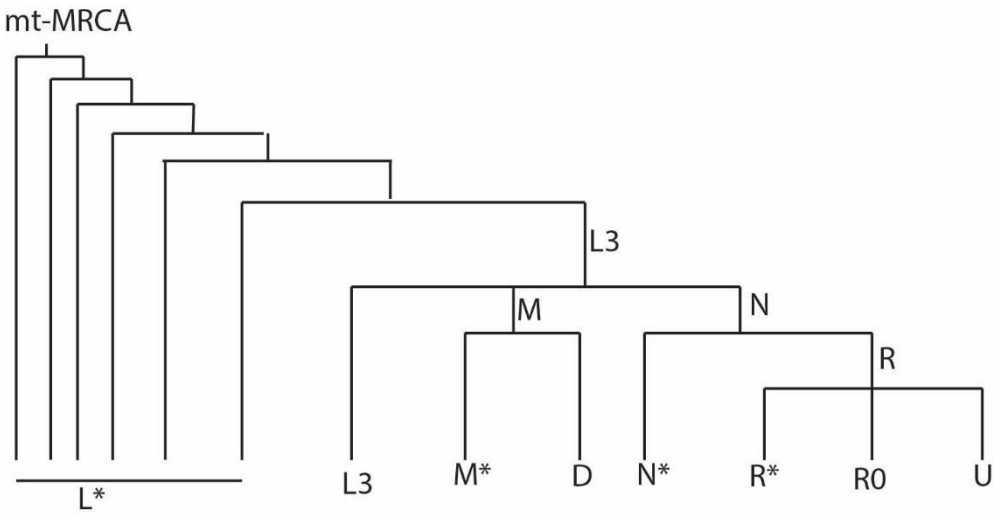


Figure 2. Simplified mtDNA tree.

Modified after van Oven and Kayser, 2009.

Haplogroup L3, originated in East Africa, gave rise to every non-African haplogroups, after a small number of L3 migrants left Africa at around 60-70,000 years ago (Behar et al., 2008; Soares et al., 2012). The time of the MRCA (T_{MRCA}) of the group L3 (corresponding to the time of the population divergence) has been dated to 78,300 BP (Fu et al., 2013). The next divergence occurred soon afterwards, at around 77,000 BP, when the M+N clades separated from L3. M clade (including haplogroups C, E, G, Q, Z, and D) was disseminated toward Asia and to Indonesia, Australia, and America. Its sisterclade N includes almost all European and Oceanic haplogroups in addition to many Asians and Amerindians. Parahaplogroup N* contains haplogroups A, I, O, S, W, X, and Y. The most common European clade R is also within the N cluster. It is divided to R* (including B, F, J, P, and T), R0 (including HV, H, and V) and U, which latter includes haplogroup K (Figure 2).

The most frequent recent European haplogroup is H (contributing to 40-45% of the European lineages), which have the highest frequency in Western Europe. Further major European haplogroups are J, K, and T (each has a frequency up to ~10%) (M Richards et al., 2000, and a private database built by G. Brandt and our team (see for references in Brandt et al., 2013, Szécsényi-Nagy et al., 2014a).

1.5.2 The characteristics and phylogeny of the Y-chromosome

The human Y-chromosome inherits exclusively paternally. It has two major parts: the pseudo-autosomal regions (5% of the chromosome at the ends of the chromosome), and the non-recombining part (abbreviated canonically as NRY), containing 95% of the Y-chromosome. This latter part is gene-poor, and does not have a partner to recombine with (Tilford et al., 2001). An evolutionary trend is observable that the genes of the Y-chromosome, originated from a common ancestor with the X, are slowly vanishing. Non-allelic homologous recombinations (recombination between similar sequences that are not allelic) occur often on the Y-chromosome, which make it instable and more mutagenic than the female germ-line. Y-chromosome has higher mutation rate than other parts of the nuclear genome, and have small effective population size due to its haploidity. This latter can cause more genetic drift, which led to a low genetic diversity within species (Homo sapiens) and a more recent MRCA then that of the mtDNA. Furthermore, the genetic drift also causes geographic structures of the Y-chromosomal diversity.

Y-chromosome carries several hundred microsatellites and SNPs. The Y-chromosomal SNP profiles have been combined into haplotypes, traced phylogenetically, and assigned to Y haplogroups. Similarly to the mtDNA tree, a Y-chromosomal phylotree exists (see a simplified version in Figure 3). After the pioneer European phylogeographic studies, done by Semino et al. and Rosser et al. (Rosser et al., 2000; Semino et al., 2000), a unified nomenclature system for the human Y-chromosomal tree was proposed (Y Chromosome Consortium, 2002). This system is followed ever since, and has been revised by Karafet et al. and is being updated by the International Society of Genetic Genealogy (ISOGG, 2014; Karafet et al., 2008).

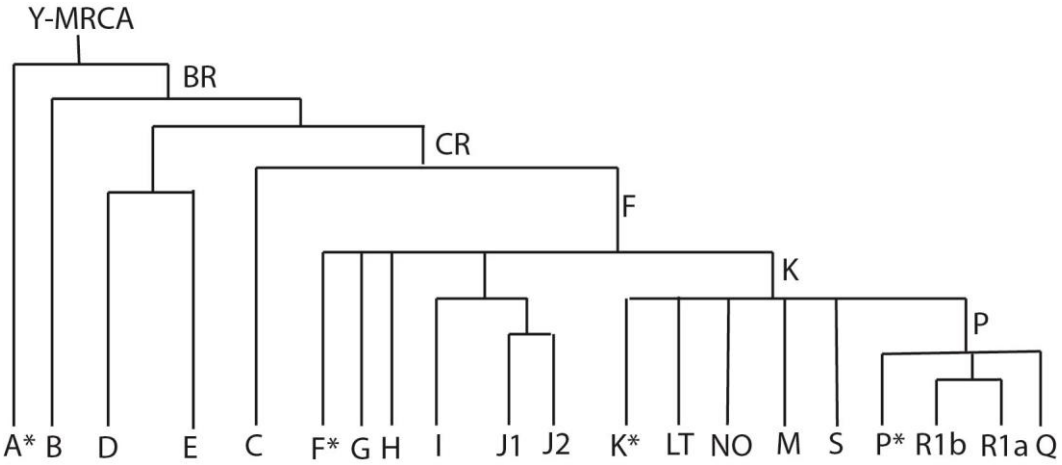


Figure 3. Simple Y-chromosomal haplogroup tree.

It is based on the Y tree of ISOGG (www.isogg.org). Paraphyletic groups (i.e., paragroups) are differentiated from monophyletic groups (i.e., haplogroups) with an asterisk.

The Y-chromosomal "Adam" lived in Africa, and belonged to the para-haplogroup A0*. The T_{MRCA} reported for the entire Y-chromosome phylogeny has been varied on a wide range: from the very old date (188,000 BP, based on a 2.6 kb DNA fragment) (Hammer, 1995) to the very young (~50,000 BP, relying on three resequenced Y genes) (Thomson et al., 2000), with some intermediate dates (e.g. 90,000 BP by Hammer and Zegura, 2002). Recently, fully sequenced diverse Y-chromosomes have become available and the discrepancy that the T_{MRCA} of the mtDNA would be older than that of the Y-chromosome seems to be resolved. After several contradictory estimations, the newest sequencing results assume a T_{MRCA} for the entire Y phylogeny being 138,000 years old (120,000 – 156,000 BP) (Poznik et al., 2013).

Regarding the Y-chromosomal haplogroup diversity, the A and B clades are virtually restricted to Sub-Saharan Africa. The very diverse haplogroup E occurs in Africa, West Asia and Europe as well, whereas the sister clade D occurs only in Asia. In the "Out of Africa" event, the haplogroups E, CR, and F took part. The CR has a T_{MRCA} of 70,000 years BP (Karafet et al., 2008) and involves C and F branches. Most of the European haplogroups are part of the F cluster, but some haplogroups of F disseminated to Asia, East Asia and to America as well. K clade has a T_{MRCA} of 47,000 BP and R has 26,800 BP (Karafet et al., 2008). The R clade contains two branches, the R1a and R1b that are particularly interesting for the European population history. R1a is likely originated in Central Asia (Iran) and disseminated to Eastern and North Europe, where it reaches 60% frequency (Semino et al., 2000; Underhill et al., 2014). Contrarily, the R1b has its highest frequency in Western Europe (57% on an average) (Semino

et al., 2000). Its best-studied subclade R1b1a2 (M269) is most commonly found among modern European populations. For example, it reaches a frequency of 95% in Great Britain (Balaresque et al., 2010). The second most important clade in Europe is the haplogroup I, the only European-specific branch. It has two major clades, showing opposing dissemination patterns. Whereas the I2a is frequent in southeastern Europe, the I1 is typical for North Europe (Rootsi et al., 2004).

The haploid Y-chromosome has become an extremely important tool in a variety of areas of particular interest to anthropologists, including DNA forensics (Ballantyne et al., 2014), genealogical reconstruction (Larmuseau et al., 2013), molecular archaeology (Lacan et al., 2011a, 2011b) and human evolutionary studies (Poznik et al., 2013; Thomson et al., 2000).

1.6 "Archaeogenetic" evidence of the modern molecular genetics

Molecular geneticists have been attempting to write human population history from mutation/polymorphism data for decades. Constructing and dating phylogenetic trees obtain a basis to draw conclusions about history. Besides phylogenetics, the present-day phylogeography (study of the geographical distribution of genetic variation) supplies evidence for the modern approach of archaeogenetics.

As Soares et al. in 2010 summarized, four major migration events were the focus points of the molecular geneticists, studying the European late Pleistocene-early Holocene epochs. The first was the pioneer colonisation of Europe by the Homo sapiens in the Upper Palaeolithic. The second was the Late Glacial re-colonisation of the Continent from southern refugia after the Last Glacial Maximum (Clark et al., 2009). The third was the postglacial re-colonization of deserted areas after the Younger Dryas [between 12,800 and 11,500 BP (Muscheler et al., 2008)] cold snap, and the last was the arrival of the Near Eastern farmers.

Most of the mitochondrial haplogroups have been interpreted as Palaeolithic haplogroups in Europe. M. Richards et al. have defined the major founder clusters in Europe. In this approach, founder haplotypes have been used as baseline, from which founder events associated with haplotype clusters could be identified and dated. They have found the U cluster as to be the oldest in Europe, dating from the Early Upper Palaeolithic. The basal U was followed by the HV, I, and U4 clusters. In the Late Upper Palaeolithic appeared the haplogroups H, T and K in Europe (Richards et al., 2000).

Among the Y-chromosomal haplogroups, the lineages approximately equivalent to R1a and R1b have been proposed to be signals of the Palaeolithic substrate and expansions from refugia of the Iberian peninsula and the present Ukraine, following the Last Glacial Maximum (20,000 to 13,000 BP) (Semino et al., 2000). Pericic et al. have stressed the expansion of NRY haplogroup R1a from east to west during the post-Last Glacial Maximum as well (Pericić et al., 2005). R1b (M269) has been the focus of a persistently controversial debate. Short tandem repeat (STR) analyses and dating of the lineage suggested the Neolithic origin of this haplogroup in Western Europe (Balaresque et al., 2010). The microsatellite-based dating was later criticised by Busby et al. (Busby et al., 2012).

According to Semino et al., NRY haplogroup I (M170) would have originated in Europe after the arrival of the Gravettian people form the Middle East between 20,000 and 25,000 years ago. I2a would have originated in the western Palaeolithic population during the LGM (Semino et al., 2000). Pericic et al. have argued for a I2a1 (P37.2) diffusion out of southeastern Europe in the Younger Dryas-Holocene (Pericić et al., 2005).

The dispersion of several haplogroups have been connected to the recolonization of Europe from the southwestern refugia during the late glacial period (~15,000 BP), such as the mtDNA sub-haplogroups H1, H3, U5b1b, haplogroup V, and some haplogroups of Y-chromosomal I clade (Achilli et al., 2004; Rootsi et al., 2004; Torroni et al., 2001).

Complete mitochondrial genome sequencing has obtained finer haplogroup resolution. For example HV3, HV4, and U4a1 have been identified as pre-Neolithic (~12,000–19,000 BP) mitochondrial haplogroups in eastern Europe by Malyarchuk et al., whereas U4a2a, U4a2*, HV3a, and R1a1 are younger, being dated between 6,400 and 8,200 BP (Malyarchuk et al., 2008).

The molecular genetic approaches, based on modern mtDNA data have yielded contradictory results about the Neolithic impact on the modern European gene pool. The estimated contribution of the Near Eastern farmers in the today's European maternal gene pool has assumed to be about 20% (Richards et al., 2000, 1997, 1996), which was interpreted as a predominantly Palaeolithic origin of the modern European population, with only minor detectable Neolithic component. According to M. Richards, only the mitochondrial haplogroups J and T1 have a Neolithic origin in Europe (Richards et al., 2000). The method was criticised by Barbujani and Chikhi as the age of a population is not the age of the common

molecular ancestor of its set of DNA sequences. They argue for a major Neolithic contribution of the European gene pool, based on the continent-wide gradients of allele frequencies (Barbujani et al., 1998).

Some early Y-chromosomal studies have rejected the demic diffusion, emphasizing the legacy of extant European lineages from the Upper Palaeolithic, with an estimation of the Near Eastern Y-chromosome contribution to 22% in the modern European gene pool (Semino et al., 2000). However, genealogical likelihood-based re-evaluation of the Semino et al. data has yielded opposing result, suggesting a significantly larger (~65%) genetic contribution of the Neolithic farmers than previously supposed (Chikhi et al., 2002). Nevertheless, a migration cannot be dated by genetic data alone; therefore, this result does not mean a direct early Neolithic contribution. Furthermore, several Y-chromosomal studies have supported the Neolithic diffusion model through observations of haplogroup frequencies, clining along a southeast to northwest axis of Europe. Among other haplogroups, G (M201), J2a (M410) (Battaglia et al., 2009; Sengupta et al., 2006), J2b (M12) (King et al., 2008), E (M35) (Semino et al., 2000) have been suggested as markers of the early farmers' colonisation.

Further population genetic studies, using not only uniparental, but also autosomal markers, estimate a higher Neolithic genetic contribution rate, which supports the demic diffusion model of the Neolithisation (Chikhi et al., 1998; Rosser et al., 2000).

The high ratios of parallel substitutions found in the mtDNA genome make the mtDNA T_{MRCA} estimations unreliable. The field of modern molecular genetics is moving on, thanks to technological advances and the increasing availability of whole genome data. Next-generation sequencing enables by now to make direct measurements of the mutation rate in modern humans (Scally and Durbin, 2012). However, even if the molecular clock becomes estimated more precisely, a certain migration still cannot be dated by the genetic data alone. Modern molecular genetic studies cannot consider lineages or genomic variation, which did go extinct in the past, or estimate precisely past populations' diversities. Population processes, such as founder events (reduced genetic diversity in a population founded by a small number of people) and bottlenecks (reduction of genetic diversity due to a reduction in population size) keep hindering the look into past population genetic events (Jobling et al., 2013).

1.7 Studying ancient DNA

1.7.1 Introduction

The retrieval of ancient DNA (aDNA) from a specimen is a challenging task, because the target DNA might be present only in small quantity. The invention of polymerase chain reaction (PCR) technique [by K. Mullis, (Mullis and Faloona, 1987)] has made it possible to amplify and analyse DNA of low copy number, and allowed the field of archaeogenetics a rapid development. Ancient (mitochondrial) DNA was amplified and sequenced for the first time from a museum specimen of a quagga (Higuchi et al., 1984), a subspecies of the plains zebra that became extinct in the 19th century. The PCR technique was than subsequently used for the analysis of the DNA remains of several extinct animals, such as marsupian wolf, moa, mammoth, mastodon, cave bear (reviews in Pääbo et al., 2004; Hofreiter et al. 2001b).

The history of archaeogenetics have several examples for exceptional results, which were either not reproducible or turned out to be biased by contamination. For example, a putative dinosaur DNA (cytochrome b sequences) (Woodward et al., 1994) turned out to be integrations of mitochondrial DNA into the human nuclear genome, and therefore being a human contamination (Zischler et al., 1995).

The approximate theoretical limit (Lindahl, 1993) of the survival of ancient DNA in cold conditions has been approached recently. The technical improvements in the investigation of ancient DNA coupled with the techniques of the “Next Generation or Whole Genome Sequencing” have recently allowed to sequence the draft genome from a ca. 700,000 years old horse (Orlando et al., 2013). The earliest sequenced hominin mitochondrial genome has also became older by almost an order of magnitude, when M. Meyer and his colleagues have published the sequence of a ca. 400,000 years old individual found in Sima de los Huesos in Spain (Meyer et al., 2014).

Studying ancient DNA is used not only for retrieving the DNA of extinct species (hominins), but also for reconstructing diet and behaviour. This can be studied e.g. using coprolite (e.g. Poinar et al., 2003), intestinal remains or dental plaques (Adler et al., 2013) of past people and animals. Furthermore, aDNA has applicability for medical archaeology [tracing pathogens, e.g. Mycobacterium leprae (Schuenemann et al., 2013)], for the history of domestication [e.g. (Bollongino et al., 2006)] and for population history and phylogeography.

The technique of aDNA enables a direct insight into the past. Investigating the population history and phylogeography using aDNA is much more precise than extrapolating past history from modern-day DNA results. For example, genetic varieties can be observed, which might vanished or diluted up to our time (such as mtDNA haplogroup N1a, (Haak et al., 2005). Furthermore, conclusions about the genetic variability of past populations from modern DNA data can only be drawn limitedly, because different population historic events are not clearly distinguishable, if their source region is the same. Studying ancient DNA has the benefit that past migrations and responses to natural selection can be directly tracked back (Pickrell and Reich, 2014).

1.7.2 Taphonomy of ancient DNA

The DNA of deceased organisms goes through a decaying process. Studying the course of fossilization, the transition of remains from the biosphere to the lithosphere is called as "taphonomy" by the palaeontologists (Shaw, in Shaw and Jameson, 2008, 563). During the decomposing process, the DNA is exposed to different types of damages due to environmental effects. The degradation of the DNA by microorganisms, enzymatic cleavages, and non-enzymatic hydrolytic effects result in strand breaks of the DNA molecule. It leads to size reduction of the DNA molecules. Oxidative lesions of bases and deoxyribose residues are caused by free radicals. These damages along with DNA crosslinks (intermolecular and interstrand) can block the DNA polymerase, making the template DNA un-amplifiable. (Dabney et al., 2013; Hebsgaard et al., 2005; Pääbo et al., 2004). Further types of damages are the hydrolytic lesions, which lead to deamination of the bases (loss of amino groups): from adenine becomes hypoxanthine, from cytosine uracil, from 5-methyl-cytosine thymine and from guanine becomes xanthine (Gilbert et al., 2003). These modified bases can cause incorrect nucleotide incorporation during the PCR. Consequently, it leads to false DNA sequences (with C ->T and G ->A substitutions). At such cases, the original ancient DNA sequence can either be retrieved through cloning process, or pre-treated with uracil *N*-glycosylase before amplification. Further quality control possibility is the quantification of the template DNA (Hofreiter et al., 2001a). However, phantom mutations (Brandstätter et al.,

2005), damage and jumping PCR (Hofreiter et al., 2001a) might cause difficulties in defining the endogenous sequences.

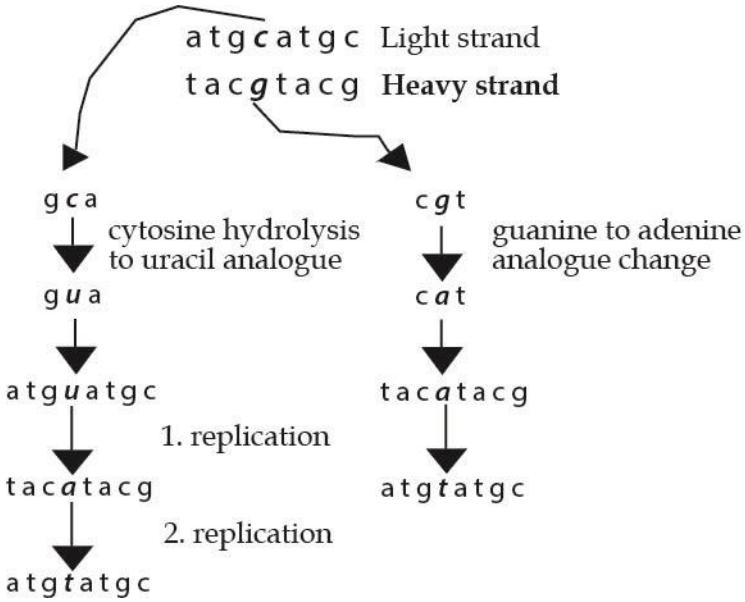


Figure 4. The process of nucleotide misincorporation at the amplification of degraded ancient DNA template.

Figure is based on Figure 1. in Gilbert et al., 2003. It demonstrates how postmortem deamination becomes apparent in PCR amplicons. This train of thought was followed several times during the evaluation of the clone sequences.

The extent of the degradation may vary from sample to sample even if they originated from the same archaeological site or have a same age. Heat can degrade DNA rapidly (Lindahl, 1993), the long-time survival of DNA in cold conditions, such as in cave or especially in permafrost is more likely [see e.g. the mammoth DNA studies, such as (Gilbert et al., 2007; Krause et al., 2006)]. Besides cold temperature, rapid desiccation and high salt concentration can help preserving DNA molecules. These conditions make the endogenous nucleases inactive or destroyed, thus they cannot degrade the DNA. Taphonomy supplies important background information for ancient DNA analyses. Knowing the degradation process of the DNA helps choosing the adequate sample and proper evaluation of the data.

1.7.3 Problems and challenges of ancient DNA work

The degraded nature of ancient DNA causes a major issue in archaeogenetic research. The above listed postmortem damages allow amplify ancient DNA only in short fragments (up to 500 bp on an average). The challenge of fragment length has been overcome by the new sequencing techniques. Whereas at the PCR technique a minimum of 60-70 bp DNA fragment is necessary, with next generation techniques the limit drops to a length of 20-30 bp, allowing older and poorly preserved material to yield useful sequence data.

The second major challenge of archaeogenetics is the danger of contamination. Archaeogeneticist researchers happen to interpret a false positive result as endogenous DNA (see chapter 1.7.1 for example studies). Such contaminations can originate from the post-excavation history of the samples (DNA of the anthropologists, excavators etc.), or it can come from intra-laboratory contamination: DNA of the processing staff, cross-contamination by another sample, carry-over contamination from the post-PCR part of the laboratory, and chemical contamination.

Since the first attempts on the field of archaeogenetics (e.g. Pääbo, 1989), methodological advances obtained safer work frames, reducing the possible sources of contamination. A series of studies have addressed the issue of contamination, establishing a set of criteria that are pivotal for an authentic ancient DNA result (Gilbert et al., 2005; Hebsgaard et al., 2005; Hofreiter et al., 2001b; Pääbo et al., 2004; Willerslev and Cooper, 2005). In 2005, the following eight points of criteria have been declared by a group of aDNA experts: (1) isolation of work areas, (2) negative controls in extractions and PCR, (3) appropriate molecular behaviour of the ancient DNA, (4) reproducibility of the results by multiple extractions and PCRs, (5) cloning of the products, (6) replication of the results by independent research groups, (7) quantification of the starting template DNA, and (8) testimony of the associated remains (Gilbert et al., 2005). These criteria have been listed in several other combinations. Furthermore, it has been improved by the “exclusion of nuclear insertions of mtDNA” and the “biochemical assays of macromolecular preservation” (Pääbo et al., 2004).

This system of criteria has been robustly changed with the introduction of the next generation sequencing techniques, which has brought new solutions but also new problems into the field of archaeogenetic. Authenticity of sequences is assessed for example by comparing the mtDNA sequence reads to the reference (mtDNA), and by considering the

plausibility of the consensus sequence in phylogenetic and geographic context. Furthermore, a generally accepted criterion is the presence of C-to-T damage patterns at the 5'-ends of DNA fragments (used for e.g. in Krause et al., 2010; Lazaridis et al., 2014).

Reviewing the advances of whole-genome sequencing, J. Pickrell and D. Reich have listed some new challenges of aDNA work. At first, sequencing has to be preceded by screening many carefully chosen and prepared samples, until a well performing subset is identified. Those samples have to be selected that have a sufficiently high proportion of endogenous DNA. Further challenges are the computational and analytical issues of the large amount of generated sequence data (Pickrell and Reich, 2014).

The last issue to mention is the correct dating of the study sample. Analysis of a misdated skeleton can lead to totally false conclusions (Bánffy et al., 2012). Systematic radiocarbon dating of the studied samples and close cooperation with the excavator archaeologists can help overcome such pitfalls.

1.8 Ancient DNA studies of prehistoric population events, focusing on the Neolithisation of Europe

Ancient DNA analyses of the Central European Neolithic have made a significant progress in the last few years. The first milestone in the research history was the study of W. Haak and his colleges of 24 LBK individuals from Germany, Austria, and Hungary (Haak et al., 2005). Mitochondrial haplogroup N1a has been found to be the most characteristic for the LBK maternal gene pool, with a frequency of 25% in the Neolithic and only about 0.2% in the today's Europe. The N1a has been supposed to signalize the immigrant farmers and its dilution would indicate further post-LBK population genetic events. The LBK sample set has reached an extent of 108 individuals to the year 2013 (Bramanti, 2008; Brandt et al., 2013; Haak et al., 2010, 2005). The frequency of N1a haplogroup has decreased (12%), but it has remained one of the major mtDNA haplogroups of the early farmers. When compared to modern-day population data, the maternal ancestry of LBK in Central Europe shows a Near-Eastern affiliation (Haak et al., 2010, Brandt et al., 2013). The Neolithic mtDNA diversity of Central Germany has been further studied, resulting in a genetic stratigraphy with four major prehistoric stages of the formation of the extant Central Europeans' gene pool. These were:

the initial colonisation of Central Europe by the LBK (event A), a bidirectional gene flux among Scandinavia and Central Europe [at ~4,100 cal BC a northward, and at ~3,100 cal BC a southward diffusion (event B1-B2)]. The last two Late Neolithic population genetic events have been connected to the Corded Ware culture (event C) and to the Bell Beaker culture (event D) in Central Europe (Brandt et al., 2013). The PCR based mtDNA studies have been extended to complete mitochondrial genome sequencing studies. Focusing on haplogroup H, similar conclusions have been drawn from the Central European transect to the assumptions of the Brandt et al. study (Brotherton et al., 2013).

A recent study has reported some mtDNA samples from the Near Eastern Neolithic (Pre-Pottery Neolithic B). MtDNA haplogroups K and N have been suggested as potential markers of the Early Neolithic expansion. The Near Eastern samples show more affinity to the Iberian Neolithic through the matches of the haplogroup N* than to the Central European LBK, where the haplogroup N1a seems to stand without antecedent in the Pre-Pottery Neolithic of Levant (Fernández et al., 2014). Nevertheless, the sample size is still very small from the earliest Neolithic Near East, which limits the possibilities of interpretation and comparative aDNA analyses.

MtDNA HVS-I records of 13 hunter-gatherers from Central and Northern Europe have shown discontinuity between the local Mesolithic hunter-gatherers and the 24 LBK farmers of the Haak et al. study (Bramanti et al., 2009). From Central Europe, further hunter-gatherer data have been published from Luxemburg (Loschbour), and Czech Republic (Dolní Věstonice) (Fu et al., 2013), and the Mesolithic aDNA dataset from Germany has been increased as well with sites of Oberkassel and Blatterhöhle (Bollongino et al., 2013; Fu et al., 2013). Furthermore, the mitochondrial gene pool of the Mesolithic is quite well researched on the Iberian Peninsula (Chandler et al., 2005; Hervella et al., 2012; Sánchez-Quinto et al., 2012). Single Mesolithic aDNA results from Sicily (Mannino et al., 2012) and the Gravettian aDNA results from the Paglicci cave (Caramelli et al., 2008, 2003) represent the maternal gene pool on the pre-Neolithic in Italian Peninsula. Recently, hunter-gatherer mtDNA data have been published from North-eastern Europe, supplementing and also diversifying the previously described Mesolithic substrate (Der Sarkissian et al., 2013) (see a map in chapter 4.3.1, Figure 10).

The number of published Y-chromosomes is still very small from the European prehistory (see a summary on Figure 5). Mesolithic genomic data have revealed the

dominance of the NRY haplogroup I [with sub-haplogroups I2a (L178)] in the Central-North European hunter-gatherer population, which is represented by two sites, Loschbour in Luxembourg and Motala in south Sweden (Lazaridis et al., 2014). NRY haplogroup C has been detected in the Mesolithic La Brana site in Spain (Olalde et al., 2014). Partial nuclear DNA data have also been published from French and Spanish Neolithic, arguing for a male diffusion through the Mediterranean route of Neolithisation (Lacan et al., 2011a, 2011b).

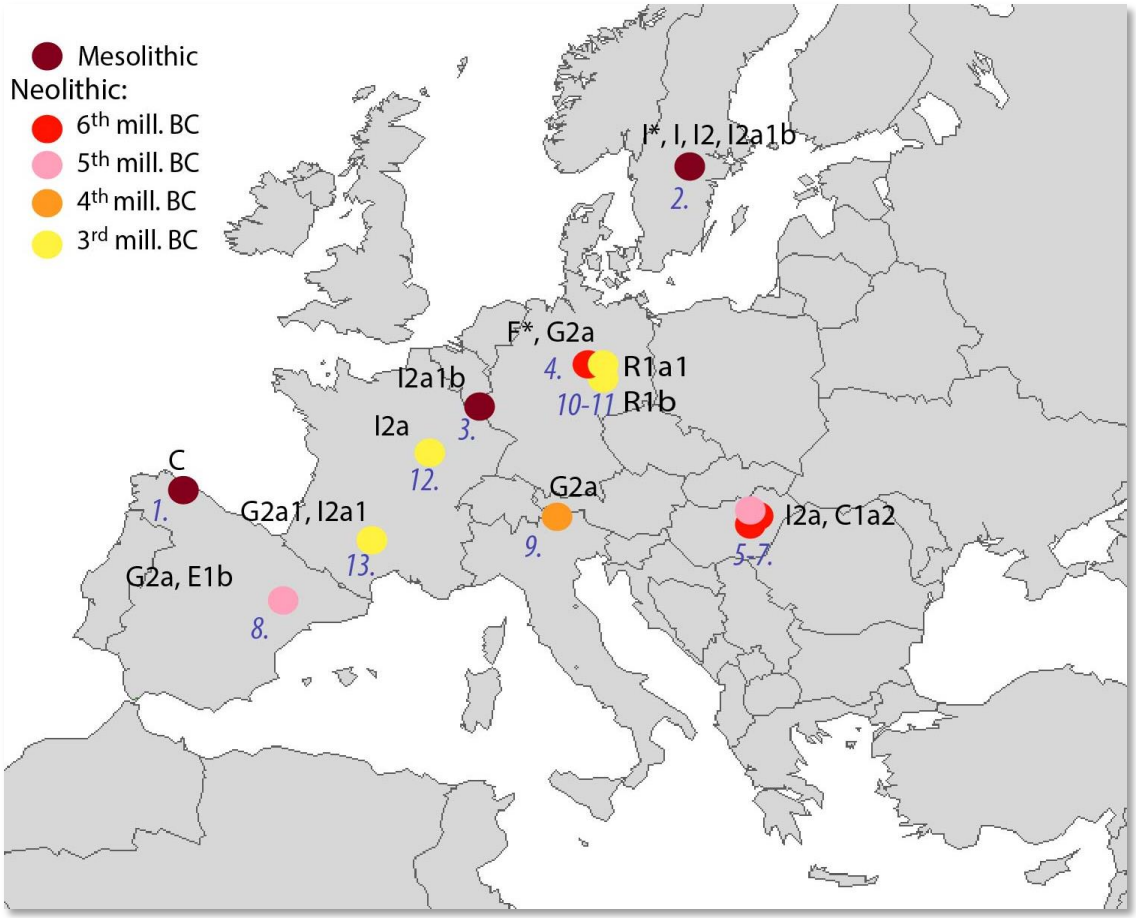


Figure 5. To date available Mesolithic and Neolithic Y-chromosomal haplogroup data in Europe.

Sites: 1. La Brana, 2. Motala, 3. Loschbour, 4. Derenburg, 5. Apc-Berekalja, 6. Tiszaszőlős-Domaháza, 7. Kompolt-Kígyósér, 8. Avellaner cave, 9. Tisenjoch Pass, 10, Eulau, 11. Kromsdorf, 12. Pierre Fritte, 13. Treilles. References (Gamba et al., 2014; Haak et al., 2010; Keller et al., 2012; Lacan et al., 2011a, 2011b; Lazaridis et al., 2014; Lee et al., 2012; Olalde et al., 2014) are detailed in the text.

Besides the majority of haplogroup G2a (P15), some E1b (V13) and I2a (P37.2) were detected in Treilles (south France) and in the Avellaner cave (Spain) (Figure 5). M. Lacan has also described a third site in her dissertation, Pierre Fritte in France (2,750 -2,725 cal BC), where she has defined two I2a1 individuals by short tandem repeat analyses (Lacan, 2011). Three LBK NRY data from the Central Germany have been published to date [two F* (M89), and one G2a (S126)] (Haak et al., 2010), and three haplogroups R1a1 (SRY 10831.2) from individuals belonging to the Corded Ware culture from site Eulau (Haak et al., 2008). Furthermore, two R1b (M269, M343) have been described from the neighbouring Bell Beaker site Kromsdorf (Lee et al., 2012). The Iceman from Austria (also called as Ötzi, lived around 3,350-3,100 cal BC), belonged to the NRY haplogroup G2a, with a terminal SNP of L91 (Keller et al., 2012). A recent study from the Hungarian Alföld region describes one NRY I2a from the Körös culture, one C6 from the Alföld distribution of the LBKT, one C6 from the ALBK, and one I2a from the Alföld distribution of the Lengyel culture (Gamba et al., 2014).

Besides PCR based studies, more and more genome-wide shotgun sequencing studies aim to investigate Mesolithic and Neolithic specimens. P. Skoglund and his colleagues have obtained genomic DNA from three hunter-gatherers from the Pitted Ware culture, and from one farmer (Gök4) belonging to the *Trichterbecher* culture in Scandinavia. The similarity between the single farmer genome and the modern southern Europeans in contrast to the similarity of the hunter-gathers to the extant North Europeans has supported the migration theory of the farmers from the south (Skoglund et al., 2012). The extended study of the Scandinavian Neolithic farmers and hunter-gatherers has revealed that the hunter-gatherers had a lower genetic diversity than the farmers (Skoglund et al., 2014). From the two millennia older Iberian hunter-gatherer samples complete mitochondrial genomes and partial genomic data have been retrieved (Olalde et al., 2014; Sánchez-Quinto et al., 2012). The genetic discontinuity between the Mesolithic and the Neolithic populations has been supported by Bayesian simulations, although it has still been based on HVS-I sequences. The 20-50,000 informative SNPs from the nuclear genome has shown that these two Mesolithic individuals are neither related to current populations of the Iberian Peninsula nor to Southern Europe (Sánchez-Quinto et al., 2012). In a recent study, J. Lazaridis and his colleagues have estimated 0-45% western hunter-gatherer ancestry in the early farmers, and they have determined the highest affinity of the farmers to the modern day south Europeans, especially to the

Sardinians. This 'early farming' group was composed of genomic data of an LBK woman (Stuttgart), the Tyrolean Iceman, and the previously described Scandinavian farmer (Gök4) (Lazaridis et al., 2014).

1.9 Osteological evidence about the Neolithisation of Europe

During the early-mid 20th century, the prevailing consensus among anthropologists was that the main transition in cranial morphology, occurring at the Mesolithic- Neolithic transition, involved a shift from dolichocephalic to brachycephalic morphology (Jelinek, 1973; Mikič, 1989; Schwidetzky, 1969). This assumption persisted in the research of the Neolithic Carpathian Basin as well, and the racial typology (Cromagnoid, Mediterranean, Nordic etc.) is still in use in anthropological studies.

The people of the Neolithic Carpathian Basin have been intensively studied with different osteological methods. Besides demographic (Köhler, 2013) and palaeopathological studies (Köhler et al., 2012), systematic craniometric analyses have been obtained on the well preserved Neolithic skeletons by Zs. K. Zoffmann. She has processed most of the Neolithic and Chalcolithic skeletons excavated in Hungary, focusing her study on the taxonomical aspects of osteology. With Penrose distance analysis (Penrose, 1954) Zs. K. Zoffmann has compared craniometric data of every Carpathian Basin Neolithic culture, and she has also analysed skeletal remains of Mesolithic population in the Iron Gates area (today's Serbia, Romania), and specimens from the Serbian distribution of the Starčevo culture. Based on these comparisons, she has made several attempts to reconstruct the population history of the prehistoric Carpathian Basin (K. Zoffmann, 2012; Zoffmann K., 2005, 2004). Collecting 56 Starčevo skeletons until 2004 from the former Yugoslavia and Hungary, she determined the Starčevo skull series as being mainly gracile Mediterranean, and according to the Penrose analysis, they show no connection to Lepenski Vir or to the following Vinča culture (Zoffmann K., 2004).

Zs. K. Zoffmann has concluded that the Earliest Neolithic Starčevo and Körös cultures neither had genetic connections to the Mesolithic population, nor to each other, nor to any of the following Neolithic and Chalcolithic populations. The LBK data shows connections to the Vinča (Serbian), Lengyel (Mórágy, Aszód), Tisza and Bodrogkeresztúr datasets, building a cluster of "Carpathian Basin". The classic Alföld LBK, and the Alföld LBK-Bükk group set off

from this cluster, showing North-eastern connections. Zs. K. Zoffmann has concluded- *"..the Early Neolithic Starčevo groups arriving to the Carpathian Basin did not interbreed with the local population to any significant extent, even though the latter adopted food production and sedentism from them, eventually rather absorbed the newcomers in the course of biological interbreeding. The earliest components of the closed Neolithic unit of the Carpathian Basin indicated by the Penrose analysis were the LBK populations of the western Carpathian Basin and Bohemia, whose origins can probably be derived from the local Mesolithic"* (Zoffmann, 2012, 314.).

When considering the results of Zs. K. Zoffmann, we have to keep in mind that osteological remains are often very fragmentary. Skulls with measurable features represent only a small part of the whole series, which limits the possible applications of osteological methods. Furthermore, the recently excavated large skeletal series from the Starčevo culture (Bánffy et al., 2010), the Sopot culture (Alsónyék-elkerülő, see chapter 3.1.1 in this study) and from the south Transdanubian distribution of the Vinča culture (Jakucs and Voicsek, 2014) would probably give some additional light to the taxonomical analyses as well.

Another issue is the methodology, the actually measured shape and size indices. The typological studies of basic measurements can underestimate the extent of intra-population variability, and morphological overlap between populations. In addition, craniometric research has found that the majority of cranial shape variation is the result of neutral evolutionary factors, whereas aspects of cranial size have been shown to fit a model of climatic differences (see a summary in Pinhasi and von Cramon-Taubadel, 2012).

Craniometric quantitative phenotypic traits have also been used for the reconstruction of the Mesolithic/Neolithic transition on a European scale by R. Pinhasi and N. von Cramon-Taubadel. Altogether 116 Epipalaeolithic, Mesolithic, and 165 Early Neolithic crania were studied from Southwest Asia and Europe (with some Körös and Alföld LBK skeletons). The statistical evaluation of the craniometric distance matrices has led to the conclusion that farming spread rather through demic diffusion of the farmers than through adoption of new technologies by the foragers (Pinhasi and von Cramon-Taubadel, 2009). These results have been complemented by further craniometric studies on the skeletal remains from the outlying (particularly the circum-Baltic) regions of Europe. R. Pinhasi and N. von Cramon-Taubadel have concluded that there were relatively low gene flow between contemporaneous populations of farmers and hunter-gatherers, which supports a mosaic model of Neolithisation, instead of

simple demic diffusion or cultural diffusion models in these eastern and northern regions of Europe (von Cramon-Taubadel and Pinhasi, 2011).

Comparing the methods of Zs. K. Zoffmann and R. Pinhasi, we can conclude that different sets of skeletons were studied, but the measured indices were partially the same. Whereas Zs. K. Zoffmann worked with ten cranial size indices, R. Pinhasi included six further traits (such as maximum and minimum frontal breadth, basion-nasion and basion-prostion length, bi-auricular and bi-asterionic breadth). The evaluations of the data were different, determined by the used statistical methods.

Although it is difficult to assess the validity of osteological results in the questions of population dynamic and movements, the works of Zs. K. Zoffmann and R. Pinhasi have to be reviewed as important anthropological antecedents of my research.

1.10 Demography of the Neolithic transition

A radiocarbon-date-based study have shown that European populations in the later Mesolithic were at historically low levels, presumably because the growing forest-coverage resulted in decreasing animal population densities (Gamble et al., 2005; Gkiasta et al., 2003).

An increase in fertility has long been proposed as either accompanying or initiating the Neolithic transition. As J. Bocquet-Appel pointed out, the proportion of 5-19 year old juveniles rose in the earliest Neolithic populations. This meant a growth of the total fertility rate of two births per woman. This demographic shift is called as Neolithic/Agricultural Demographic Transition by J. Bocquet-Appel (Bocquet-Appel, 2011). The shift to agriculture might have favoured population growth, because the food quantity rose, giving more existential safety than hunting and gathering. The carrying capacity of the land rose due to the produced surpluses. Whereas J. Bocquet-Appel assumed high-quality food, S. Shennan has stressed that the dependence on agriculture led to poorer diets and a greater incidence of infectious disease (Shennan, 2009). According to J. Bocquet-Appel, the early farmers had a population growth at first, which was certainly followed by enlarged child mortality but only with a time lag between the two phenomena. Causes of increased infant mortality would include lack of drinking water supplies, contamination by faeces, as well as reduced breastfeeding. The susceptibility of humans to new infectious diseases would have resulted from complex factors such as

modified exposure to animals, microbial adaptation, nutritional status and density of the population (Bocquet-Appel, 2008).

The above outlined reduced population control resulted in demographic revolution, which can be a rationale of the migration or diffusion, as argued by several scholars (Ammerman-Cavalli-Sforza, 1984, Renfrew, 1987). J. Bocquet-Appel has also considered that the population increase triggered the spread toward new geographic regions (Bocquet-Appel, 2011). On the other hand, ethnographic studies reveal a far less population growth-rate difference between foragers and farmers, and a wider overlap of population densities, as it is argued by the migrationists (Zvelebil, 2001).

Another aspect of this demographic event is that in agricultural societies juveniles become productive earlier. Due to the larger numbers of offspring, the farmers invested less energy into child raising than the foragers (Shennan, 2009). The children not only became productive earlier but also started working earlier, which further supported the new social-economic system.

Demography of the Neolithic transition and the modelling of the LBK dispersal are two closely related fields of studies. If the migration of the first farmers from southeastern into Central Europe is assumed, it has to be supported by demographic (archaeological) evidence as well. From several estimations, only two studies are highlighted here. In the first, J. Petrasch has estimated a population growth rate of about 1.4% pro year in the Older/*Ältere* LBK during the expansion phase. He based his deterministic estimation on settlement data, presuming hunter-gatherers in the origin territory, as ancestors of the LBK people (Petrasch, 2001). Nevertheless, his model has not accounted for the uncertainty associated with archaeological parameters. The intensive participation of the Mesolithic population in the formation of the Central European LBK has been assumed by P. Galeta and J. Bruzek as well. In their stochastic demographic model, the growth rate of the farming population ranges from 0.64% to 1.96% per year, and the estimates of total fertility rates vary from around 6 to 13 children per woman. These rates are higher than observed in other ethnographic and demographic studies, and that of the critical values, which led the scholars to come to the conclusion, that "*it is more likely that LBK fertility was not high enough to allow farmers to spread over Central Europe without admixture with local foragers*"(Galeta and Bruzek, 2009). In their recent study, congruent with the previous one, Galeta et al. have described three criteria for the LBK colonisation from the demographic point of view: "(*1) more than 37% of women survived to*

mean age at childbearing, (2) Neolithic expansion in Central Europe lasted more than 150 years, and (3) the population of farmers grew in the entire settled area." Otherwise, the Neolithic dispersal has to be regarded as a result of LBK-hunter-gatherer admixture in their view (Galeta et al., 2011).

Summarizing the above mentioned, there is an inconsistency between archaeogenetic (Bramanti et al., 2009; Haak et al., 2010), craniometric data (Pinhasi and von Cramon-Taubadel, 2009), and demographic modelling of the LBK dispersal at the current state of research. Recent or new excavations in the LBK source region (western Carpathian Basin) and further radiocarbon data might refine the contradictory picture.

1.11 Records and theories of residential rules and marital systems in the Neolithic Europe

Social and cultural anthropology distinguishes different marriage and residential systems. The review of these categories is important for the proper evaluation of the Neolithic maternal and paternal genetic variation. The first basic social anthropological term is the exogamy, which means that marriage is allowed only outside of a given social group. Such group can be a family, clan, village, or tribe as well. Its opposite is the endogamy, when marriages occur only within a specific social group. Whereas the first results in increasing genetic variation, the second leads to inbreeding and loss of genetic variation within a certain group.

The residential rules are closely related expressions to the previous ones, referring to social systems. Patrilocal residential rule means that women move to their husband's birthplace after the marriage. Ethnographic studies have suggested a shift in residential rules at the advent of Neolithisation, showing different trends among modern foragers and nonforagers (Marlowe, 2004). Increasing sedentism promotes territorial defence and control of resources, favouring men in the inheritance of land and property, which consequently led to patrilocal post marital residence system (Marlowe, 2004). These rules are closely connected to a system of descent along the father's line (patrilineality) in farming communities.

Patrilocality has also been suggested for the Neolithic in recent bioarchaeological studies. It has been supported by aDNA evidence for the Treilles Neolithic community (Lacan

et al., 2011a), and by stable isotope studies for the LBK and Corded Ware culture in Central Europe, detecting higher $Sr^{87/87}$ isotope rate variability among women than among men (Bentley, 2012; Bentley et al., 2012; Haak et al., 2008).

An interesting aspect of this theory is the marriage alliances between hunter-gatherer women and farming men, as part of the "frontier mobility" (Zvelebil, 2001). If it prevails in a unidirectional way (foragers to farmers), it can lead to a population where the paternal lineages originate from the farmers, but the maternal gene pool becomes diluted by the forager women.

Recent model-based statistical analyses of contemporary NRY and mtDNA data have revealed a shared admixture history for men and women, but not the same demographic history. R. Rasteiro and L. Chikhi have shown that female had a larger effective population size, which was likely based on differential effects of social and cultural practices (Rasteiro and Chikhi, 2013). The sedentism of the farmers possibly led to a decrease in male gene flow, whereas female gene flow would either have remained constant or would have increased (by exogamy of patrilocality) (Rasteiro and Chikhi, 2013). In the light of these results, aDNA data is especially useful in considering the possible sex specific patterns of admixture and demographic history of the Transdanubian Neolithic populations.

2 Aim of this study

2.1 Introduction to the project "Population history of the Carpathian Basin in the Neolithic period and its influence on the colonization of Central Europe"

My thesis is a part of a large multidisciplinary project, titled as "*Bevölkerungsgeschichte des Karpatenbeckens in der Jungsteinzeit und ihr Einfluss auf die Besiedlung Mitteleuropas*", translated in English as "Population history of the Carpathian Basin in the Neolithic period and its influence on the colonisation of Central Europe". It was funded by the German Research Foundation (DFG) between 2010 and 2014, led by Prof Dr Kurt W. Alt and Prof Dr Eszter Bánffy. Over the three years of research in Hungary and Germany (Mainz), our team (see the list of partners and team members in chapter 12) collected DNA and isotope (Sr^{87}, C^{13}, N^{15}) samples from more than 600 skeletons, dated from the Late Mesolithic to the Middle Chalcolithic. The sampling was focused on the territory of today's Hungary, but we got access to samples from Croatia and Slovakia as well. The DNA investigations were split up to two regions (western Carpathian Basin and Alföld or eastern Hungary) between my colleague Victoria Keerl and myself. The whole project was aimed to study and reveal possible population changes or population dynamic events behind the cultural changes of the Neolithic epochs in the Carpathian Basin. Using the available pre-Neolithic mtDNA data and the information and experience from the precursor German Research Foundation project of our team (entitled as „*Kulturwandel = Bevölkerungswechsel? Die Jungsteinzeit des Mittelelbe-Saale-Gebietes im Spiegel populationsdynamischer Prozesse*", or in English "Culture changes = Population changes? The Early Neolithic Middle-Elbe-Saale region in the light of population dynamic processes") we searched for evidence in the questions of Mesolithic/Neolithic transition, spread of farming, and defining Neolithisation routes in Europe.

2.2 Scope of the dissertation: ancient DNA study of the Neolithic Transdanubia

The purpose of my own investigation was to reveal the genetic background of the sixth-fifth millennia BC cultural changes in the western part of the Carpathian Basin, concentrating mainly on western Hungary. The data analyses were complemented by the research of V. Keerl

(Keerl, 2014), whose results have provided a rich database of comparative ancient mtDNA data from eastern Hungary.

Specific topic of my study was the search for the genetic traces of the Neolithisation process in the Carpathian Basin and in Central Europe. The postulated Near Eastern origin of Central Europe's LBK farmers has only been inferred from modern-day population data so far, while the contemporaneous genetic diversity of potential source populations, on the vast territory from the Fertile Crescent to Central Europe has remained almost unknown. The western part of the Carpathian Basin was a crucial region on this route, an intermediate stop in the first half of the sixth millennium BC, before the farmers spread further toward northwest on the continent. The Carpathian Basin remained a zone of contacts and interactions in the last centuries of the sixth and during the fifth millennium BC as well. The study region was subsequently reached by cultural effects and possibly by migrations from southeastern Europe, and the people of Transdanubia probably acted as mediators toward the northern regions of Central Europe.

Consequently, my aims were to

i) study the genetic diversity of the Early, Middle, and Late Neolithic Carpathian Basin cultures' populations (sixth-fifth millennia BC) from both the mtDNA and Y-chromosome perspectives, and specify the main population genetic events during this period
ii) look for geographic patterns within the Transdanubian genetic data
iii) compare the results with prehistoric data, especially with the European pre-Neolithic and the Central European Neolithic datasets
iv) investigate the contribution of the southeastern European Starčevo culture's people to the genetic variability of the LBK and its successor populations in Central Europe
v) study the subsequent periods' Neolithic genetic legacy in the Carpathian Basin
vi) reveal the potential genetic origins of the Transdanubian Neolithic populations, by comparison with prehistoric and modern-day genetic data
vii) investigate, whether men and women had different Neolithisation or migration histories through the comparison of the observed mtDNA and Y-chromosomal diversities

3 Materials

3.1 Archaeological sites and dating

Thirty-two sites from the Neolithic and Copper Age were chosen for ancient DNA analyses of this study. The concept of our team was to sample the major sites from these periods, covering all cultures from the early Neolithic to the Middle Copper Age, and study the possibly largest area of each culture's distribution territory. We aimed to collect approximately equal amount of samples from the Starčevo, LBKT, Sopot, Lengyel and Balaton-Lasinja cultures, considering length of their existence in Hungary, and the records about their settlement structure and population density. The Vinča culture, with two south Transdanubian sites came additionally to the above listed five cultures, whose appearance in Hungary has been discovered only recently (Jakucs and Voicsek, 2014).

Our team sampled the major Starčevo sites in Transdanubia. Only a few Starčevo burials remained unsampled from this epoch. Furthermore, the Starčevo sample set was rounded out by three sites from the Croatian Syrmia region. The major LBKT sites (with more than three burials per site) were chosen from all around Transdanubia. All LBKT sites in Hungary were settlements, with sparse burials among the houses. The Slovakian site Nitra is the only graveyard, which was separated from the LBK settlement. All studied LBK sites from Transdanubia completed with three Nitra samples are discussed as LBKT. From the next epoch, our team collected almost every Sopot skeletal remains from Hungary, which is known to be scattered in this region. With 21 graves at the site Alsónyék-elkerülő 2. lh. we managed to sample the largest excavated Sopot graveyard. The succeeding Lengyel culture, represented by large cemeteries, was the only one, where we had the luxury of choosing between sites and grave groups. Here we collected samples from larger sites and from different regions of Transdanubia. At sites as Bátaszék-Alsónyék or Mórágy-Tűzkődomb, where we could not analyse all the excavated graves, we sampled specific grave groups, considering the material's preservation. There are only a few published excavated series from the Balaton-Lasinja culture. Most of the remains are either cremation burials or not real burials but sacrifices/victims of special events as for example house foundation. Here we focused on continuity (as in the case of Veszprém) or specific ritual or kinship (as in Keszthely) issues during the sampling process.

Most of the studied archaeological sites were settled in several archaeological periods, which circumstance made it difficult to date burials without grave goods or archaeological information. At these ambiguous cases, we made the human skeletal remains measured in two laboratories (CEZA in Mannheim, Germany and Beta-Analytic in Miami, USA). Any other graves were datable by grave goods, archaeological context, and stratigraphic position (further details in Supplementary Table 1).

3.1.1 Sampled archaeological sites in detail, listed in alphabetical order

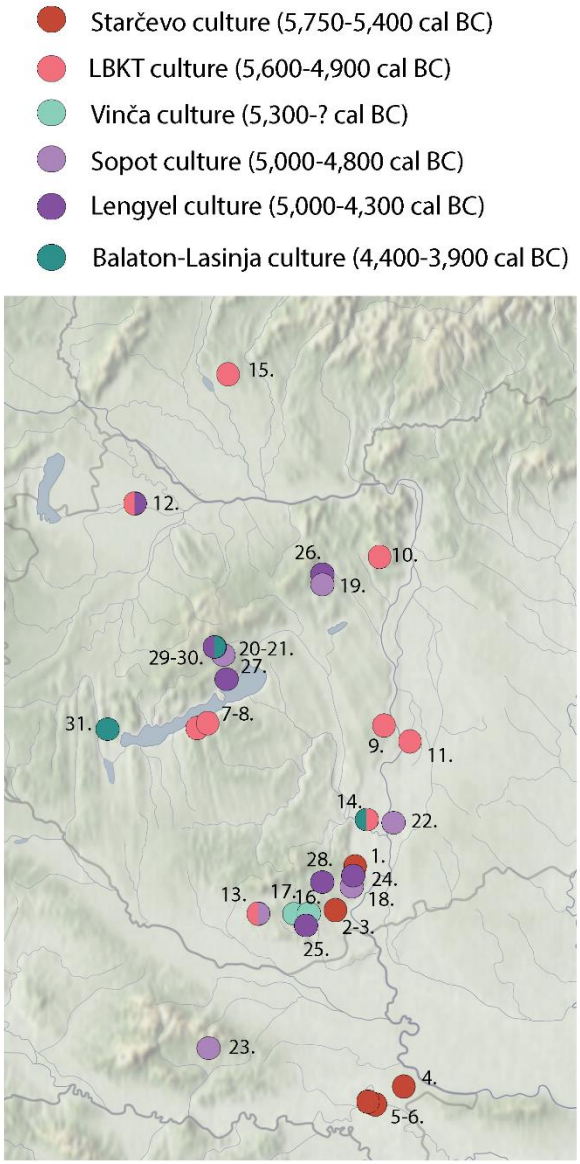
Alsónyék-Bátaszék, Mérnöki telep

Between 2006 and 2009, two settlement periods of the Starčevo culture, features of the LBK and late Neolithic Lengyel cultures were unearthed in Mérnöki telep by the Institute of Archaeology, HAS (excavators: A. Osztás and I. Zalai-Gaál). The uncovered Starčevo settlement reached the extension of 80 hectares. Out of the 1568 excavated features more than 400 surely belonged to the Starčevo culture (from the Linear B-Spiraloid B phases), concentrated mainly in the southern part of the area. The majority of the features were pits, in various shapes and sizes. Besides pits, ditches and several types of ovens were excavated. Altogether 26 burials belonged to the Starčevo period on this site. They were found either inside of ovens or in pits, in contracted position, without any trend in body orientation. Only one burial contained grave good (Bánffy et al., 2010). The stratigraphic uncertain skeletons, buried without grave goods, were radiocarbon dated in the CEZA laboratory in Mannheim, Germany (Supplementary Table 1).

Alsónyék-elkerülő 2. lh.

Another sub-site of the site complex Alsónyék-Bátaszék is called Alsónyék-elkerülő or Alsónyék-Hosszú-dűlő, 2. lh. It was excavated in 2008 by the Wosinsky Mór Museum (excavator: J. Ódor), and its archaeological material is still under evaluation. The people of the Sopot culture settled on this area, in the early fifth millennium cal BC. Beside photos of grave goods, the radiocarbon dating was used to omit skeletons from the later periods. We have managed to sample 18 skeletons; all of them have been safely assigned to the Sopot culture. This means an exceptional number of Sopot individuals, there is not any other known Sopot

site with so many burials in the whole dissemination area of this Late Neolithic culture (J. Ódor, T. Marton, A. Osztás personal communication).



Site number	Site name
1	Alsónyék-Bátaszék, Mérnöki telep
2	Lánycsók-Csata-alja
3	Lánycsók-Gata Csotola
4	Vinkovci-Nama
5	Vinkovci-Jugobanka
6	Vukovar-Gimnazija
7	Balatonszárszó-Kis-erdei-dűlő
8	Balatonszemes-Bagódomb
9	Bölcske-Gyűrűsvölgy
10	Budakeszi, Szőlőskert-Tangazdaság
11	Harta-Gátörház
12	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II
13	Szemely-Hegyes
14	Tolna-Mözs
15	Nitra
16	Versend-Gilencsa
17	Szederkény-Kukorica-dűlő
18	Alsónyék-elkerülő 2. lh.
19	Bicske-Galagonyás
20	Nemesvámos-Kapsa utca
21	Nemesvámos-Baláca
22	Fajsz-Garadomb
23	Radovanci
24	Bátaszék-Lajvérpuszta
25	Borjád-Kenderföldek
26	Csabdi-Télizöldes
27	Felsőörs-Bárókert
28	Mórágy-Tűzkődomb, B1
29	Veszprém-Felszabadulási út
30	Veszprém-Jutasi-Munkácsy út
31	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő

Figure 6. Location of the studied 31 Neolithic and Chalcolithic sites from the western Carpathian Basin.

The sites are grouped by cultures, which are listed in a chronological order, presented in the upper left corner of the figure. The calibrated BC dates refer to the approximate periods of the Transdanubian disseminations of the cultures.

Balatonszárszó-Kis-erdei-dűlő

During a rescue excavation on the motorway track M7, approximately ten hectares of an LBKT settlement were excavated between 2001 and 2006 by the Institute of Archaeology,

HAS (excavators: K. Oross and T. Marton). Besides 59 post houses 43 LBK burials and a 160 m long neolith ditches were recovered at the south part of the LBK settlement. In the pottery sequence, examples of the early LBKT (*Bicske–Bíňa/Bény* phase) and typical vessels to the younger *Keszthely* and *Zselíz* group are known (Marton and Oross, 2010). The deceased were interred on the settlement, in upper parts of garbage pits. The orientation of the graves did not show any tendency. Except two skeletons, the deceased were buried in contracted positions. Some of them had vessels, shell items (spondylus) and stone artefacts as grave goods. 16 skeletons were radiocarbon dated in the ORAU laboratory of Oxford (as part of the Life ways project of A. Whittle) and one in the CEZA laboratory in Mannheim (Supplementary Table 1, Oross and Marton, 2012; Zoffmann, 2012).

Balatonszemes-Bagódomb

On the site Balatonszemes-Bagódomb, traces of several cultures came to light between 1999 and 2001, during a rescue excavation by the Directorate of Somogy County Museums (excavator: V. Kiss). Traces of the LBKT, Balaton-Lasinja culture, Roman period (4th century), and four Avar burials were uncovered on a total surface of 27 hectares. The LBKT population lived here for a longer period: from the early *Bicske-Biňa* phase up to the late *Keszthely* phase of the culture. Evidence on one post-built house was found, and 10-15 further houses were deducted from the settlement plan.

There were six LBKT graves uncovered; four of them were buried in separate pits, and two in upper parts of larger storage pits. All deceased were laid in contracted position. An adult man was covered with ochre in grave 410. One stone blade was found as single grave good (Kiss and Sebők, 2007).

Bátaszék-Lajvérpuszta

A grave group at the site Bátaszék-Lajvérpuszta is dated to the late Neolithic Lengyel culture. It was excavated in 2009 by the National Heritage Protection Centre of the Hungarian National Museum, Department Pécs (excavator: V. Majerik). The archaeological evaluation of the findings is still ongoing (J. Ódor, personal communication). We sampled and studied 25 out of the 34 Lengyel period skeletons, using the skeleton's preservation as sampling criterion.

Bicske-Galagonyás

Ten skeletons were uncovered on a Sopot settlement in three excavation campaigns, in years 1933, 1959, and 1974. The osteological remains from 1933 and 1959 are fragmentary

and defective (Zoffmann, 1978). The site was uncertainly dated in the first publications (F. Petres, 1959). First in the 1970es was the material assigned to the Sopot culture (Makkay, 1975; Makkay et al., 1996).

Borjád-Kenderföldek

During a small-scale rescue excavation, one richly furnished Lengyel grave was excavated by the Janus Pannonius Museum in Pécs in 2010 (excavator: G. Bertók). Our team had the opportunity to sample the skeleton directly after the uncovering. The grave contained pottery, copper beads, and traces of a funeral construction (Cs. Gáti, G. Bertók personal comm.).

Bölcske-Gyűrűsvölgy

On the M6 motorway track at Bölcske-Gyűrűsvölgy, 1.6-hectare surface was excavated in 2008 by the ELTE University (excavator: K. Sebők). Altogether 257 features from four archaeological periods (LBKT, Bronze Age, La Téne B-C, and Árpád Age) were uncovered on the site. LBKT features from the *Zseliz/Želiezovce* phase with typical red-yellow painted pottery and Tisza import ware were reported from the site. Five LBKT graves came to light: from a clay pit (feature no. 36) a triple grave, and from another pit a double burial. Two grave goods are mentioned: a large globular pot and a vessel with *Notenkopf* decoration (Sebők, 2013, 2008).

Budakeszi, Szőlőskert-Tangazdaság

In the site Budakeszi, Szőlőskert-Tangazdaság an early Neolithic and one Early Bronze Age settlements were unearthed in 2006 by the Directorate of Pest County Museums (excavator: A. Czene). On a surface of one hectare, 632 features were excavated; most of them belonged to the *Notenkopf* and *Zselíz/Želiezovce* phase of the LBKT. Several clay pits and storage pits were found, with large quantity of animal bones. A segment of a building construction was indicated by a sequence of postholes. On the area of the LBKT settlement, 15 burials were uncovered. Most of them laid in pits, in contracted positions without any grave good. One of them had a rich inventory of vessels (six pieces, partly with *Notenkopf* pattern) (Czene and Ottományi, 2007; Czene, 2008). Four skeletons were dated in the CEZA laboratory in Mannheim. The ^{14}C results date the burials to the *Notenkopf* and *Zselíz/Želiezovce* phase of the LBKT (Supplementary Table 1). Anthropological analyses were performed by K. Köhler (Köhler, 2014).

Csabdi-Télizöldes

On the site Csabdi-Télizöldes, a settlement and a graveyard from the early phase of the Lengyel culture were excavated by J. Antoni between 1978 and 1986. 32 graves were found among the settlement features, which are only preliminary published (Antoni, 1982). The archaeological material remains unpublished. The anthropological analyses were done by K. Köhler (Köhler, 2004).

Fajsz-Garadomb

The site Fajsz-Garadomb was investigated from 2006 to 2008, as part of a collaboration between the Institute of Archaeology (HAS) and the University of Tubingen. The excavations were directed by E. Bánffy and J. Petrasch. The archaeological records are still under processing. LBK and Sopot settlement layers were found over the occupation traces of the Körös culture. Three graves of the Sopot culture were included in our project, dated by Sopot ceramic objects and stratigraphic observations (T. Marton personal comm.).

Felsőörs-Bárókert

In 2008, a large and multilayer Neolithic settlement was unearthed on a rescue excavation by the Directorate of Veszprem County Museums (excavator was J. Regenye). The area was settled from the second half of the sixth millennium BC through the fifth millennium BC. Five graves were unearthed from the second phase of the Lengyel culture, which are still under processing (Regenye in *Felsőörsi Hírmondó* 2011 V/2, personal comm.).

M85 Enese elkerülő 02. Kóny, Proletár-dűlő II

Altogether, a surface of 18,000 m^2 was excavated between 2008 and 2009 on the site Kóny, Proletár-dűlő II, by the Hungarian Field Service for cultural Heritage (excavator K. Varga). The 941 archaeological features are dated to the following cultures: LBKT, Lengyel, Bronze Age, Celtic, and Roman periods, Arpad age.

Two LBKT graves included pottery grave goods. One skeleton was found on the left side, in contracted position, NE-SW oriented (grave 55). The other skeleton laid on her back, with legs flexed under her body, in NE-SW position. A little cup and a bowl and a stone tool were her grave goods (grave 286) (Oross and Marton, 2012). In the grave 612 laid two children without grave goods. Our radiocarbon dating assigned this grave to the late LBKT period. One grave belonged to the early Copper Age by the radiocarbon dating (grave 223), and another had a

vessel assigning it to Lengyel III culture (grave 826) (K. Varga, preliminary report, K. Oross personal comm., for the ^{14}C dates see Supplementary Table 1).

Harta-Gátőrház

A rescue excavation of the Visky Károly Museum (led by R. Kustár) took place in 2002 and 2003 on a total surface of over two hectares at Harta-Gátőrház. The unearthed 1004 archaeological features belong to several periods: Middle Neolithic, Roman (Sarmatian), Arpad Age, and Modern Age. Most of the features are from the *Zselíz/Želiezovce* phase of the late LBKT. Six LBKT longhouses were detected in parallel orientations with the adjacent longpits. Six LBKT graves were unearthed, two in refilled storage pits, two in clay pits and two in separate pits. Four skeletons were in crouched position. Two of the burials contained vessels as grave goods, and in one a spondylus jewellery was found (Kustár, 2003).

Lánycsók – Csata-Alja

The site was found in 2008, in the frame of a rescue excavation before constructing the motorway M6. On two hectares, 433 Starčevo, Balaton-Lasinja, Vučedol, Roman and Árpád Age archaeological features were unearthed by the Janus Pannonius Museum in Pécs (Vajda-Kiss, 2009). After radiocarbon dating of the supposed Starčevo specimens in CEZA laboratory in Mannheim, three burials were surely assignable to the Starčevo period, and the rest is from the Balaton-Lasinja and from the Early Bronze Age Vučedol period (see Supplementary Table 1).

Lánycsók-Gata Csotola

Remains of several archaeological periods were excavated in 2008 at Lánycsók-Gata Csotola: prehistory (Early Neolithic, Copper Age, Middle and Late Bronze Age, and Celtic period), Migration Period (Avar and Conquest period graves) and Middle Ages. The excavation was led by V. Voicsek from the Hungarian Field Service for cultural Heritage, Department Pécs. Overall, 545 objects were unearthed on a surface of 3.5 hectares. About 30 Starčevo objects (pits, pit complexes, ovens, postholes, ditches) were found on the site. Typical Starčevo pottery for example low pedestals with applied barbotine and nail impressed ribs decorated shards, and clay bobs came to light from the pits. Out of 143 graves, five belonged to the Starčevo culture. They were found in pits, on the west part of the excavated area, in contracted positions (Voicsek, 2010).

Keszthely-Fenékpuszta Pusztaszentegyházi dűlő

In 2000, two pits of the Balaton-Lasinja culture were unearthed in Keszthely-Fenékpuszta, with several human skeletal remains. J. P. Barna led the excavation from the Balaton Museum in Keszthely. In feature 45, seven sculls were uncovered besides other stray human bones. In feature 46, along the pit wall, child pectoral bones in anatomical order came to light. In the fill of the pit, several other human bones were found as stray findings. The anthropologist Zs. Zoffmann has managed to classify the bones to 14 children, two men, and three women. Kinship relations between the sacrificed, killed or reburied individuals could not be defined because of the fragmentary state of the skeletal remains (Zoffmann K, 2005).

Mórágy-Tűzkődomb B

Between 1978 and 1990, Lengyel settlement and 91 Lengyel graves were unearthed in Mórágy-Tűzkődomb, in grave group B1 by the Institute of Archaeology HAS (excavator: I. Zalai-Gaál). The graves are dated to the Lengyel II period (Zalai-Gaál, 2002). Zs. K. Zoffmann performed the anthropological analyses (Zoffmann, 2004).

Nemesvámos-Baláca

The site Nemesvámos-Baláca was excavated from 1976 onwards by Sylvia Palágyi from the Laczkó Dezső Museum. In year 1984, rich Neolithic finds came to earth beneath the Roman Age villa-farm. In pit L, a human skeleton laid on his back with his left hand behind the body. A vessel fragment close to the skeleton could have served as grave good. In the 1988/L room, a crouched skeleton was found in a deep oval shaped pit. The grave was covered with stones, and it contained some stone blades and vessel fragments. The grave belongs to the younger phase of the Sopot culture (Regenye, 1996b).

Nemesvámos-Kapsa utca

A richly furnished Sopot grave (contained pot, spondylus jewels, and beads) was excavated in 2002 by S. Palágyi. The grave is still under archaeological evaluation, and it neither has not been processed anthropologically (J. Regenye, personal communication).

Nitra

One of the largest LBK graveyard, the Nitra site was unearthed between 1964 and 1965 by J. Pavúk (Pavúk, 1972). The LBK settlement has not been excavated. The 76 graves formed rather lines then clusters, largely oriented north-west to S-E. The radiocarbon dating within

the *Lifeways* project of Prof. A. Whittle and his colleagues has defined a start date for the cemetery at around 5,370-5,220 cal BC and an end at 5,210-4,980 cal BC [two sigma values] (Seren Griffiths in Bickle and Whittle, 2013, p. 449).

Out of the 76 graves, we sampled 30 for the aDNA project, but only five of them were tested for aDNA preservation. The samples were assigned to the LBKT group, even if they are outlying from Transdanubia.

Radovanci

We have managed to sample one Sopot burial from the site Radovanci in Slavonian Croatia. The skeleton was uncovered by J. Balen and H. Potrebica in 2006 (Balen and Potrebica, 2006). The anthropological investigation was done by M. Šlaus (M. Šlaus, personal comm.).

Szederkény-Kukorica-dűlő

The site Szederkény was unearthed between 2005 and 2008 by the archaeologists of the Janus Pannonius Museum, Pécs (excavator: G. Kovaliczky). A Neolithic site of 9-hectare extension was identified by our colleagues in Szederkény, besides features of Copper, Bronze and Iron Ages. The archaeological material shows a mixed Sopot-LBK-Vinča characteristic. 64 LBK houses have been defined so far, and 54 Neolithic graves have been assigned to this period (Jakucs and Voicsek, 2014). Only 11 Vinča graves were involved in our project, because at the time of the sampling, the chronology of the rest of graves was uncertain. The five radiocarbon dates that were measured in the CEZA laboratory of Mannheim, fall between 5,380-4,850 cal BC on the two-sigma confidence level (see Supplementary Table 1).

Szemely-Hegyes

The site was excavated on the Baranya county part of the M6 motorway track in 2006 and 2007. On a surface of over four hectares ~1,400 archaeological features were documented from the *Keszthely* and *Notenkopf* phase of the LBKT, Sopot culture, and Copper age Balaton-Lasinja and *Furchenstich* periods (T. Paluch, K. Somogyi, J. Jakucs personal comm., Nagy, 2007). Twenty houses, pits, ditches, and ovens were uncovered on the LBKT settlement. From the ten Neolithic burials, we radiocarbon dated six in the Bioanalytic laboratory in Miami. Two graves are late LBKT and four are dated to the Vinča-Sopot phase (Supplementary Table 1). The population composition of the community lived in Szemely is especially interesting, because it is located in the contact zone of the LBKT, Vinča, and Sopot cultures' territories (J. Jakucs, personal comm.).

Tolna-Mözs

In 2008, the Institute of Archaeology HAS led a rescue excavation in Tolna-Mözs, along the Tolna county part of M6 motorway track (excavation was led by T. Horváth). LBKT, *Furchenstich*, Hallstatt, Celtic, Avar and Middle Age features came to light. It was possible to reconstruct 48 LBKT houses, from post-holes and longpits along the former walls. The settlement was inhabited from the early *Bicske/Biňa* LBKT phase to the *Notenkopf* LBKT phase. The archaeological materials of the houses show a sequence of occupation in different phases: the early house group contained ceramic material with connection to Vinča and late Starčevo pottery types, whereas the later northern house group contained mainly *Notenkopf* ceramic type. A LBK grave was uncovered from a longitudinal ditch (nr 1649). Based on the archaeological context and the radiocarbon date, the grave was younger than the houses of this settlement part. The second radiocarbon dated skeleton was found in a crouched position, without any grave good (Marton and Oross, 2010; Oross and Marton, 2012).

Vela Spila cave

Vela Spila cave is located in southern Croatia, on the Korčula Island. The excavations took place between 1986 and 2004, directed by D. Radič (Radic, 2005). Until 2004, five Mesolithic skeletons came to light from the cave. We sampled “Stanko”, who was unearthed in 2004 from the stratum 12. “Stanko” died as an adult man (Slaus, 2005) at around 7,200±30 BP (6,205-6,000 cal BC on two sigma level), dated by the material of layer 7/4 (Komso, 2006).

Versend-Gilencsa

The site Versend-Gilencsa was excavated by V. Voicsek from the Directorate of the Baranya County Museums, between 2005 and 2008. The settlement and graves of the Vinča culture were identified during the sampling campaign of our project. Twenty-five Vinča graves were involved into the aDNA analyses out of the excavated twenty-seven. We had four skeletons radiocarbon dated in the CEZA laboratory of Mannheim, their two-sigma values fall between 5,360-5,040 cal BC (J. Jakucs, V. Voicsek personal comm. see Supplementary Table 1).

Veszprém Felszabadulás út

In 1972, three graves from the early Copper Age phase of the Lengyel culture (Lengyel III phase) came to light. Two skeletons were in crouched positions, one on his back. The two adults have grave goods such as bone tool and pottery. The crouched skeleton of a 6-6.5-year-

old child was found under an apses shape building, who could have been a foundation sacrifice. The site was crucial in the research history of the Lengyel culture, since it enabled the definition of the third phase of the Lengyel culture, which persisted until the end of the Early Copper Age (Raczky, 1974).

Veszprém Jutasi út

In the crossroad of the Jutasi and Munkácsy M. street, 250 m away from Felszabadulási út, sixteen graves were unearthed from the Neolithic and Copper Age in 2003 by J. Regenye (Laczkó Dezső Museum). Eight graves belong to the Neolithic Lengyel culture and four to the middle Copper Age Balaton-Lasinja culture. The members of the small grave group were richly furnished with grave goods (especially noteworthy is grave 5). Graves 1-2 and 6-7 were double burials. The settlement was densely inhabited, but the area of the graves remained undisturbed. Seven out of eight Lengyel skeletons were found in crouched position, facing toward south, and one skeleton (grave 8) laid on her back. The Balaton-Lasinja double grave nr. 13-14 was covered with calcar stones, which shows a unique rite in Transdanubia (Regenye, 2006). We had four skeletons radiocarbon dated in the CEZA laboratory of Mannheim. They two sigma level values vary between 4,800-3,960 cal BC (Supplementary Table 1).

Vinkovci-Nama

K. Minichreiter and I. Janošić from the Vinkovci Town Museum (Gradski muzej Vinkovci) excavated the site Vinkovci-Nama in 1977. The Starčevo findings belong to the Linear B and Spiraloid B phases of the culture (Janosic, 1984; Minichreiter, 2002, 2001, 1992). Out of the seven excavated Starčevo skeletons, we included six specimens into the aDNA project.

Site number	Culture(s)	Site name	Sum of burials from the studied culture	Sum of sampled burials
1.	Starčevo, Lengyel	Alsónyék-Bátaszék, Mérnöki telep	27	27
2.	Starčevo, Balaton-Lasinja, Bronze Age	Lánycsók-Csata-alja	13	8
3.	Starčevo	Lánycsók-Gata Csotola	5	5
4.	Starčevo	Vinkovci-Nama	7	6
5.	Starčevo	Vinkovci-Jugobanka	5	4
6.	Starčevo	Vukovar-Gimnazija	5	4
7.	LBKT	Balatonszárszó-Kis-erdei-dűlő	43	23
8.	LBKT	Balatonszemes-Bagódomb	6	6
9.	LBKT	Bölcske-Gyűrűsvölgy	7	5
10.	LBKT	Budakeszi, Szőlőskert-Tangazdaság	15	15
11.	LBKT	Harta-Gátörház	6	5
12.	LBKT, Lengyel, Balaton-Lasinja, Bronze Age	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	12	7
13.	LBKT, Vinča, Sopot	Szemely-Hegyes	2	2
14.	LBKT, Balaton-Lasinja	Tolna-Mözs	4	4
15.	LBKT	Nitra	77	30/5
16.	Vinča	Versend-Gilencsa	27	25
17.	Vinča	Szederkény-Kukorica-dűlő	54	11
18.	Sopot, Bronze Age	Alsónyék-elkerülő 2. lh.	21	19
19.	Sopot	Bicske-Galagonyás	12	5
20.	Sopot	Nemesvámos-Kapsa utca	1	1
21.	Sopot	Nemesvámos-Baláca	2	2
22.	Sopot	Fajsz-Garadomb	4	3
23.	Sopot	Radovanci	1	1
24.	Lengyel	Bátaszék-Lajvérpuszta	34	25
25.	Lengyel	Borjád-Kenderföldek	2	1
26.	Lengyel	Csabdi-Télizöldes	31	30
27.	Lengyel	Felsőörs-Bárókert	5	5
28.	Lengyel	Mórágy-Tűzkődomb, B1	91	25
29.	Lengyel	Veszprém-Felszabadulási út	3	3
30.	Lengyel, Balaton-Lasinja	Veszprém-Jutasi-Munkácsy út	12	12
31.	Balaton-Lasinja	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	19	8+6 single samples
32.	Mesolithic	Vela Spila cave/Island Korčula	5	1

Table 1. Summary of the sampling.

Vinkovci-Jugobanka

K. Minichreiter and I. Janošić from the Vinkovci Town Museum (Gradski muzej Vinkovci) excavated the site Vinkovci-Jugobanka in 1977 and 1978. The unearthed Starčevo findings belong to the Linear B and Spiraloid B phases of the culture (Janosic, 1984; Minichreiter, 2002, 2001, 1992). Out of the five excavated Starčevo skeletons, we sampled four specimens for aDNA analyses.

Vukovar-Gimnazija

M. Dalic from the Vukovar Town Museum (Gradski muzej Vukovar) excavated this site in 1999. The material has typical features of the Spiraloid B phase of the Starčevo culture (Minichreiter, 2002). Four Starčevo skeletons were sampled for aDNA analysis.

3.2 Samples and sampling

In the frame of the above-described DFG project, including the former Balatonszárszó project (carried out between 2008-2009), our team has sampled more than 600 samples for aDNA analyses from Transdanubia, Croatia and Slovakia. We aimed to sample at least 50 individuals per culture, in order to gain statistically evaluable amount of data from each period of the Neolithic Transdanubia. This plan was modified by the fact, that some cultures (Vinča and Balaton-Lasinja) do not have so many uncovered graves in Transdanubia yet.

I processed 633 samples (two samples per skeleton on an average) between 2008 and 2013 in the aDNA laboratory of the Johannes Gutenberg University in Mainz (Figure 8). Whenever possible, teeth were favoured for aDNA analysis, otherwise pieces of compacta from long bones or petrous parts of temporal bones were sawn out using a cleaned diamond drill. We took two to five samples per skeleton, from different skeletal elements. Most of them (61.14%) were molars, just a small amount were other permanent teeth (7.9%). Among the long bones, femur was preferred, which came to 14.6% of the total sample amount.

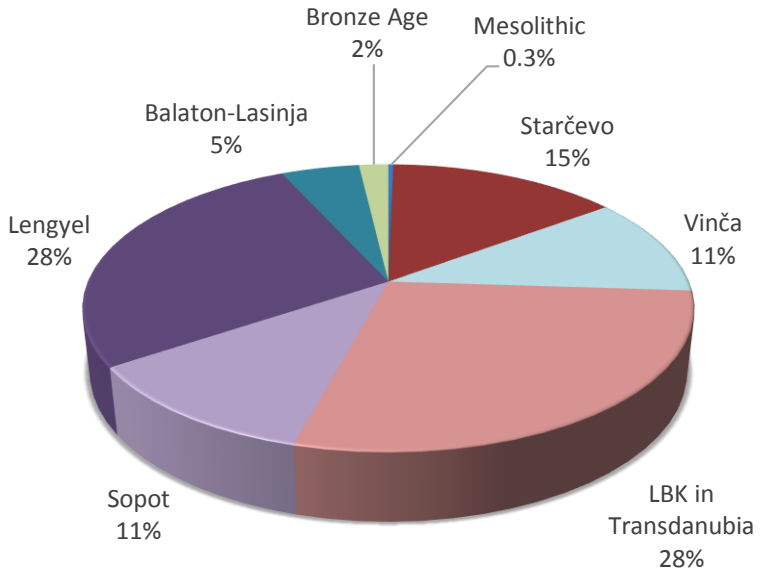


Figure 7. Distribution of the sampled individuals (n=323) per culture or population.

Nitra site was involved in this diagram, which was only partially processed (5/30 sampled skeletons).

With the exception of the Croatian sites, Mórágy-Tűzkődomb, Veszprém Felszabadulási út, and Csabdi-Télizöldes, all sites were excavated in the last decade, during rescue excavations before motorway constructions or other major investments. From eight sites, we managed to get samples before washing the bones (two of Lánycsók, Alsónyék-Bátaszék, Szemely-Hegyes, Tolna-Mözs, Versend-Gilencsa, Szederkény-Kukoricás, Fajsz-Garadomb). They could not have been contaminated with modern DNA, except during the excavation itself. The rest of the skeletons had been washed and anthropologically processed before the sampling. In each case, we took and analysed swab samples from the anthropologists, washerwomen, and archaeologists who had contact with the skeletons (see for their haplotype data).

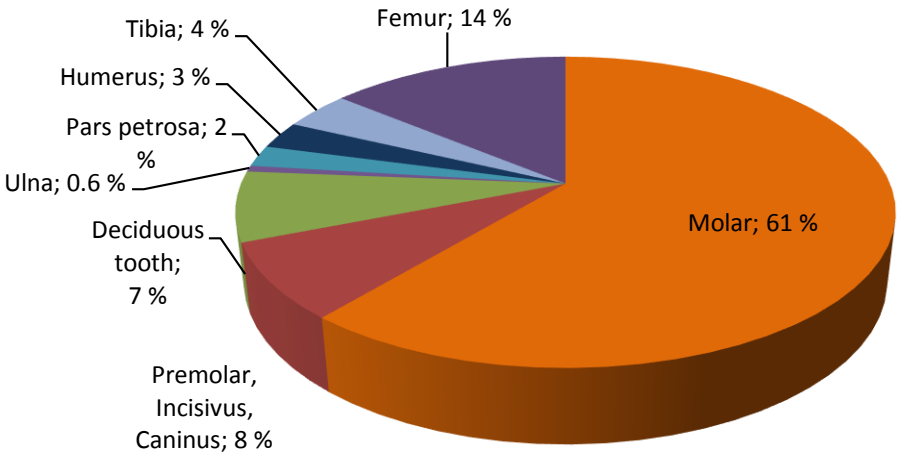


Figure 8. Frequencies of the processed sample types.

All processed samples (A-B-C-D samples per individual) are involved (n=633).

The sampling was performed by J. Jakucs, M. Fecher, V. Keerl, and myself with the assistance of anthropologists B. G. Mende and K. Köhler. If more than one skeleton were buried in a pit, bones were separated professionally by the processing anthropologists. The samples were taken wearing gloves, facemask, under possibly clean conditions, with bleached materials. The samples were then directly transferred to the laboratory in Mainz, and were stored at -20 °C.

4 Methods

All ancient DNA analyses were performed in the aDNA facilities of the Anthropological Institute of Johannes Gutenberg University in Mainz. Two separate laboratories were used for the two steps of the aDNA work: a pre-PCR area for sample preparation, milling, DNA extraction, and set up of PCRs, and a post-PCR laboratory in a different building used for gel electrophoresis, sequencing, and cloning.

4.1.1 Sample preparation

The tooth and bone samples were irradiated with UV-C (260nm) light for 35 minutes on two sides. Subsequently, the entire surface of the bone pieces and teeth were removed with a spot blasting unit P-G 400 (Harnisch & Rieth, Winterbach, Germany) with aluminium oxide abrasive. It was followed by a second UV-irradiation for 35 minutes on both sides. The decontaminated samples were ground to fine powder, using a mixer mill (Retsch, Haan, Germany) and stored at 4°C until use. To avoid cross contamination of the samples, the grinding jars were cleaned with bleach, DNA free water, and silicon dioxide (Roth, Karlsruhe, Germany) after each sample. The sterile box was also cleaned extensively with soap and bleach between the samples. Every tenth sample was a grinding blank, milling DNA free Hydroxylapatite (Roth, Karlsruhe, Germany) powder. This control was processed as a normal sample during the following extraction and DNA amplification steps.

4.1.2 Ancient DNA extraction

The DNA extraction process followed the standard protocol of our team (Haak et al., 2005; Brandt et al., 2013). On the first day of a DNA extraction process, 0.2-1 g bone or tooth powder were aliquoted in an end volume of 3.33 ml, with 1,500 mol EDTA (pH 8.0, Applied Biosystems/Ambion, Darmstadt, Germany), 15 mg N-laurylsarcosine (Merck, Darmstadt, Germany, solubilized previously) and 600 mg Proteinase K (Roche, Mannheim, Germany). The solution was mixed and incubated at 37°C on a rotary mixer for 24 hours.

DNA was extracted on the next day with 3 ml phenol/chloroform/isoamyl alcohol (25:24:1, ph 8.0, Fisher scientific) added to the solution, and centrifuged for 10 min at 4,000 rpm. The lower aqueous phase was transferred to a fresh tube. The bone powder and the organic phase were discarded. The solution was washed once more with phenol/chloroform/isoamyl alcohol. In the third washing step, trichloromethane/chloroform (Roth, Germany) was used to clean the solution from phenol. The supernatant was transferred to Amicon Ultra-15 filter units with 50

kDa NMWL (Millipore Billerica, USA) filters and the solution was concentrated and washed with 4-10 ml UV-C irradiated HPLC water (centrifuging three times on an average at 6,590 rpm for 5-7 min) until the extract was clear. 120-200 µl DNA extracts were aliquoted in DNA free HPLC water and stored at -20°C until amplification. 8-22 samples were processed at once, with 1-2 extraction blanks and one grinding blank at each extraction event.

4.1.3 PCR

4.1.3.1 Amplification of the mitochondrial genome

The amplification of the HVS-I of the mtDNA control region was performed using three primer pair systems (Figure 9, Supplementary Table 26). In the case of well-preserved samples, the DNA was amplified in two overlapping fragments, in a range of np 16,046-16,401. In case of the samples with an average DNA preservation, four overlapping primer pairs were used (amplifying between np 15,997-16,409). Some samples, in order to complete the HVS-I sequence, had to be amplified with the six primer pairs system, covering a range of np 16,019-16,401. HVS-II was amplified only from those samples that showed the same HVS-I motive from the same archaeological site, in order to detect potential intra-site maternal kinship relations.

The amplification reaction for HVS was set up in a volume of 25 µL. It contained 1x PCR Gold Buffer (Applied Biosystems, Darmstadt, Germany), 2.5 mM $MgCl_2$ solution (Applied Biosystems), 0.2 mM dNTP Mix (Qiagen, Hilden, Germany), 2.5 U AmpliTaq Gold® DNA Polymerase (5U/ µl Applied Biosystems), 0.2 µM each primer, 10 µg BSA (20 mg/ml Roche, Mannheim, Germany) and 1-5 µl DNA from bone or teeth extracts. The HPLC water used in the reaction was UV-C irradiated, and were tested for possible DNA contamination along with the primers. At each PCR, two to five reaction blanks were used, depending on the number of the amplified samples. Different samples (A, B) from the same skeleton were always amplified separately.

The amplifications were carried out in a Mastercycler gradient thermocycler (Eppendorf, Hamburg, Germany). The cycle conditions consisted of an initial denaturation at 95°C for 6 min, 45-50 cycles of 15 sec at 95°C, 15-30 sec at 50-67°C and 20-40 sec at 72°C, followed by a final extension at 72°C for 10 min. The presence of PCR products of expected size was visualized on a 2% ultrapure agarose (Invitrogen, Karlsruhe, Germany) electrophoresis gel, using GeneRuler allelic ladder (Thermo Scientific).

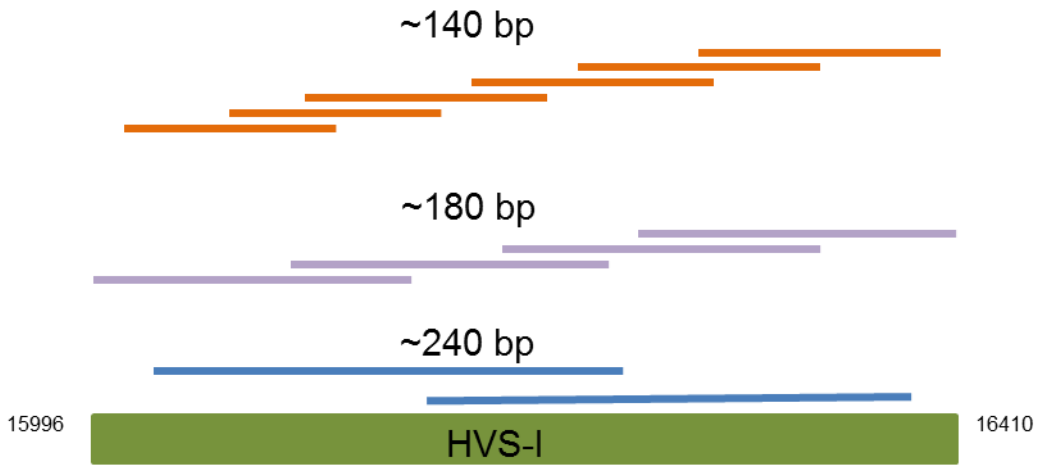


Figure 9. The three different amplification strategy for the studied HVS-I (np 15,996-16,410) of the mitochondrial genome.

Amplicons were purified with MultiScreen94 PCR Plates (Millipore Billerica, USA) by eluting the PCR product in 180 µl of HPLC-water (Applied Biosystems), transferring the mix to the MultiScreen94 PCR Plates, pumping for 15 minutes, washing with 25 µl of HPLC-water, pumping for 5 min, and resuspending the PCR product in 18 µl of HPLC-water.

Additionally, 22 phylogenetically significant coding region SNPs of the mitochondrial genome were analysed using the previously published GenoCoRe22 SNP multiplex assay (Haak et al., 2010, Supplementary Table 4). The PCS setup contained in a final volume of 25 µl (adjusted by HPLC water), 1x PCR Gold Buffer (Applied Biosystems), 6.5 mM $MgCl_2$ solution (Applied Biosystems), 0.6 mM dNTP Mix (Qiagen, Hilden, Germany). Furthermore, 1.25 U AmpliTaq Gold® DNA Polymerase (5U/ µl Applied Biosystems), 0.01-0.06 µM each primer, 20 µg BSA (20 mg/ml Roche, Mannheim, Germany) and 2-4 µl DNA were added to the reaction compound. The PCR cycling conditions were the followings: 95 °C, 10 min polymerase activation time; then 35 cycles (95 °C for 30 sec; 60 °C for 45 sec; 65 °C for 30 sec.) with a final elongation at 65 °C for 6 min.

The amplicons were purified by EcoI-SAP purification. 2.5 µl PCR product was mixed with 1 U of FastAP (Thermosensitive Alkaline Phosphatase), 0.4 U of ExoI enzyme (both from Thermo Fischer Scientific), and incubating at 37°C for 10 min, followed by heat inactivation at 75°C for 5 min.

4.1.3.2 Amplification of the Y-chromosome

25 phylogenetically informative SNPs were typed on the Y-chromosome, using the GenoY25 SNP multiplex assay (Haak et al., 2010) with some smaller modifications. Multiplex PCR was conducted in 16 µl volumes using 1× Buffer Gold, 8 mM $MgCl_2$ (Applied Biosystems), 0.7 mM dNTPs (Qiagen), 0.01-0.2 µM of each primer 13.4 µg BSA (Roche), 1.25 U of Amplitaq Gold Polymerase (Applied Biosystems), and 4 µl of DNA extract. Thermocycling conditions consisted of initial enzyme activation at 95°C for 6 min, followed by 37–45 cycles of denaturation at 95°C for 10 sec, annealing at 60°C for 1 min, and elongation at 65°C for 30 sec, followed by a final extension at 65°C for 6 min.

PCRs were visually checked by electrophoresis in 3% ultrapure agarose (Invitrogen) gels. PCR products were purified by mixing 2.5 µl of PCR product with 1.5 U of FastAP and 0.6 U of ExoI, 1 µl 10x FastAP buffer (all chemicals ordered from Fermentas, Thermo Scientific) and 1 µl HPLC water, and incubating at 37°C for 10 min, followed by heat inactivation at 75°C for 5 min.

After the initial multiplex screening of all the samples, I designed several singleplex PCRs for further typing of the I, G and F* haplogroups (see Table 2 for primer details). The PCR conditions were the same as at the mitochondrial DNA amplification. The PCR products, after detecting in 2% agarose gel, were purified by Exo I-FastAP enzymes, as it was written earlier. The CycleSeq reaction and the clean-up of the sequencing products were also the same as at the mitochondrial DNA analysis.

Y haplogroup	Position	Primer name	Sequence 5'-3'	Product Size (bp)	Annea-ling (C°)	Reference
I1	M253	M253_L	TGCTCAGCTAACTAGTCCTGT	75	52°	this study
		M253_H	CATTCAATGAAGAACCTGGAGA			this study
I2	M438(=P215)	M438.1_L	GGGCCTGGAATGTAGACTAATGGT	62	58°	this study
		M438.1_H	AAGCCAAATTCACAAATTAAAAACAAAT			this study
I2	M438	M438.2_L	AGTTTTGGGCCTGGAATGTAGA	61	55°	this study
		M438.2_H	ATTCACAAATTAAAAACAAATCTGTATGC			this study
I2a1 (I2a formerly)	P37.2	P37.2_L	ATCTCCTGAGACAAGCATAGTGATAGGGT	83	60°	this study
		P37.2_H	ATGGTTGTGGGGGGCCTTTAAAT			this study
I2a2 (I2a2a, I2b formerly)	M223	M223_L	AGACTCTGTGTTACTAGCTGAAGATG	68	54°	this study
		M223_H	TGCACATTGATAAATTTACTTACAGT			this study
G2	P287	P287_L	CTCTGGAACCTCCTGACTGACAT	63	54°	this study
		P287_H	GCTAAAGCCACTGGCACTGAA			this study
G1	M285	M285_L	CATTTCTCATCATCTACATTTCTCCT	84	52°	this study
		M285_H	TCGAATCCGCTATCCAGACT			Haak et al. 2010
G2a	P15	P15.2_L	TCCTCACATGAATAGAGCCAATGCT	72	59°	this study
		P15.2_H	ACTTTTGCAACTTTCATCTGCCTTCA			this study
G2a2b	S126/L30	S126_L	TCAGAGCCCTTAGTAGCTCATTTTA	75	54°	this study
		S126_H	TAGGCAAGACCATGTAGAGTATAGAATG			this study
H	M69	M69_L	TTCAGGAGGCTGTTTACACTC	63	52°	this study
		M69_H	ATCTTTATTCCCTTTGTCTTGC			this study

Table 2. List of primers, used for Y-chromosomal singleplex reactions.

Primer's names indicate forward [L-strand (L)] and reverse primer orientation [H-strand (H)]. Primers denote the typed position at the case of the Y-chromosome primers.

4.1.4 Cloning and sequencing

During the cloning and sequencing we (our lab assistant S. Möller-Rieker and myself) followed our standard protocols as described previously (Haak et al., 2005). Ligation was carried out with 4 U T4 Ligase (Fermentas, Thermo Scientific), 1x Ligation Buffer (Fermentas, Thermo Scientific) and 50 ng T-Vector (self-production of S. Möller-Rieker) and 4 µl purified PCR product. The T vector was made of 73.5 µl 7 µg/ml pUC 18 vector cut by SmaI restriction enzyme, 10 µl of 10x PCR buffer, 10 µl $MgCl_2$, 1.5 µl of 100 mM dTTP, 5 µl (5U/ µl) Taq polymerase (all ordered from Thermo Scientific).

Ligation mix was incubated at 16°C overnight. On the second day, the samples were shaken for one hour at 450 rpm. The T4 ligase was inactivated at 70°C for 10 min. This step was followed by a chloroform extraction: 10 µl water and 10 µl chloroform was pipetted to the ligated samples and after vortexing it was centrifuged for 10 min at 12,000 rpm. After discarding the lower organic phase, the DNA was precipitated from the chloroform by adding 2 µl of 3 M sodium acetate (pH 4.6) and 50 µl of 100% ethanol. After 10 min incubation, we centrifuged the solution for 30 min at 12,000 rpm. We discarded the supernatant, and the resulting pellet was washed once again with 250 µl 70% ethanol, and centrifuged for 15 min at 10,000 rpm and then dried for 15 min at 36 °C. The purified plasmid was re-suspended in 10 µl HPLC water.

The plasmid DNA was transformed to electrocompetent E. Coli cells. The plasmids were mixed with 50 µl E. Coli cells, and an electroporation was performed by an electroporation system (Easyject Prima 2,500V by Equibio, Ashford, UK) with electroporation cuvettes (peQLab Biotechnologie, Erlangen, Germany). The shocked cells were suspended in sterile 1 ml LB medium (5g tryptone, 2.5g yeast extract, 2.5 g sodium chloride in 500 ml of HPLC water) and 30-100 µl was plated out. The agar plates were cooked from 10 g tryptone, 5 g yeast 5 g sodium chloride, and 15 g agar-agar in 1,000 ml HPLC water. Additionally 100 mg ampicillin, and 47.6 mg IPTG and 100mg X-Gal were added to the solution. Bacteria that did not contain any plasmid (with ampicillin resistance gene), could not grow on this medium.

The clones were selected by blue-white selection after a day of incubation at 37 °C on the agar-agar plates. The white colonies contained the plasmid with the inserted aDNA fragment. In these bacteria, the β-galactosidase remained inactive, since the inserted aDNA fragment interrupted the transcription of the LacZ gene and so no functional β-galactosidase could be formed. The non-recombinant plasmids (which contain only the vector) grow into blue colonies.

Clone PCR products were purified by adding 0.33 U of FastAP and 2 U of Exo I (MBI Fermentas, Thermo Scientific) directly to the PCR product, incubating at 37°C for 10 min and inactivation at 75°C for 5 min.

The PCR from the colonies was set up in a volume of 50 µl, with 10x PCR Buffer, 2.5 mM $MgCl_2$, 2 mM dNTP, 0.2 µM M13 forward and reverse primer, 1U Taq DNA polymerase (all from Fermentas, Thermo Scientific).

Cycle sequencing was set up in 10 µl reaction volume, including 0.5-1 µl Big Dye Terminator v. 1.1 (Applied Biosystems), 1.5 µl 5x Big Dye Terminator Sequencing Buffer (Applied Biosystems), 1 µM primers (M13 universal for the clones or the adequate reverse of forward primer for the direct sequencing), HPLC water and 1-4 µl PCR product. The thermocycling conditions were the followings: 25 cycles of denaturation on 93° for 15 sec, annealing on 56-58° for 15 sec, elongation in 60° for 2.30 min.

Sequencing products were digested with 1U FastAP (Fermentas), and then purified using the MultiScreen$_{384}$ SEQ filter Plates (Millipore Billerica, USA). We eluted the products in 50 µl of HPLC-water, transferred the mix to the MultiScreen$_{384}$ SEQ Plates, vacuum for 15 min (until the wells were empty), and resuspended the DNA in 20 µl Hi-Di formamide (Applied Biosystems). Samples were run on an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) using POP-6 (Applied Biosystems). The electropherograms were first visually checked and edited in Sequencing Analysis Software (Applied Biosystems), then the sequences were cut and edited in SeqMan software, using the IUB (International Union of Biochemistry) codes for the ambiguous or double nucleotides. Afterwards the sequences were assembled and aligned to the rCRS (Anderson et al., 1981; Andrews et al., 1999) in Megalign software (DNASTAR Lasergene software package).

4.1.5 SNaPshot typing

The single base extension (SBE) reaction of the GenoCorRe22 assay (Haak et al., 2010) was performed with the published SBE primers and the SNaPshot Multiplex Kit of the Applied Biosystems. 2.5 µl Ready Reaction Mix, 0.5 µl SBE primer mix, 1 µl HPLC water and 1 µl PCR product was mixed together. The cycling conditions were the following: 35 cycles of denaturation at 96 °C for 10 sec, annealing at 55 °C for 5 sec, elongation at 60 °C for 30 sec.

The SBE reaction of the GenoY25 multiplex assay was performed as published before (Haak et al., 2010). The reaction mix was composed of 4 µl Ready reaction mix (SNaPshot Multiplex Kit

by Applied Biosystems), 2 µl HPLC water, 1 µl primer mix and 1 µl purified PCR product. The cycling conditions were similar, but with fewer (30) cycles: denaturation at 96 °C for 10 sec; annealing at 55 °C for 5 sec; elongation at 60 °C, for 30 sec. The GenoCoRe22 and GenoY25 amplicons were purified with 1U FastAP (Fermentas) with the same cycling conditions as the PCR products.

For the electrophoresis, 2 µl SBE product was mixed with 11.8 µl formamide + 0.2 µl Liz 120 size standard (Applied Biosystems). Samples were run on an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) using POP-6 (Applied Biosystems). Evaluation and analysis of SNaPshot typing profiles were performed within the GeneMapper version 3.2 Software (Applied Biosystems), using custom settings.

The mitochondrial haplogroups were determined with the HaploGrep software (Kloss-Brandstätter et al., 2011) checking the results manually on the mitochondrial phylotree version 14 (van Oven and Kayser, 2009). The Y-chromosomal haplogroups were defined based on the subsequent SNP typing and the terminal SNP. Therefore I used the definitions of the ISOGG (International Society of Genetic Genealogy) Y-DNA haplogroup tree (version 9, 2014, http://www.isogg.org/tree) and the latest published update of the Y-chromosome phylogenetic tree (Karafet et al., 2008).

4.2 Authentication criteria

During my aDNA analyses, I took the following precautions to prevent and identify modern DNA contamination:

1. All the samples were collected under clean conditions, partly directly after the excavation. If they were washed or touched by other experts before sampling, we took swab samples from their buccal mucosa, and compared their DNA profile with the haplotypes generated from the prehistoric samples (Supplementary Table 27).
2. We used a special clean room area at the Johannes Gutenberg University in Mainz for all the analytical steps prior to the DNA amplification. The pre- and post-PCR areas were in separated buildings. The workflow and material transport was only in post-PCR direction from pre-PCR, there were strict precautions against carry over contamination from the post-PCR area. The pre-PCR laboratory was used only for ancient DNA analysis, no modern DNA work

was carried out in it. The amplifications, cloning, and sequencing were performed in the post-PCR lab.

3. In the clean room area, all workers changed street clothes to UV-C irradiated overall with hood and overshoes. Facemask, cap, visor, and three pairs of gloves were worn during the routine work, following strict protocols. The different work steps were performed in separate rooms, and they were all carried out in special clean boxes. The lab and the work surfaces and equipment were regularly cleaned with filtered, UV-C irradiated soapy water, and bleach (2.8%) or DNA-ExitusPlus™ (Applichem, Darmstadt, Germany). All items entering the lab were decontaminated with soapy water, bleach, and UV-C exposure.

4. All chemicals used in the pre-PCR phase were tested with blank and positive reactions for DNA contamination before using. Grinding, extraction and PCR were all monitored by blank controls, which were all amplified and at case of contamination sequenced. The gained positive blanks were checked for origin of contamination and compared with the samples.

5. The samples were collected by V. Keerl, J. János, M. Fecher, B. G. Mende and myself. All samples were processed exclusively by myself (mtDNA haplogroup T1a) in the pre-PCR steps. The mitochondrial haplotypes of all co-workers who possibly had contact with the samples are listed in Supplementary Table 27.

6. All mitochondrial HVS-I sequences were reproduced by at least three different PCRs per fragment from at least two extracts from anatomically different bones or teeth (12 independent reactions at the default system of four overlapping primers). HVS-II sequences were amplified at least twice from the two extracts, in four overlapping fragments. The coding region multiplex PCR (GenoCore22) was performed once per DNA extract to verify the haplogroup assignments based on the HVS-I region. The Y25 multiplex PCR was carried out at least four times, twice from each extract. Heterogeneous HVS-I, and HVS-II sequences were cloned at an average of five clones per amplicon. Poorly preserved samples were cloned entirely. Inconsistent results were clarified by a third extraction from a third anatomically different sample.

4.3 Mitochondrial DNA population genetic analyses

4.3.1 Additional and comparative prehistoric mtDNA data

The comparative mitochondrial DNA data from literature was arranged into cultural units modelling prehistoric populations (Supplementary Table 6). 19 hunter-gatherers from Central/North Europe (HGCN) (Bollongino et al., 2013; Bramanti et al., 2009; Fu et al., 2013), 14 hunter-gatherers from Portugal and Spain (HGSW) (Chandler et al., 2005; Chandler, 2003; Hervella et al., 2012; Sánchez-Quinto et al., 2012) and 14 individuals from west Russia (HGE) (Bramanti et al., 2009; Der Sarkissian et al., 2013; Krause et al., 2010), were involved in the analyses (Figure 10).

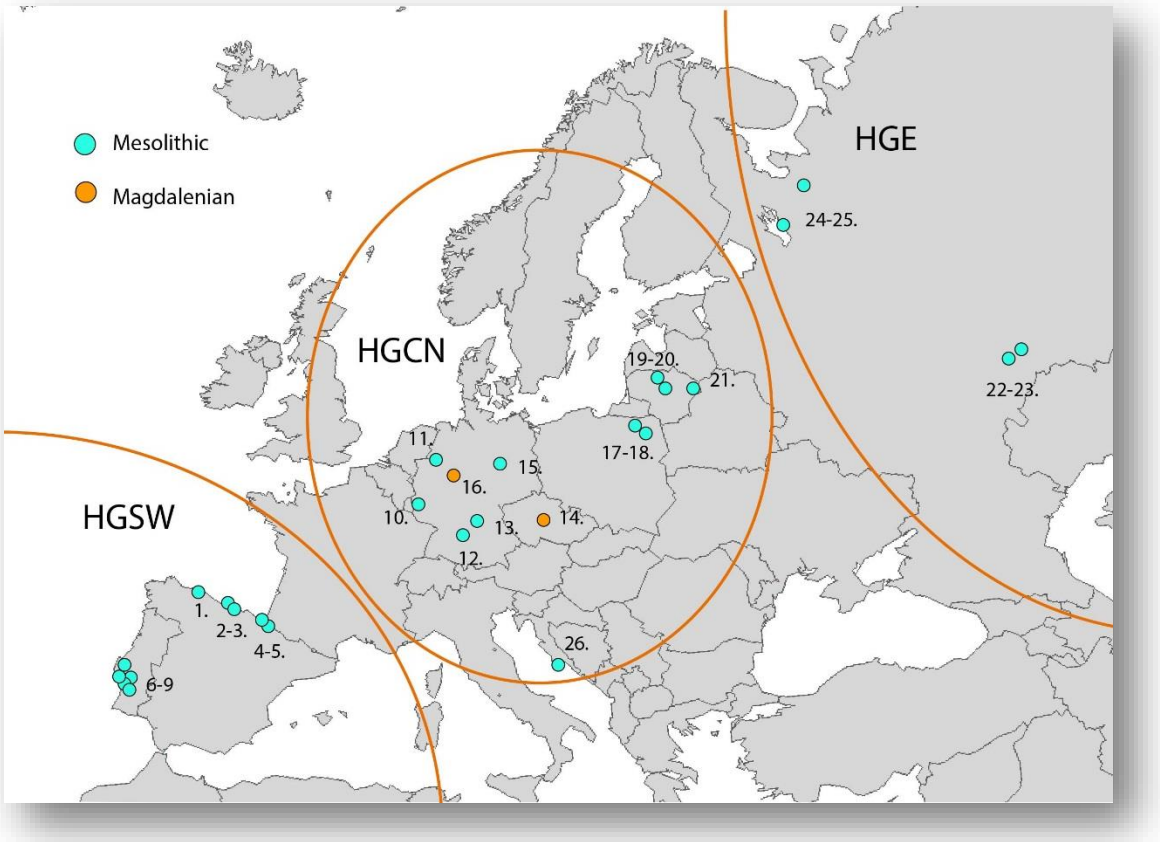


Figure 10. Distribution of the comparative European pre-Neolithic hunter-gatherer aDNA data.

Plotted sites: La Braña (1), La Pasiega (2), La Chora (3), Erralla (4), Aizpea (5), Toledo (6), Arapouco (7), Cabeço das Amoreiras (8), Cabeço de Pez (9), Reuland-Loschbour (10), Oberkassel (11), Hohler Fels (12), Hohlenstein-Stadel (13), Dolni Vestonice (14), Bad Dürrenberg (15), Blätterhöhle (16), Dudka (17), Drestwo (18), Spiginas (19), Donkalnis (20), Kretuonas (21), Lebyazhinka (22), Chekalino (23), Popovo (24), Yuzhnyy Oleni Ostrov (25), Vela Spila (26). Comparative hunter-gatherer groups: HGCN=Hunter-Gatherer Central/North, HGSW=Hunter-Gatherer Southwest, HGE=Hunter-Gatherer East.

In addition, a series of Neolithic populations were used from Central Europe as comparative data. These are the LBK sample set with 108 individuals (Bramanti, 2008; Brandt et al., 2013; Haak et al., 2010, 2005), 17 individuals from the Rössen culture (RSC) (Brandt et al., 2013; Lee et al., 2013), 33 individuals from the Schöningen group (SCG), 19 from the Baalberge culture (BAC), 29 from the Salzmünde culture (SMC), 17 from the Bernburg culture (BEC), 36 from the Corded Ware culture (CWC) (Haak et al., 2008; Brandt et al., 2013), 35 individuals form the Bell Beaker culture (BBC) (Brandt et al., 2013; Lee et al., 2012) and 94 from the Únětice culture (UC) (Brandt et al., 2013).

Furthermore, 18 Neolithic individual from North Portugal (NPO) (Chandler et al., 2005, 2003), 20 from Cardial and Epicardial culture in Spain (CAR) (Gamba et al., 2012; Lacan et al., 2011b), 43 from the Neolithic period of Basque Country, Navarre and Cantabria (NBQ) (Hervella et al., 2012) were involved the analyses. Additionally, 29 individuals from the Treilles cultural group in France (TRE) (Lacan et al., 2011a), 19 individuals from the Pitted Ware culture (PWC) from Gotland, south Sweden (Malmström et al., 2009), 10 individuals from the Funnel Beaker culture (FBC) from Frälsegården, south Sweden and Ostdorf, Germany (Malmström et al., 2009; Bramanti et al., 2009) were involved In the extended population genetic analyses. Moreover, 11 Individuals from Siberian Bronze Age (BAS) (Keyser et al., 2009), eight individuals from Bronze Age Kazakhstan (BAK) (Lalueza-Fox et al., 2004), nine individuals from Eneolithic on the North Pontic-Caspian Steppe (ENL), 26 individuals from Yamnaya culture (YAM), 24 individuals from Catacomb culture (CAT) (Wilde et al., 2014) were used in the population genetic studies.

Due to the small sample sizes and too short HVS-I sequences, the 11 individuals from Camıi de Can Grau, northeast Spain (Sampietro et al., 2007), six samples from the Neolithic Ukraine (Nikitin et al., 2012) were not used. The large time span between the hunter-gather data made it reasonable to omit the Gravettian aDNA data from the analyses (Bramanti et al., 2009; Caramelli et al., 2008; Fu et al., 2013; Krause et al., 2010). The study of Guba et al., 2011 was omitted from the analyses because of misdating most of the Neolithic samples (Bánffy et al., 2012).

4.3.2 Data structuring

I used several combinations of the above listed cultures or groups for the population genetic analyses, which are proxies of different prehistoric populations or metapopulations. These variations of grouping were necessary for the comprehensive analyses of the

Transdanubian cultures' genetic affinities. Two types of datasets were differentiated in the cases of the "Regions Transdanubia." and "Carpathian-Basin". The difference is the inclusion of the HGCN group, which inhibited the evaluation of the multidimensional scaling and molecular variance analyses, though was a very important element of the principal component analysis.

Altogether 29 cultures/populations from the pre-Neolithic times to the Bronze Age on a scale of Europe and West/Central Asia were used for the population genetic analyses in the following combinations. Assemblages of 29 cultures (29POP) for the Eurasian scale comparison, 20 cultures (20POP), 16 cultures (16POP), 15 cultures (15POP) for the European scale comparison, and 8-7 cultures (8-7POP) assemblages were used for the comparisons inside Transdanubia (Table 3).

The 20POP and Carpathian Basin-1 prehistoric datasets were classified in two ways: (1) The complete dataset of each culture/population consists of all individuals with reproduced HVS-I sequences. (2) Datasets signed by "*" after the culture/population abbreviation are based on the first datasets but with a potential kinship taken into account. Maternal lineages with identical HVS-I, and if provided, identical HVS-II sequences from the same archaeological site were counted only once.

The Transdanubian genetic results were also compared to modern mtDNA genetic data. In order to identify affinities of our prehistoric sample sets in the maternal gene pool of present-day Eurasian and African populations, our team gathered 67,996 mitochondrial HVS-I sequences from the literature (Brandt et al., 2013). Different mtDNA datasets were generated, which were used for principal component analysis (PCA) and genetic distance maps. The modern-day data were pooled into different populations according to geography or ethnicity, as described in the original publications. For PCA with mtDNA data, the present-day samples were grouped into 73 populations. This dataset was composed of 50,688 sequences with an average sample size of 694 samples per population (Brandt et al., 2013). All sequence data were ordered to haplogroups by using phylogeny of phylotree.com (van Oven & Kayser, 2009, www.phylotree.com, built 11, accessed on 07 February 2010).

List of the involved cultures/populations in a specific arrangement	Name of the arrangement	Haplogroup frequencies	Molecular Diversity Indices	MDS	PCA	Ward clustering	Fst analysis	ASHA	TPC	AMOVA
HGCN, HGSW, HGE, STA, LBKT, VIN, SOP, LGY, BL, LBK, RSC, SCG, BAC, SMC, BEC, CWC, BBC, UC, CAR, NPO, NBQ, TRE, PWC, FBC, BAK, BAS, ENL, YAM, CAT	29POP			C	C	C	C			
HGCN, HGSW, STA, LBKT, VIN, SOP, LGY, BL, LBK, RSC, SCG, BAC, SMC, BEC, CWC, BBC, UC, CAR, NPO, NBQ	20POP			C + *	C + *					
HGCN, STA, LBKT, VIN, SOP, LGY, BL, LBK, RSC, SCG, BAC, SMC, BEC, CWC, BBC, UC	16POP							C		
STA, LBKT, VIN, SOP, LGY, BL, LBK, RSC, SCG, BAC, SMC, BEC, CWC, BBC, UC	15POP									C
HGCN, STA, LBKT, VIN, SOP, LGY, BL, LBK	8POP	C	C				C		C	
HGCN, STA, LBKT, VIN, SOP, LGY, BL,	7POP_1							C		
STA, LBKT, VIN, SOP, LGY, BL, LBK	7POP_2									C
STA, LBKTsou, LBKTno, VIN, SOPsou, SOPno LGYsou, LGYno, BL +HGCN	Reg. Transd. _1				C + *					
STA, LBKTsou, LBKTno, VIN, SOPsou, SOPno LGYsou, LGYno, BL	Reg. Transd. _2			C + *						C
STA, LBKTsou, LBKTno, VIN, SOP, LGYsou, LGYno, BL, Körös, ALBK-Szatmár, ALBK-Tiszadob/Bükk, ALBK-Esztár, ALBK-Szakálhát, Tisza +HGCN	Carpathian-Basin _1				C					
STA, LBKTsou, LBKTno, VIN, SOPsou, SOPno LGYsou, LGYno, BL, Körös, ALBK-Szatmár, ALBK-Tiszadob/Bükk, ALBK-Esztár, ALBK-Szakálhát, Tisza	Carpathian-Basin _2			C						C

Table 3. Assemblages of the prehistoric datasets.

Letter "C" indicates the set of complete datasets for the given analysis. Sign * indicates that reduced datasets (without possible maternal related individuals) were also involved into a particular analysis. Abbreviations: MDS=multidimensional scaling, PCA=principal component analysis, ASHA=ancestral shared haplotypes analysis, TPC=test of population continuity, AMOVA=analysis of molecular variance. Culture/population abbreviations are resolved in the text and in Supplementary Table 6.

Mitochondrial genetic distance maps (GDM) were generated from HVS-I sequences of 130 modern-day populations. Whenever possible, the administrative subdivisions of a country were considered in order to increase the phylogeographic resolution. In this dataset, population data with a minimum sequence range of np 16,068-16,365 were included, in order to exclude biases of varying sequence ranges. Each population is represented by a maximum of 140 randomly selected individuals, which resulted in a total amount of 17,074 sequences used in the analysis.

4.3.3 Molecular diversity indices

I calculated molecular diversity indices for HVS-I sequence in the range of np 16,056-16,400 for the HGCN, STA, LBKT, LBK, VIN, SOP, LGY and BL datasets (Table 6). For the calculation, the function "Gene flow and genetic differentiation" in the program "DNA Sequence Polymorphism" or DnaSP, version 5.10.01 was used (Librado and Rozas, 2009). DnaSP estimated the following measures: haplotype diversity (Hd) (Nei, 1987, equation 8.4), average number of nucleotide differences (K) (Tajima, 1983, equation A3), nucleotide diversity (π, Pi) (Nei and Li, 1979). Tajima's D was calculated in Arlequin 3.5.1 software (Excoffier and Lischer, 2010). Haplotype definition was inferred from distance matrix.

4.3.4 Genetic distances

F_{st} values were computed in Arlequin 3.5.1 (Excoffier and Lischer, 2010) based on HVS-I sequences (np 16,056-16,400) of the 29POP and 8POP datasets. The Tamura & Nei substitution model (Tamura and Nei, 1993) was used and an associated gamma value of 0.325, inferred from the software FindModel based on PAML likelihoods. (www.hcv.lanl.gov/content/sequence/findmodel/findmodel.html) and tested significant variations in F_{st}-values by 10,000 permutations. The p values were adjusted post hoc to correct for multiple comparisons with the Benjamin and Hochberg method, using the function p.adjust in R. 3.0.2.

4.3.5 Test of population continuity

I performed a tests of population continuity (TPC) as described in Brandt et al. (Brandt et al., 2013) using the absolute haplogroup frequencies of the hunter-gatherers from Central/North Europe, STA, VIN, LBKT (n=39, without Nitra site), LBK in Central Europe, SOP, LGY, and BL datasets. In order to apply conservative parameters (i.e. maximizing the chances of genetic drift), I used the terminal dates of the Mesolithic in the Carpathian Basin (6,000 cal

BC) and of each Neolithic culture's timespan. These were the following dates: 5,400 cal BC for STA, 5,000 cal BC for VIN, 4,950 cal BC for LBKT, 4,800 cal BC for LBK in Central Europe, 4,800 cal BC for SOP, 4,300 cal BC for LGY, 3,900 cal BC for BL. The difference in time between populations was defined in n generations of 25 years. I also ran each of the pairwise tests of all possible group combinations with three different effective population sizes (Ne=500, 5,000 and 30,000) (Table 7). The TPC script is available at https://github.com/joepickrell/tpc.

4.3.6 Principal component analysis

I performed PCA analyses, comparing the Transdanubian results with prehistoric and modern mtDNA data. MtDNA haplogroups were condensed into 25 haplogroups at the case of the Eurasian-scale (29POP) comparison. These haplogroups were the followings: C, D, Z, H, HV, V, I, J, K, N, N1a, R, T, T1, T2, U, U2, U3, U4, U5, U5a, U5b, U8, W, and X. I used the following sets of 23 (sub-) haplogroups for the comparison of the 20POP sets of data: H, H5, HV, HV0, V, I, J, K, N, N1a, R, T1, T2, U, U2, U3, U4, U5, U5a, U5b, U8, W, and X. Each of these haplogroups were observed at least in one of the ancient datasets. At the 20POP level of the analyses, I tested the affinities with two datasets per culture/population. Besides complete dataset, I used the reduced (*) dataset as well (see chapter 4.3.2). The Transdanubian PCA was done with 20 haplogroups: H, H5, HV, HV0, V, J, K, N1a, T1, T2, U, U2, U3, U4, U5, U5a, U5b, U8, W, X. The Carpathian-Basin-scale PCAs contained two more haplogroups, R and N1b, because of their occurrence in the Alföld (eastern Hungary).

The variable correlations with the PCs were displayed in the PAST program (PAleontological STatistics), in version 2.15 by Øyvind Hammer (Supplementary Figure 1).

The comparative PCAs between prehistoric cultures/populations and modern populations were carried out by using the mtDNA haplogroup frequencies of the whole dataset of each Transdanubian culture and 73 modern populations. The applied 21 haplogroups cover the most frequent haplogroups of the six ancient datasets and the modern Eurasian populations: N1a, I, W, X, HV, HV0/V, H, T1, T2, J, U, U2, U3, U4, U5a, U5b, U8, K, African haplogroups (L), Asian haplogroups (A, B, C, D, E, F, G, Q, Y, and Z) and other (all remaining haplogroups).

All PCAs were carried out using the *prcomp* function for categorical PCA implemented in the R 2.13.1 package (R Core Team (2012), http://www.r-project.org).

4.3.7 Hierarchical clustering

Adjoining to certain PCAs, I performed hierarchical clustering analyses, testing several algorithms (Ward type (Ward, 1963), paired group, single linkage) and several similarity measurement methods (Manhattan, Euclidean, correlation). All PCs (all genetic variations) were used for the clustering. The Ward clustering results were visualized in R as a dendrogram by using the hclust function in R.2.13.1. Significance of each cluster was evaluated by 10,000 bootstrap replicates using the pvclust function in R.2.13.1. Cluster significance is given as AU (Approximately Unbiased) p-value, which is computed by multi-scale bootstrap resampling with 10,000 replicates.

The unweighted pair-group clustering method (UPGMA) was used for the hierarchical cluster analysis in PAST software (PAleontological STatistics) version 2.15, by Øyvind Hammer. At the unweighted pair-group average, the clusters are joined based on the average distance between all members in the two groups. Several distance calculations were tested, and the best choice, with high cophenetic coefficient was selected. Bootstrapping was performed 10,000 times, and the percentage of random replicates was given at the root.

4.3.8 Multidimensional scaling

Using the Multidimensional scaling (MDS) method, I compared the complete datasets in the 29POP, Regions Transdanubia_2, Carpathian Basin_2 combinations, and the two versions of datasets in the assembling of 20POP. The distance calculations were based on the HVS-I sequence range of np 16,056-16,400, and was performed in Arlequin software version 3.5.1 (Excoiffer et al., 2010). Pairwise F_{st} value and Slatkin's distance matrixes were computed. The evolution model of Tamura & Nei (Tamura and Nei, 1993) with a gamma value of 0.325 was chosen by FindModel program based on PAML likelihoods (http://hcv.lanl.gov/content/sequence/findmodel/findmodel.html). The F statistic was based on 10,000 permutations with a significance level of 0.05. The haplotype definition was inferred from the distance matrix. MDS was applied on the matrix of linearized Slatkin F_{st} values (Slatkin, 1995) and visualized in a two dimensional space using the *metaMDS* function based on Euclidean distances implemented in the *vegan* library of R 2.13.1 (R Core Team (2012), http://www.r-project.org/).

4.3.9 Analysis of molecular variance

Analysis of molecular variance (AMOVA), based on HVS-I sequences (np 16,056-16,400) was performed with the 15POP, 7POP_2, Regions Transdanubia_2, Carpathian Basin_2 datasets. Complete datasets were used in the analysis. Variances, fixation indexes, and significant values (p) were computed, using the standard AMOVA function implemented in Arlequin 3.5.1. F_{st} values were tested on significance by 10,000 permutations. The haplotype definition was inferred from the distance matrix. The cultures or groups were arranged into different two or three clusters (models) and AMOVA was conducted for each assembly. From several tested combinations, one best model was resulted, with the greatest "among-groups" and the least "within-groups" variances and F_{st} values (Supplementary Table 12, 16, 20).

4.3.10 Genetic distance mapping

The comparative modern datasets of the mtDNA distance mapping (GDM) was formed out of 130 present-day populations. Whenever possible, the administrative subdivisions of a country were considered. From these datasets, I chose randomly maximum 140 sequences per population (n=17,074 sequences altogether), in order to balance the differences in sample sizes. The sequence length was uniform, ranged between np 16,068-16,365. I have calculated the genetic distances of these modern populations from the six Transdanubian Neolithic cultures, using the complete aDNA datasets. The analysis was performed with the Arlequin program, using the Tamura and Nei model (Tamura and Nei, 1993), with a gamma value of 0.177. For the haplotype definition, the original definition was used. F_{st} values between the cultures and each modern population were combined with longitudes and latitudes (according to sampling information in literature). The F_{st} values and coordinates were interpolated with the Natural Neighbor method implemented in ArcGis version 10.0 (Arcmap, Environmental Systems Research Institute (Esri) Inc, Redlands, USA) (Supplementary table 23).

4.3.11 Ancient shared haplotype analysis

The ancient shared haplotype (ASHA) analysis with prehistoric comparative data consisted of a two-step principle as written in Szécsényi-Nagy et al., 2014a. The first step was a classical shared haplotype analysis among the whole datasets of 16POP and 7POP_1 (Supplementary Table 13a, 15a), and the second was a successive reduction of the shared haplotypes by taking the following hypothetical chronology into account: HGN, STA, LBKT, LBK,

SOP, LGY, RSC, BL, SCG, BAC, SMC, BEC, CWC, BBC, UC for the 16POP and HGCN, STA, VIN, LBKT, SOP, LGY, SOP, BL for the 7POP_1 (see Supplementary Table 6 for abbreviations). Each haplotype was traced back to its earliest appearance in the analysed cultures, and got his name from the culture where it appeared at first (Supplementary Table 13b, 15b).

Similar principle was followed in the ASHA comparing the Transdanubian haplotypes to modern population genetic data. In the first comparison, I followed the method published in Haak et al., 2010. The modern comparative datasets were the same as used for the previous GDM. I counted shared haplotypes between the STA, LBKT and 130 modern populations from Europe, Asia, and North Africa. The number of individuals sharing the same haplotypes was regarded, similarly to the previous ASHA. I classified the STA and LBKT haplotypes into three groups: (1) common haplotypes, (2) rare haplotypes that are informative, (3) unique prehistoric haplotypes. The percentage of the matches of the STA informative haplotypes was displayed on Supplementary Figure 4.

In the second comparison, I used Lombard mtDNA results from the Transdanubia (Alt et al., 2014), Hungarian Conquest period (9-10th century) data (Tömöry et al., 2007), Cumanian mtDNA haplotypes from the Alföld (Bogácsi-Szabó et al., 2005), and the two available modern Hungarian datasets (Irwin et al., 2007; Tömöry et al., 2007). An ancestral shared haplotype calculation was made with these datasets (Table 9).

4.4 Population genetic analyses with Y-chromosomal data

4.4.1 Comparative prehistoric Y-chromosomal data

The Y-chromosomal data were compared with the Neolithic and Mesolithic parallels (Haak et al., 2010; Lacan et al., 2011a, 2011b; Lazaridis et al., 2014; Olalde et al., 2014; Gamba et al., 2014). For population genetic analysis of NRY data, I combined my results with three published LBK data (Haak et al., 2010) to enlarge the prehistoric dataset up to twelve individuals.

I collated a binary marker (SNP) frequency database from literature data (Szécsényi-Nagy et al., 2014a). These 49,500 Y-chromosomal SNP data from Eurasia and Africa were the basis of the selections, used in PCA and GDM analyses. After a successive reduction of the datasets, following haplogroup resolution criteria (see further details in the specific method

section). I pooled the modern-day data into different populations, according to geography or ethnicity, as described in the original publications. Present-day population data were only considered when the Y-chromosome sub-haplogroups I, G, F, K, and R1a were differentiated, since these are the most frequent haplogroups in the prehistoric data. Pooling the R1b with the R1 group was necessary, due to the poor resolution of the R1b outside Western Europe. At some Near East populations O was not characterized, but the NO cluster was detected (El-Sibai et al., 2009; Zalloua et al., 2008), therefore I condensed the two sister-haplogroups in one cluster. The poor resolution of the I1-I2 and the G1-G2 subgroups in the majority of the literature did not enable any further cluster division.

PCA was carried out with 24,464 samples from 80 present-day populations with an average sample size of 305 individuals per population. The Y-chromosome genetic distance map consists of 100 modern-day populations with 215 samples per population on average, using of 21,478 individuals from my database (Supplementary Table 24).

4.4.2 Principal component analysis of the Y-chromosomal data

In case of the PCA with Early/Middle Neolithic NRY data, the reduction of the collected data yielded 80 modern Eurasian and North African populations with sufficient SNP resolution for the most informative shared derived markers. The haplogroups were condensed to 13 clusters (AB*, E, DHC*, F*, G, I, J, KTS*, L, NO*, N, PR*, and R1a), based on phylogeny and phylogeography of the subgroups.

When analysing the late Neolithic NRY data, 16 haplogroups/haplogroup clusters were differentiated: AB, DE*, E, F*, G, H, C, I, J, K*, L, NO*, N, PR*, R1, and R1a. The SOP and LGY datasets were pooled in this analysis, since they have very similar Y-chromosomal diversity. As comparison, NRY data of 79 modern populations were used for the analysis. The difference between the two PCAs was that Kuwait (KUW) was an additional population in the SOP-LGY PCA, and Oman and Saudi Arabia were merged into the group “Arabs” (ARA) in this second analysis (Supplementary Table 24).

The 3D scatterplots were performed with an R script (see an example in chapter 15.1.3), using the *prcomp* function for categorical PCA implemented in the R 2.13.1 package (R Core Team (2012), http://www.r-project.org/). The variable correlation with the PCs were displayed in PAST (PAleontological STatistics) version 2.15, by Øyvind Hammer (Supplementary Figure 5Supplementary Figure 7).

4.4.3 Genetic distance mapping of the Y-chromosomal data

Based on 16 NRY haplogroup clusters (AB, DE*, E*, F, G, H, C, I, J, KTS*, L, O, NO*, PR*, R1, and R1a), I calculated genetic differentiation of 100 modern populations from the combined STA-LBKT-LBK, and SOP-LGY datasets. Pairwise F_{st} values were computed in Arlequin 3.5.1 program (Excoffier and Lischer, 2010), using conventional F statistics, and the original definition as haplotype definition. The genetic distances among the cultures and each modern population were combined with longitudes and latitudes (according to sampling information in literature), and interpolated with the Natural Neighbor method implemented in ArcGis version 10.0 (Arcmap, Environmental Systems Research Institute [Esri] Inc, Redlands, USA) (Supplementary Table 25).

5 Results

5.1 Mitochondrial DNA results

5.1.1 HVS-I amplification and DNA contamination

The HVS-I region of the mtDNA was amplified following a hierarchical amplification strategy with each sample (see in chapter 4.1.3.1). All well-preserved samples (n=566, with a "good look" after surface removing) were amplified with the two-fragment primer pair system at first. If the DNA was too degraded for the amplification in longer segments, the amplification was successively followed by the four-fragment and the six-fragment systems (Figure 9, Supplementary Table 26).

The success rates of amplifications regarding the sample types versus DNA fragment lengths relations were summarised on Figure 11. The samples from the site Balatonszárszó (n=67) are not involved in this demonstration, since they were amplified in a pilot study, using different amplification strategy (only the four overlapping primer pairs were used). Using the two-fragment system, the permanent and deciduous teeth, and the petrous part of the temporal bones (pars petrosa ossis temporalis) had better amplification rates than the long bones (47-63% versus 12%).

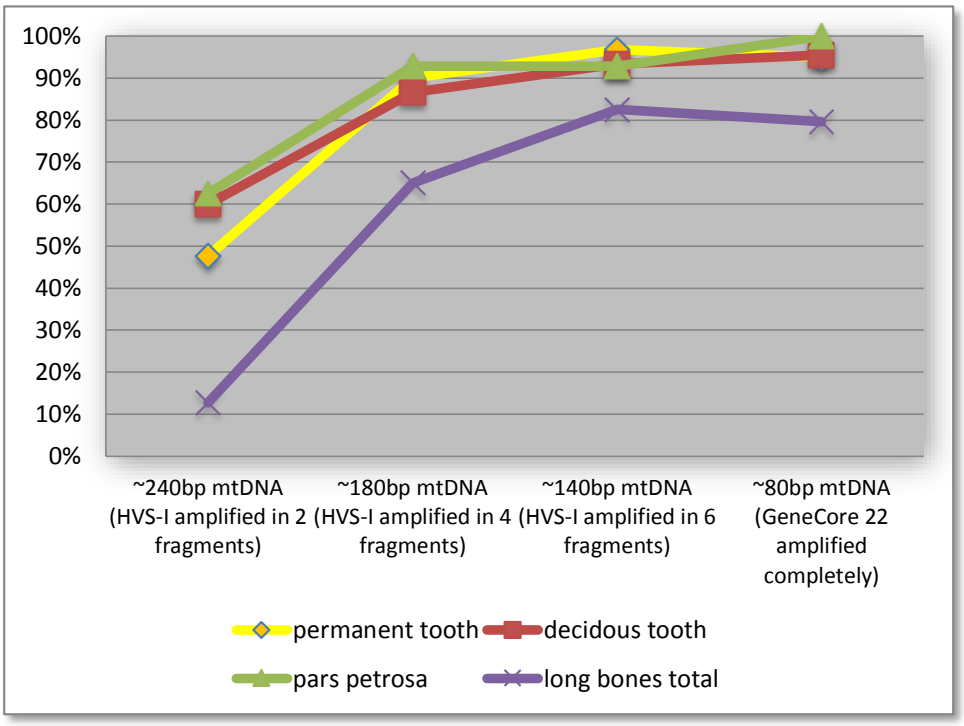


Figure 11. Success rate of amplification with different sample types, amplifying different lengths of fragments.

The four-fragment system was used for those samples, which were poorly preserved, or did fail in the two-fragment amplification (n=442 samples). The improvement of the amplification rate was significant, around 90% for the teeth and pars petrosa, and about 65% for the long bones. Completing the analyses with the six overlapping fragment amplification systems the success rates were 92-96% for the teeth and partes petrosa, and 82% for the long bones. The GenoCoRe22 multiplex SNP assay (Haak et al., 2010) was completely amplified (haplogroup was identifiable) in 95% of the teeth samples and 100% of the partes petrosa (n=14), and ~80% of the long bones.

Ancient DNA was much better preserved in teeth and petrous part of the temporal bones than in long bones. Furthermore, the shorter the fragment, the better the amplification it was, which is an excepted phenomenon in aDNA research (Pääbo et al., 2004), and it is even used as argument for authenticity of the detected ancient DNA (e.g. Der Sarkissian et al., 2013). On the other hand, there were no major amplification rate differences among different types of bones or teeth. The DNA preservation rate might altered by other factors (e.g. such as individual preservation).

Blank controls were used in grinding, extraction, and PCR setup phases of the pre-PCR laboratory work. Altogether, 3,044 blank controls were amplified during the two and a half year laboratory work (between 20.02.2010-05.09.2012). Out of the 3,044 blank controls, 132 contained amplifiable DNA, and 126 were sequenceable (4.14%). Each PCR setup involved at least two PCR blank controls, and additionally one blank control per further sets of five samples. Each extraction set (10-22 samples) contained one hydroxyl apatite control (grinding control) and at least one extraction reagent blank control. These controls were amplified at least once with the four-fragment system of the HVS-I. Altogether, 34 out of 372 grinding and extraction blank controls were positive for DNA (9.14%). There were three extractions with consistent contamination in one of the blank controls (extraction AE, AF, AT showing contaminations of different types of haplogroup H). In other cases, the blank controls were not amplifiable for the whole HVS-I sequence, these contaminations were sporadic and partial. The positive blank controls were compared to the samples from the same grinding day, extraction, or PCR reaction. In case of haplotype matches, the given analyses were repeated, or started with a new sample. In case of two or more positive blank controls in a PCR reaction, the whole PCR was repeated. I never found my HVS-I haplotype (16126C 16163G 16186T 16189C 16264T 16294T, T1a) in any of the blank controls.

With the aid and co-work of our assistant, S. Möller-Rieker, 607 PCR fragments were cloned through pUc18 plasmids into E. Coli cells, and regained by blue-white selection and colony PCRs. Four to six colony PCR products pro original PCR fragment were sequenced and assembled. The sequenced PCR products were aligned to the original direct sequences, and thus helped to assess the consensus sequence of the sample, and to separate possible contaminations from endogenous sequences.

5.1.2 Success rate of the mitochondrial DNA analyses

Considering all studied individuals (n=298), the HVS-I region was reproduced in 85.9% of the cases (Table 4). Furthermore, HVS-II regions were reproduced in 80 individuals, in order to detect potential intra site maternal kinship relations. A general haplogroup definition was obtained in 276 cases out of the 298 individuals (92.6%), detecting 21 control region SNPs and one deletion with the GenoCoRe22 multiplex PCR system (Haak et al., 2010).

Period	Culture	n processed individuals	n successfully typed for HVS-I	success rate
Mesolithic		1	1	100%
Early Neolithic	**Starčevo**	47	44	93,6%
Middle Neolithic	**Vinča**	36	31	86,1%
	LBK in Transdanubia	66	42	63,6%
Late Neolithic	**Sopot**	37	37	100,0%
	Lengyel	89	82	92,1%
Chalcolithic	**Balaton-Lasinja**	16	13	81,3%
Bronze Age		6	6	100,0%
Summa		**298**	**256**	**85,9%**

Table 4. Success rates for the HVS-I region analyses of the mitochondrial DNA.

5.1.3 Mitochondrial haplogroup compositions of the Transdanubian datasets

The published Upper Palaeolithic (Magdalenian) and Mesolithic hunter-gatherer mtDNA results from Central and North Europe (abbreviated as HGCN) built the base of the

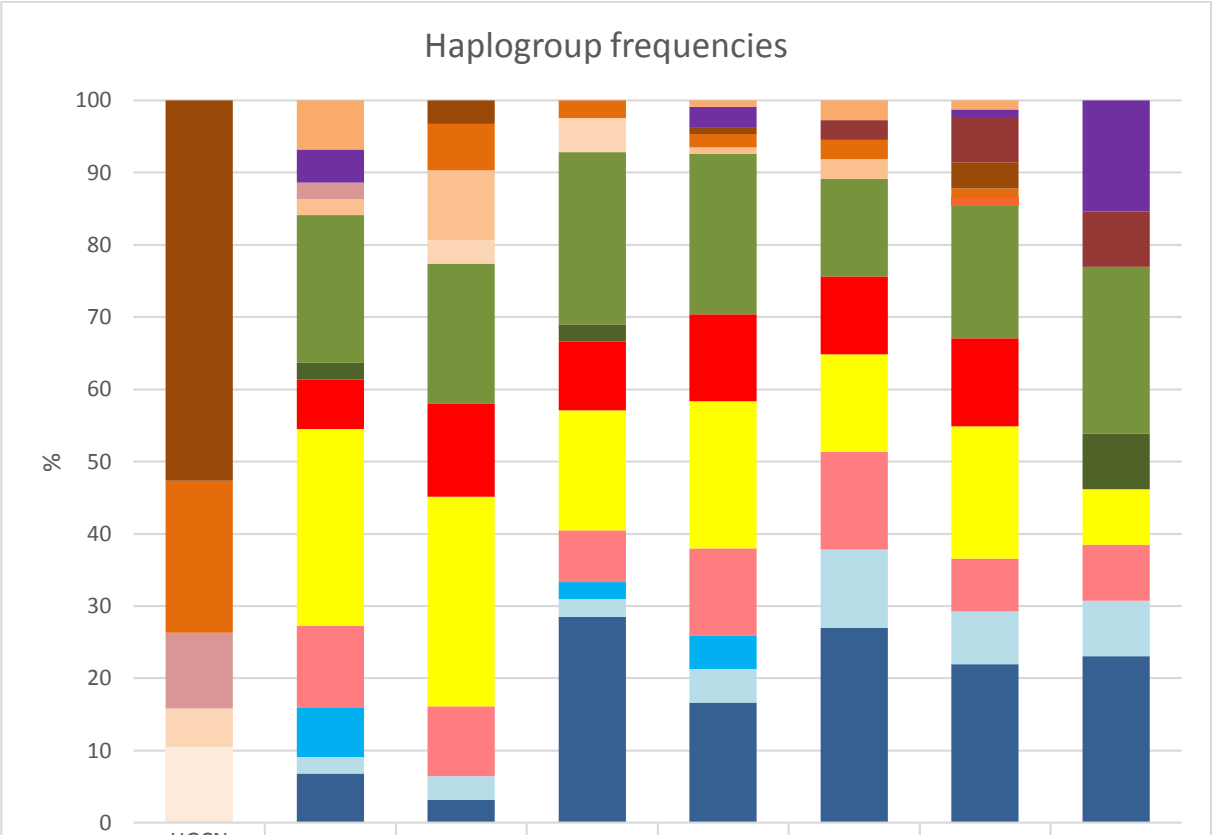
comparative studies. The dominant mitochondrial haplogroup of the pre-Neolithic hunter-gatherer dataset is the haplogroup U. Within the U cluster, U2, U4, U5a-U5b and U8 represent the hunter-gatherers' maternal gene pool (Bollongino et al., 2013; Bramanti et al., 2009; Fu et al., 2013).

I added one result of my own to the HGCN dataset, from a Mesolithic specimen (ca. 6,200-6,000 cal BC) found in Vela Spila cave on the Croatian Adriatic Korčula Island. The mtDNA type of this individual (sample "STANKO") belongs to the haplogroup U5b2a5, which represents the eighth haplogroup U5b among the 19 published HGCN data. This result further supports the theory of a homogeneous Mesolithic mtDNA substratum in Central/North Europe. These 19 hunter-gatherer data were projected to the Carpathian Basin, where the Mesolithic skeletal remains are exceptional scarce.

Comparing the aDNA results from the Carpathian Basin to the Central/North European hunter-gatherers, it is noticeable that the mitochondrial diversity rose significantly during the period of the earliest Neolithic Starčevo culture. A broad spectrum of new haplogroups appeared with the first farmers, such as K, J, T2, N1a, HV, and V, practically replacing the Central/North European Mesolithic substrate, which was supposed for the pre-Neolithic time of the Carpathian Basin too.

Besides typical Early Neolithic signature (N1a, J, K, and T2), the genetic data of the Vinča culture show different, potential hunter-gatherer haplogroups, such as U5a, U5b, and U2 (Table 5). The Early Neolithic impact either came from the southeastern core area of the Vinča culture (that was the same source region for the preceding Starčevo culture too), or these "Neolithic package" haplogroups originated from local elements, descended from the people of the Starčevo culture.

The people of the LBK in Transdanubia share a maternal genetic diversity with the Starčevo culture on the level of haplogroups T1, T2, K, J, N1a, V, and HV (Table 5). Nevertheless, the elevated prevalence of the haplogroup H, the emergence of haplogroup U5a and U2 and the absence of W and X in the LBKT dataset signalize a sort of divergence from the STA.



	HGCN (n=19)	STA (n=44)	VIN (n=31)	LBKT (n=42)	LBK (n=108)	SOP (n=37)	LGY (n=82)	BL (n=13)
X	0	6,82	0	0	0,93	2,70	1,22	0
W	0	4,55	0	0	2,78	0	1,22	15,38
U8	0	0	0	0	0	2,70	6,10	7,69
U5b	52,63	0	3,23	0	0,93	0	3,66	0
U5a	21,05	0	6,45	2,38	1,85	2,70	1,22	0
U5	0	0	0	0	0	0	1,22	0
U4	10,53	2,27	0	0	0	0	0	0
U3	0	2,27	9,68	0	0,93	2,70	0	0
U2	5,26	0	3,23	4,76	0	0	0	0
U	10,53	0	0	0	0	0	0	0
T2	0	20,45	19,35	23,81	22,22	13,51	18,29	23,08
T1	0	2,27	0	2,38	0	0	0	7,69
N1a	0	6,82	12,90	9,52	12,04	10,81	12,20	0
K	0	27,27	29,03	16,67	20,37	13,51	18,29	7,69
J	0	11,36	9,68	7,14	12,04	13,51	7,32	7,69
V	0	6,82	0	2,38	4,63	0	0	0
HV	0	2,27	3,23	2,38	4,63	10,81	7,31	7,69
H	0	6,82	3,23	28,54	16,67	27,00	21,95	23,08

Table 5. MtDNA haplogroup frequencies in the studied Carpathian Basin and Central European populations.

Different haplogroups (from X to H) are marked with different colour shadings. Abbreviations: hunter-gatherer in Central and North Europe (HGCN), Starčevo (STA), Vinča (VIN), LBK in Transdanubia (LBKT), Central European LBK (LBK), Sopot (SOP), Lengyel (LGY), and Balaton-Lasinja (BL) cultures/populations. Frequencies are given in percent (%).

Haplogroup U5a and U2 were both detected in the Vinča and LBKT, though the increased number of haplogroup H makes the LBKT unique in the sequence of the Transdanubian Neolithic cultures. Haplogroup H could not be further differentiated in our analyses focusing on the HVS-I, and HVS-II parts of the mtDNA. Some subgroups of H might have come with the first farmers to Europe, as it has been assumed recently by P. Brotherton and his colleges (Brotherton et al., 2013), whereas others might signalize forager genetic substrate. Complete mitochondrial genome analyses could answer this question.

Regarding the following two late Neolithic, and the Chalcolithic Balaton-Lasinja datasets, we can assume that the majority of the Neolithic mtDNA haplogroup variability was present from the Early Neolithic period in the western Carpathian Basin (Table 5). Late Neolithic novel haplogroups are the U8b, appearing first in the Sopot culture and T2f, emerging in the Lengyel culture. U8 is one of the oldest sub-haplogroup in Europe, beside the U2. U8 have an estimated age of 50,000 years in Europe (Soares et al., 2010), and it has also been detected in pre-ice age Upper Palaeolithic context in Dolní Věstonice (Fu et al., 2013). The first occurrence of U8b in the Carpathian Basin is earlier in the Alföld region (Alföld LBK-Szatmár group, then Tiszadob and Esztár groups) than in Transdanubia (Keerl, 2014). In central Germany, it appeared sporadically in the Schöningen culture (haplogroup U8b), while the U8a variant occurred in the Baalberge group, and then both haplogroups re-emerged in the Únětice culture (Brandt et al., 2013).

Haplogroup T2f appeared first in the Körös culture in Hungary, but it was also present in the Alföld-LBK (Keerl, 2014). It reached Central Germany with the earliest farmers (LBK), and persisted in the succeeding Rössen and Schöningen groups as well (Haak et al., 2010, Brandt et al., 2013). Since these T2f haplogroups from Alföld and Germany are the same haplotype that was detected first in the Lengyel period of Transdanubia, it might be just a coincidence, that T2f has not been found yet in the previous SOP, LBKT or STA datasets.

The haplogroup variability decreased in the Balaton-Lasinja period, but the small sample set (n=13) can also cause frequency biases. After a hiatus, T1a haplogroup appeared in the Balaton-Lasinja dataset again. T1 or T1a has only been detected in the Carpathian-Basin in the sixth-fifth millennia BC. From the Alföld region, T1a is known from the Alföld-LBK-Szakálhát, and Tiszadob groups and from the succeeding Tisza culture as well (Keerl, 2014). Following the Central European Baalberge culture (with one hit from the first half of the fourth millennium BC) T1a became common in the Late Neolithic (local term) Central Germany

(Corded Ware culture, third millennium BC) (Brandt et al., 2013). Besides Central Europe, T1 appeared in the Eastern European and Central Asian Bronze Age as well (Keyser et al., 2009; Lalueza-Fox et al., 2004; Wilde et al., 2014).

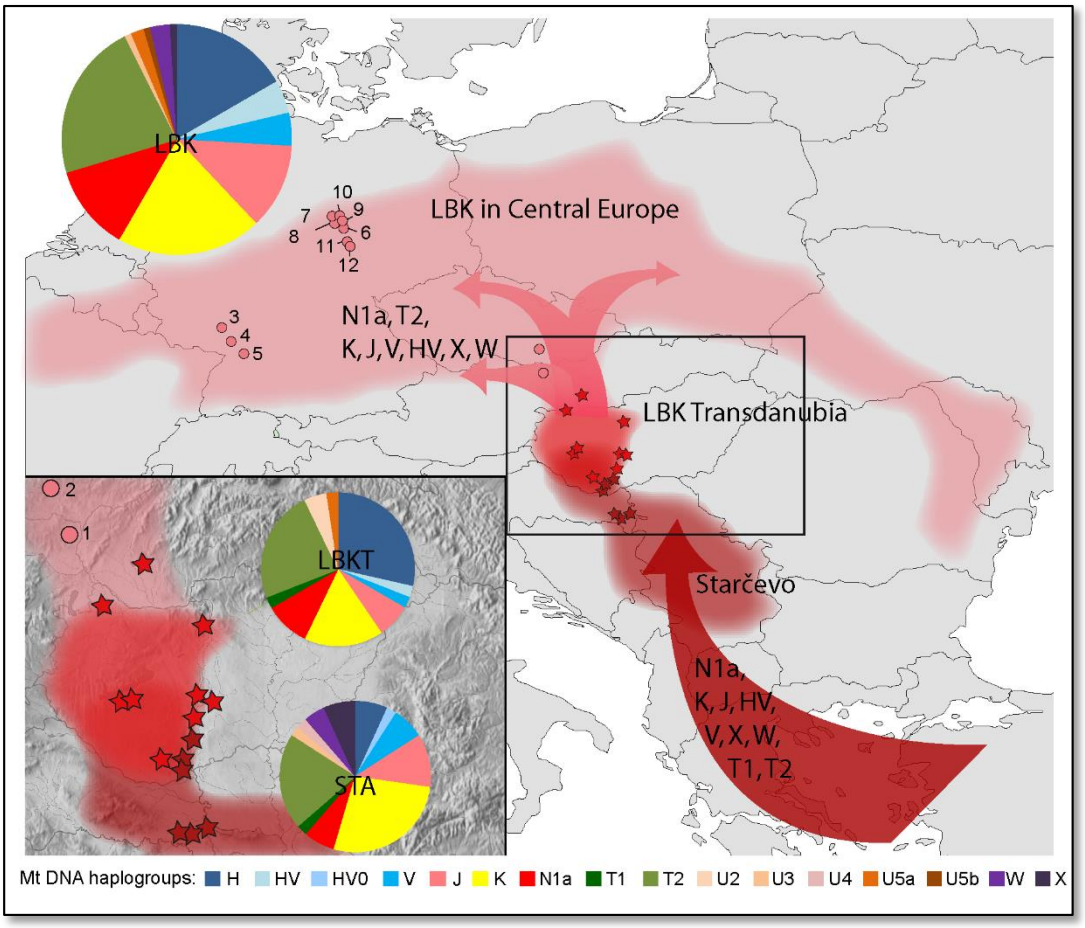


Figure 12. The mitochondrial "Neolithic package" and the Continental Route of the Neolithic dispersal.

For the STA, LBKT site names see Figure 6. Studied LBK sites in Central Europe are the followings: 1. Asparn Schletz, 2. Vedrovice, 3. Flomborn, 4. Schwetzingen, 5. Vaihingen, 6. Seehausen, 7. Derenburg, 8. Halberstadt, 9. Oberwiederstedt, 10. Eilsleben, 11. Karsdorf, 12. Naumburg (Bramanti, 2008; Brandt et al., 2013; Haak et al., 2010, 2005).

Focusing on the sixth millennium BC, it is worth to mention that the mitochondrial diversity of the Carpathian Basin Starčevo and LBKT populations had a general similarity to the LBK in Central Europe (Figure 12, Table 5). Previous studies have shown that haplogroups N1a, T2, J, K, HV, V, W, and X were most characteristic for the Central European LBK and have described these haplogroups as the mitochondrial 'Neolithic package' that reached Central Europe in the mid-sixth millennium BC (Haak et al., 2010; Brandt et al., 2013). Interestingly, most of these haplogroups show comparable frequencies between the STA, LBKT and LBK,

consisting the majority of mtDNA variation in each population (STA=86.36%, LBKT=61.9%, LBK=79.63%). However, there is some inconsistency among the three populations' haplogroup compositions as well. The W and X were characteristic haplogroups of the STA and the LBK in Central Europe, but they have not been found in the LBKT yet (Figure 12). T1 was only present in the Transdanubian Early and Middle Neolithic; it probably did not reached central Germany until the first half of the fourth millennium BC. In contrast to the frequency of the mitochondrial "Neolithic package", hunter-gatherer haplogroups were rare in the STA and in both LBK groups (Figure 12). U3, U5a, and U5b occurred both in the Carpathian Basin and in Central Europe, but U2 and U4 has only been detected in the Transdanubian Early and Middle Neolithic datasets.

The mtDNA haplogroup N1a was probably the most characteristic marker of the first farmers' maternal gene pool (Haak et al., 2005; Palanichamy et al., 2010). Its frequency is very low in modern Europe (~0.2%, Haak et al., 2005), slightly higher (~2%) in today's Saudi Arabia (Abu-Amero et al., 2008), Yemen (Kivisild et al., 2004), and it has been found to be frequent on the Croatian Cres Island (Jeran et al., 2009). N1a occurs in North Africa, Ethiopia and in Asia as well. In Central Asia it is known among e.g. the Kazakhs (Gokcumen et al., 2008), Altaians, Buryats (Derenko et al., 2007). The modern N1a variation is divided into the European/Central Asian and African/South Asian branches, based on specific genetic markers. The European/Central Asian branch is characterized by 16147A, 3336C, and 16320T positions (Haak et al., 2005; Derenko et al., 2007).

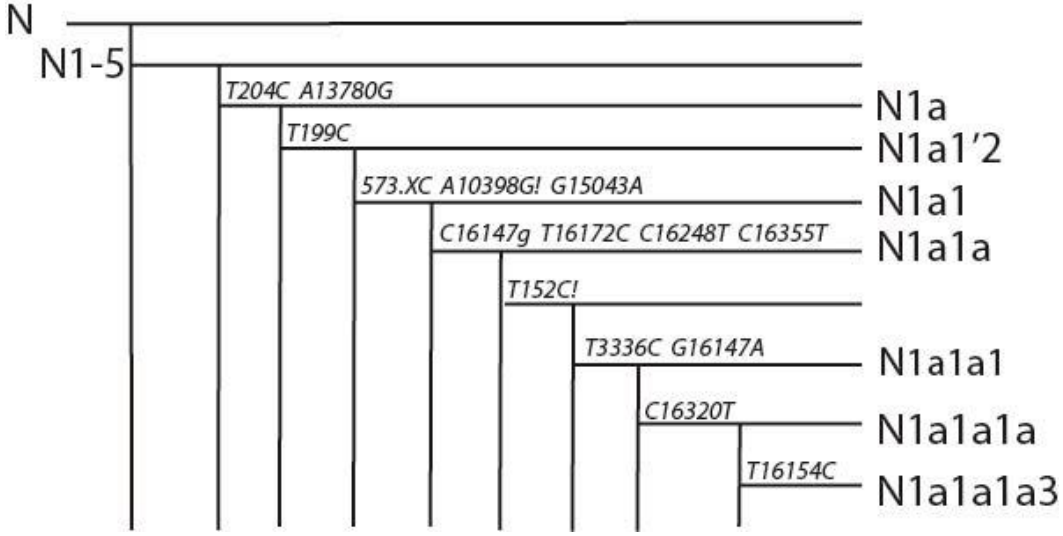


Figure 13. Simplified N1a phylogeny with the mentioned SNPs.

[Based on www.phylotree.org. Build 16 (19. Feb 2014)]

All the 25 N1a haplotypes from the Transdanubian Neolithic carry the position 16147A, but only a smaller part (13/25) harbours the 16320T [after the current nomenclature (phylotree 16, see in Figure 13) it is called as N1a1a1a, see the haplotypes in Supplementary Table 3. Position 3336C was not study by me, hence it is outside of the control region (i.e. HVS-I, II) of the mitochondrial genome. The 12 Transdanubian N1a haplotypes, which do not show the derived 16320T, can be ordered as N1a1a1*. They represent an archaic sub-cluster, which has probably gone extinct by now. I detected one N1a individual (LGCS2) with the position 16189C, which SNP is described as the marker for the Central Asian branch of the N1a1a1a. Since this haplotype does not carry the 16320T, it could not belong to the Central Asian branch either. It is interesting, that it shares its HVS-I positions with the Alföld-LBK N1a individual from Ecsegfalva 23 (Haak et al., 2005).

I calculated molecular diversity indices for the Transdanubian cultures and for the two most relevant comparative datasets, the Central/North European hunter-gatherers and the Central European LBK. The column "S" in Table 6 contains the number of nucleotide positions at which a polymorphism was found (called as the number of segregating sites). After the LBK, the SOP has the second most segregating sites among the studied datasets. The haplotype diversity (Hd) is the lowest in the HGCN group (0.92398) and rises notably in the Starčevo dataset (0.97674). Among the Neolithic groups, the LBKT has the lowest haplotype diversity value. The average number of haplotype differences (K) is the highest in the Vinča dataset. The nucleotide diversity (π) signalizes the average number of nucleotide differences per site between any two DNA sequences chosen randomly from a certain dataset (Nei and Li, 1979). Nucleotide diversity is the lowest in the HGCN group too (0.00875), and the highest in the Vinča dataset (0.1828).

Considering the Tajima's D values and their significances, it can be assumed, that they are all negative, which indicates population size expansion (e.g., after a bottleneck or a selective sweep) and/or purifying (negative) selection (Tajima, 1989). The deviation of D from zero was only in the SOP dataset significant ($p<0.05$), meaning that the null hypothesis of "neutrality" (randomly evolving DNA sequences) can be rejected in this case. The Tajima's D value of the SOP dataset suggests non-random process, such as selection or demographic expansion, contraction or genetic introgression.

Culture	n	S	h	Hd	K	π	Tajima's D	Tajima's D p-value
HGCN	19	14	12	0.92398	3.01754	0.00875	-0.9150	0.1940
STA	44	33	30	0.97674	5.27801	0.01539	-1.0399	0.1482
LBKT	42	34	23	0.95238	4.98374	0.01453	-1.2740	0.0820
VIN	31	32	23	0.97419	6.27097	0.01828	-0.7837	0.247
SOP	37	38	25	0.96697	4.87988	0.01423	-1.6452	0.027
LGY	82	36	36	0.96868	5.0801	0.01481	-0.9427	0.18
BL	13	21	11	0.97436	4.97436	0.0145	-1.1373	0.141
LBK	108	39	40	0.95483	4.94393	0.01446	-1.0324	0.1510

Table 6. Summary statistics.

Molecular indices were computed based on HVS-I sequences (np 16,056-16,402) of the Transdanubian cultures and the hunter-gatherer and LBK groups. Culture information about HGCN and LBK are presented in Supplementary Table 6. Abbreviations: n = sample size, S = number of segregating sites, h = number of haplotypes, Hd = haplotype diversity, K = average number of nucleotide differences, π = nucleotide diversity, D = Tajima´s D. For abbreviations of the cultures, see the Glossary or the legend of Table 5.

5.1.4 Population genetic analyses in an Eurasian prehistoric context

5.1.4.1 PCA and Ward clustering with 29 prehistoric datasets

Comparing the Transdanubian datasets to the available Eurasian aDNA records from the Upper-Palaeolithic to the Bronze Age (see Methods, chapter 4.3.6), A PCA with 29 populations/metapopulations was performed, which displayed the frequencies of 25 haplogroups in two dimensions (Figure 14, Supplementary Table 7). The first two principal components (PCs) represent 28.5% of the total variance. The third PC contains 9.6% of the variance and each further PC shows lower than 7.9% of the total variation with a gradually decreasing proportion. The first component (X-axis) separates the different hunter-gatherer groups (HGE, HGCN, HGSW, PWC) from the sixth-fourth millennia BC Central European and Carpathian Basin populations. The haplogroups U2, U4, and U5a that are frequent in the hunter-gatherer groups, have the highest loading toward the negative direction on PC1, while J, K, N1a, and T2 haplogroups have the highest loading with an opposite trend. The J, K, N1a, and T2 haplogroups cluster the sixth-fourth millennia BC populations into one cluster on the plot (Supplementary Figure 1). The Bronze Age groups of Europe and Asia come apart along the second component from the Iberian Neolithic and the hunter-gatherer datasets. The haplogroups D, T, T1, I, U8, and W have the highest loading toward the positive direction on the second PC, differentiating the eastern Bronze Age datasets. Haplogroup-vectors U5b, N, and V point toward the negative direction along the second component, separating the Iberian Neolithic populations and the hunter-gatherers.

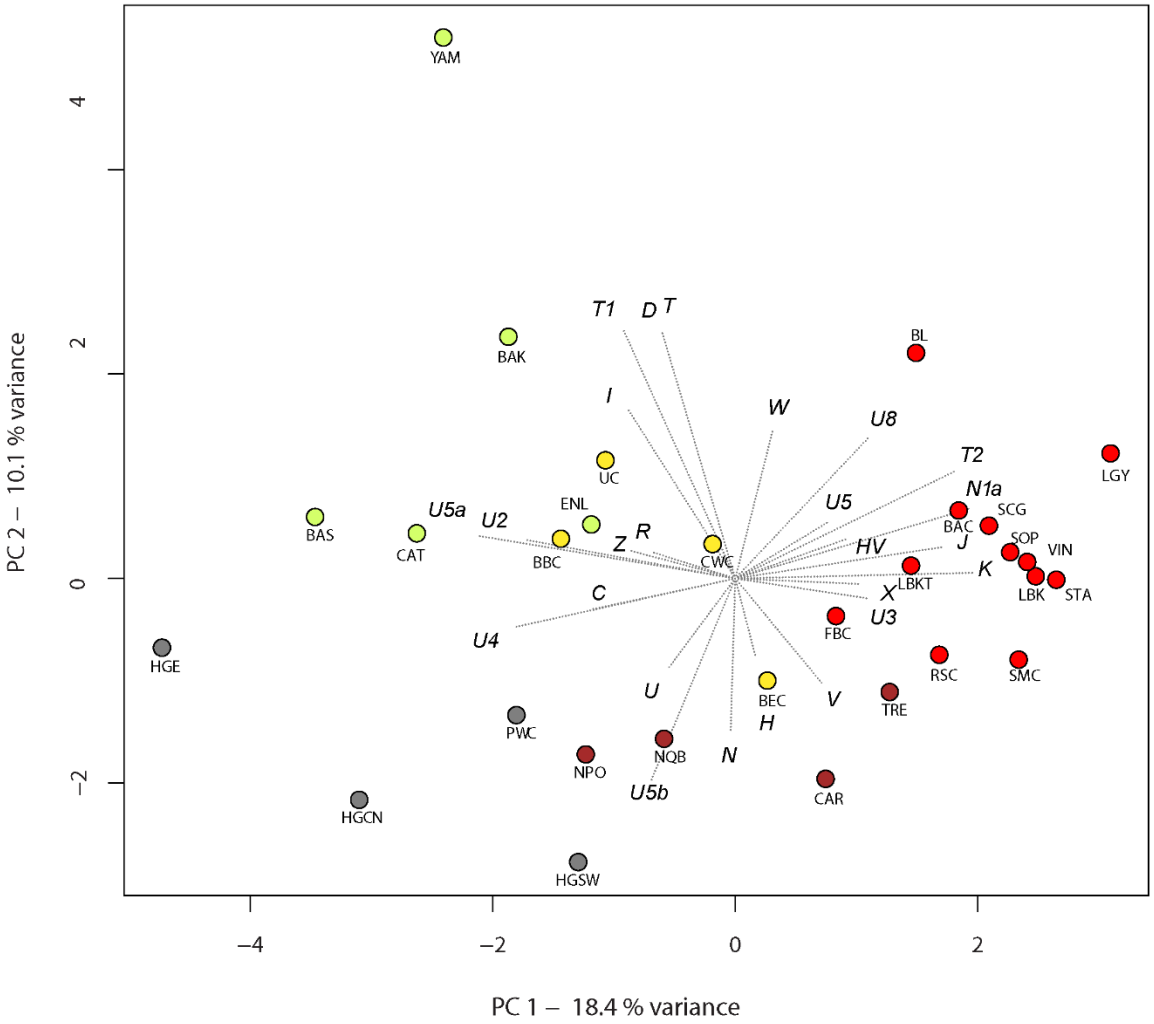


Figure 14. PCA with 29 prehistoric populations (29POP) from Europe and West/Central Asia.

The PCA represents 28.5% of the total haplogroup variation. The sixth-fourth millennia BC Central European and Carpathian Basin populations are signalised with red-filled circles, the third-second millennia BC Late-Neolithic/Early Bronze Age datasets in Central Europe are yellow. The eastern Eneolithic/Bronze Age populations are shown as green circles. The hunter-gatherers are in grey, and the Iberian populations are brown. The frequency data are presented in Supplementary Table 7. Culture/population abbreviations: hunter-gatherers in Central and North Europe (HGCN), hunter-gatherers in South-West Europe (HGSW), hunter-gatherers in East Europe (HGE), LBK in Central Europe (LBK), Starčevo culture (STA), LBK culture in Transdanubia (LBKT), Vinča culture (VIN), Sopot culture (SOP), Lengyel culture (LGY), Balaton-Lasinja culture (BL), Rössen culture (RSC), Schöningen group (SCG), Baalberge culture (BAC), Salzmünde culture (SMC), Bernburg culture (BEC), Corded Ware culture (CWC), Bell Beaker culture (BBC), Únětice culture (UC), Cardial and Epicardial culture (CAR), Neolithic Basque Country and Navarre (NBQ), Neolithic Portugal (NPO), Pitted Ware culture (PWC), Funnel Beaker culture (FBC), Treilles culture (TRE), Bronze Age Siberia (BAS), Bronze Age Kazakhstan (BAK), Eneolithic on the North Pontic-Caspian Steppe (ENL), Yamnaya culture (YAM), Catacomb culture (CAT).

There is a clear separation of the Neolithic sixth-fourth millennia BC Central European and Carpathian Basin datasets from the third-second millennia BC Bronze Age populations on the PCA plot. The hunter-gatherers are also separated, but they are closer to the Iberian Neolithic populations than to other Neolithic and Bronze Age datasets. This result is due to the frequency of haplogroups U5a, U5b, U4, and N*, which appear in both clusters. The group of the sixth-fourth millennia BC datasets has a central core, composed of the STA, LBK, VIN, and SOP populations. They both share the haplogroups N1a, K, J, X, and U3, which haplogroup-vectors point to their direction. LBKT is slightly apart from this central group, probably due to the elevated frequency of haplogroup H in this dataset. BL and LGY are also somewhat separated, because of the haplogroups U8 (in LGY and BL), T1, and W (only in the BL). Interestingly, the south French Treilles and the northern German/southern Sweden Funnel-Beaker (*Trichterbecher*) datasets are close to the Central-European/Carpathian-Basin cluster. Their affinities to the Central European Middle Neolithic have been revealed by Brandt et al. as well (Brandt et al., 2014, 2013). A set of individuals assigned to the Bernburg culture is situated between the sixth-fourth and the third-second millennia BC Central European populations, it shows several characteristics (U5a-U5b, no N1a), that point toward the Endneolithic of this region (Brandt et al., 2013).

The score data of the PCs (used for the PCA plot), were reused for a hierarchical clustering analysis of Ward type with Euclidean distance calculation method. The dendogram on Figure 15 helps verifying the PCA results. Out of the two major branches, the first contains the sixth-fourth millennia BC populations mixed with two third millennium BC populations, the Bernburg dataset from Central Germany and the Treilles dataset from France. The LBK, LBKT, and SOP are neighbours to each other, with high probability values. Surprisingly, the STA is closest to the Bernburg population, though it has a low probability value (67%).

The Eneolithic and Bronze Age datasets from East Europe, Central Asia, and three groups from the third-second millennia BC Central Europe are on the second branch, adjacent to the hunter-gatherers and the Iberian Neolithic groups (with a zero p value, which means that this major clusterings are not robust or sure). It is shown on the dendogram that the Neolithic populations from Portugal and Basque Country and Navarra are on neighbouring branches as well as the Iberian hunter-gatherers and the (Epi)Cardial population. From the third-second millennia BC, the Únětice, Corded Ware, and Bell Beaker are also in neighbouring

branches, with high p values (%) in the latter two cases. They are also close to the Eneolithic of the north Pontic steppe and to the Transdanubian Chalcolithic Balaton-Lasinja population.

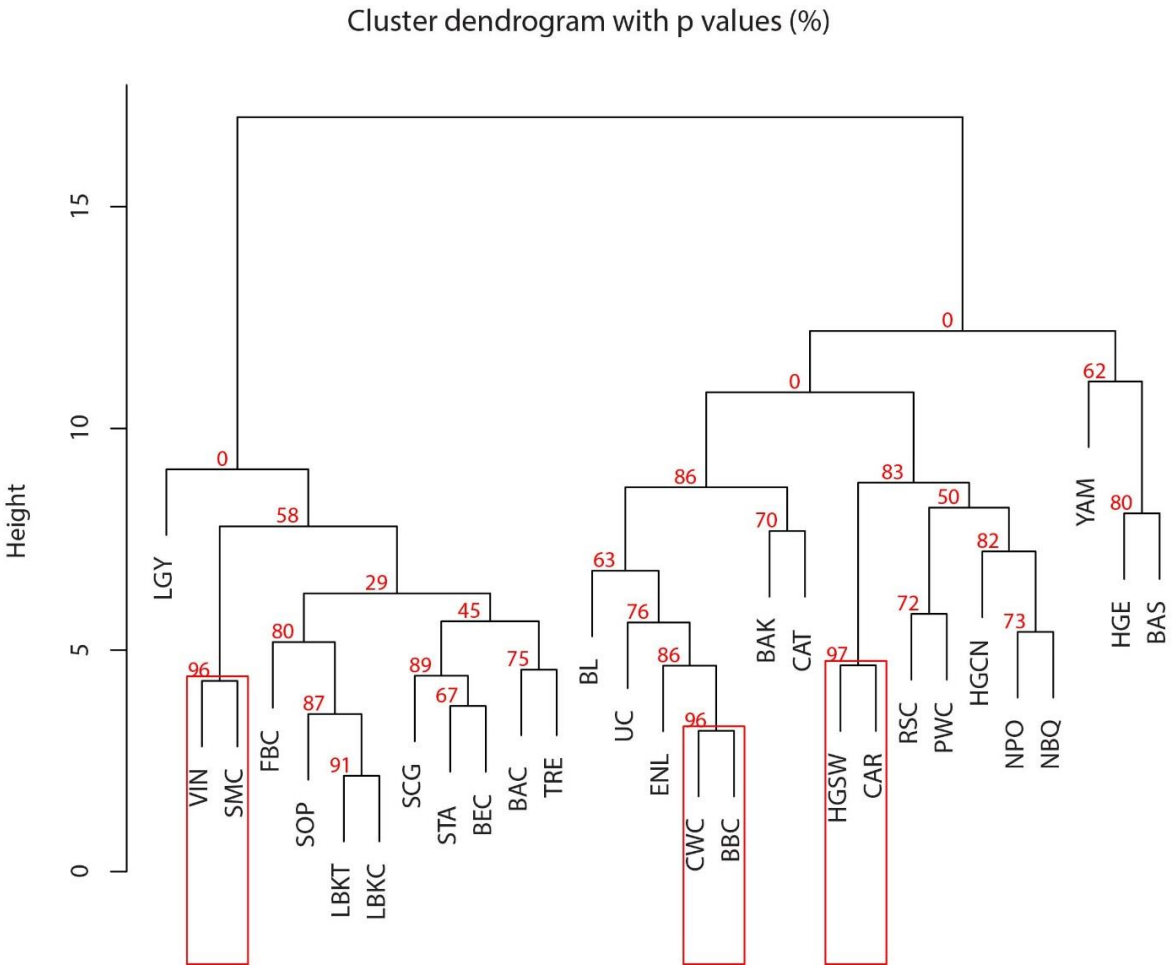


Figure 15. Ward clustering of 29 prehistoric populations (29POP) calculating with Euclidean distances.

25 principal components were used for the analysis. Culture/population abbreviations are resolved in legend of Figure 14. Probability values (p values in %) are written in red. Clusters that are highly supported by the data are marked with red rectangles.

The dendogram presented in Figure 15 is mainly consistent with the PCA, although some smaller alterations exist on the branch tails, which is probably a result of methodical differences. Whereas in the PCA the first two components were captured, in the hierarchical clustering all PCs (PC1-25) were considered. Reducing the number of PCs in the cluster analysis results in a more similar dendogram to the PCA, where STA, LBK, VIN, and Salzmünde populations are on adjacent branches (data not shown). Because each of the PC3-25 presents

the proportion of the variance in gradually decreasing amount, I used all PCs in the Ward clustering, in order to keep the analysis unbiased.

5.1.4.2 Genetic distances between 29 prehistoric cultural groups

The HVS-I sequences were used for genetic distance calculations (Figure 16, Supplementary Table 8). Significant genetic distances were detected between the three hunter-gatherer groups and almost every Neolithic, Chalcolithic, and Bronze Age datasets. All of the Carpathian Basin populations differ significantly from the hunter-gatherers (HGCN, HGSW, and HGE) and from the Pitted Ware and Catacomb populations. Except the pair of the STA and Cardial, all Transdanubian datasets are in significant distances from the Iberian Neolithic. The Carpathian Basin datasets build one cluster with the Central European sixth-fourth millennia BC populations (LBK, Rössen, Schöningen group, Baalberge culture, Salzmünde culture), showing no significant differences from each other. Furthermore, the Funnel-Beaker, Siberian Bronze Age group, Kazakhstan, and the Yamnaya show relative small distances from the studied Transdanubian populations as well. From the Late Neolithic/Early Bronze Age Central Europe, the Bernburg population shows smaller (no significant) distances from the Transdanubian Neolithic groups. The Corded Ware is significantly different from the LBKT, VIN, and LGY datasets, and the Bell Beaker and Únětice are significantly different from all Transdanubian Neolithic populations. Interestingly, the Balaton-Lasinja data shows affinity to the Late Neolithic/Early Bronze Age Central European cultures' people (CWC, BBC, and UC).

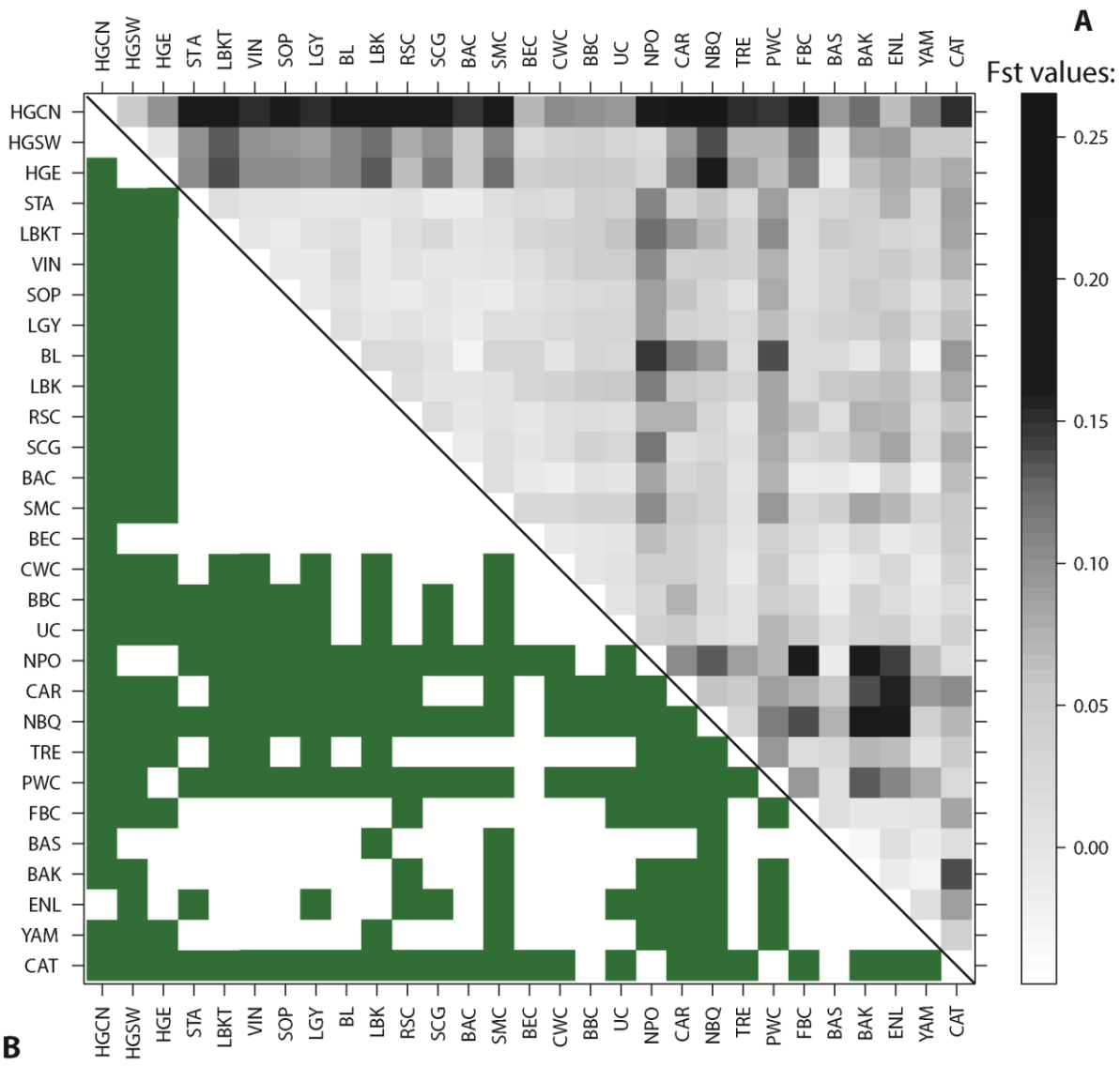


Figure 16. F_{st} level-plot (A) and significant ($p<0.05$) F_{st} values (B) of 29 prehistoric populations (29POP) from Europe and West/Central Asia.

In the upper diagonal (A), the greater genetic distances are shaded black or dark grey, shorter distances are brighter grey or white. The significant F_{st} values ($p<0.05$) are shown with green colour in the lower diagonal (B). Culture/population abbreviations are according to Figure 14. Specific F_{st} values and p values are seen in Supplementary Table 8.

5.1.4.3 Multidimensional scaling of 29 prehistoric cultural groups

The F_{st} values, plotted in Figure 16, were linearized to a Slatkin matrix, and used for the multidimensional scaling (Figure 17, Supplementary Table 9). The stress value (reliability or goodness of fit statistic to the Shepard diagram) was 0.1496 of this MDS outcome. The datasets were coloured and abbreviated as on the Figure 14. The sixth-fourth millennia BC cultures' people (red circles) are in one sector of the plot. Hunter-gatherers are toward the negative range along the first coordinate, separated from the sixth-fourth millennia BC populations by the Late Neolithic/Bronze Age third-second millennia BC datasets of Central Europe and Eastern Europe/Central Asia. The cluster of the Iberian Neolithic populations in the negative range on the second coordinate is divided by the Pitted Ware and the Catacomb

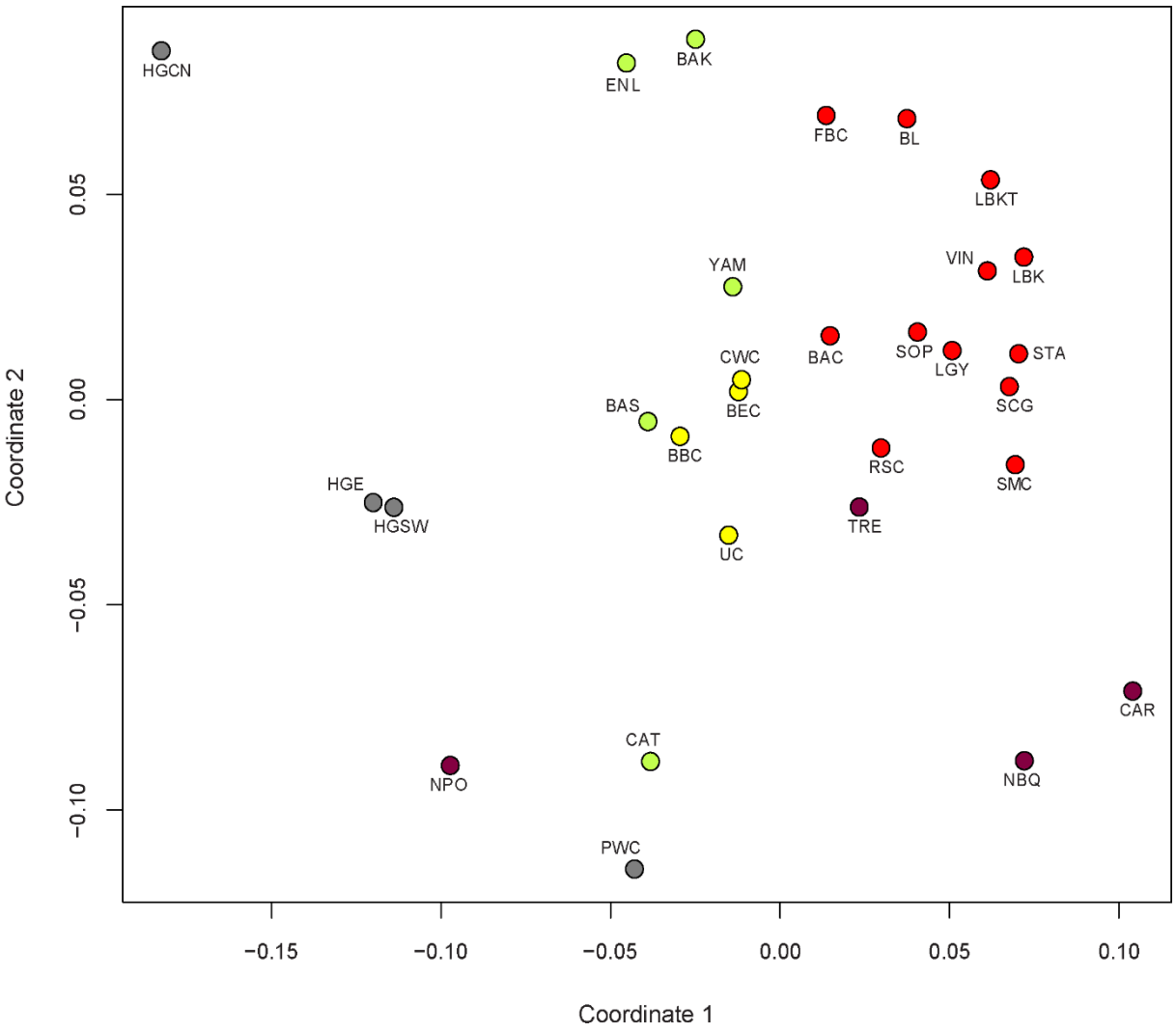


Figure 17. MDS with 29 prehistoric populations (29POP) from Europe and West/Central Asia.

Culture abbreviations are according to the legend of Figure 14. Stress value is 0.1496.

cultures' populations. The Portugal Neolithic data have a greater affinity to the hunter-gatherers than to the (Epi)Cardial population and to the Neolithic samples of Basque County and Navarra region.

Summarizing the MDS results, although the sixth-fourth millennia BC and the third-second millennia BC Central European cultures' people group in two clusters, they are adjacent to each other and to the eastern Late Neolithic-Early Bronze Age groups. The hunter-gatherers and the Iberian Neolithic are clearly detached on the MDS plot, along with the Pitted Ware and the Catacomb populations, which do not order in to their supposed geographic or chronologic clusters.

5.1.4.4 PCA and MDS of 20 prehistoric populations using two parallel datasets

A PCA and a MDS were performed with 20 prehistoric populations (referred as 20POP dataset), focusing on Europe, especially on Central Europe and the Carpathian Basin (Supplementary Table 10, 11). Nine populations were omitted from the 29POP dataset analysed in the previous chapters. The Central European populations, the Iberian Neolithic, the Iberian hunter-gatherers, and the Central/North European hunter-gatherers were retained. Two sets of data were used for each population: I marked those with an asterisk (*), which are free from potential maternally related individuals. There were no intra-site haplotype matches within the Baalberge dataset (BAC), therefore the complete and the reduced datasets are the same. It can be assumed from the Figure 18 that the intra-site maternal relationships do not influence the result of the PCA relevantly. The reduced datasets (signed with an "*") are close to the location of their original complete datasets. The Neolithic Carpathian Basin and the sixth-fourth millennia BC Central European populations form one cluster, echoing the 29POP PCA. Only the Rössen and the Balaton-Lasinja are slightly split off along the second PC from the sixth-fourth millennia BC cluster. They are both represented by relatively few samples, which have to be considered, when conclusions or hypotheses are drawn from these results.

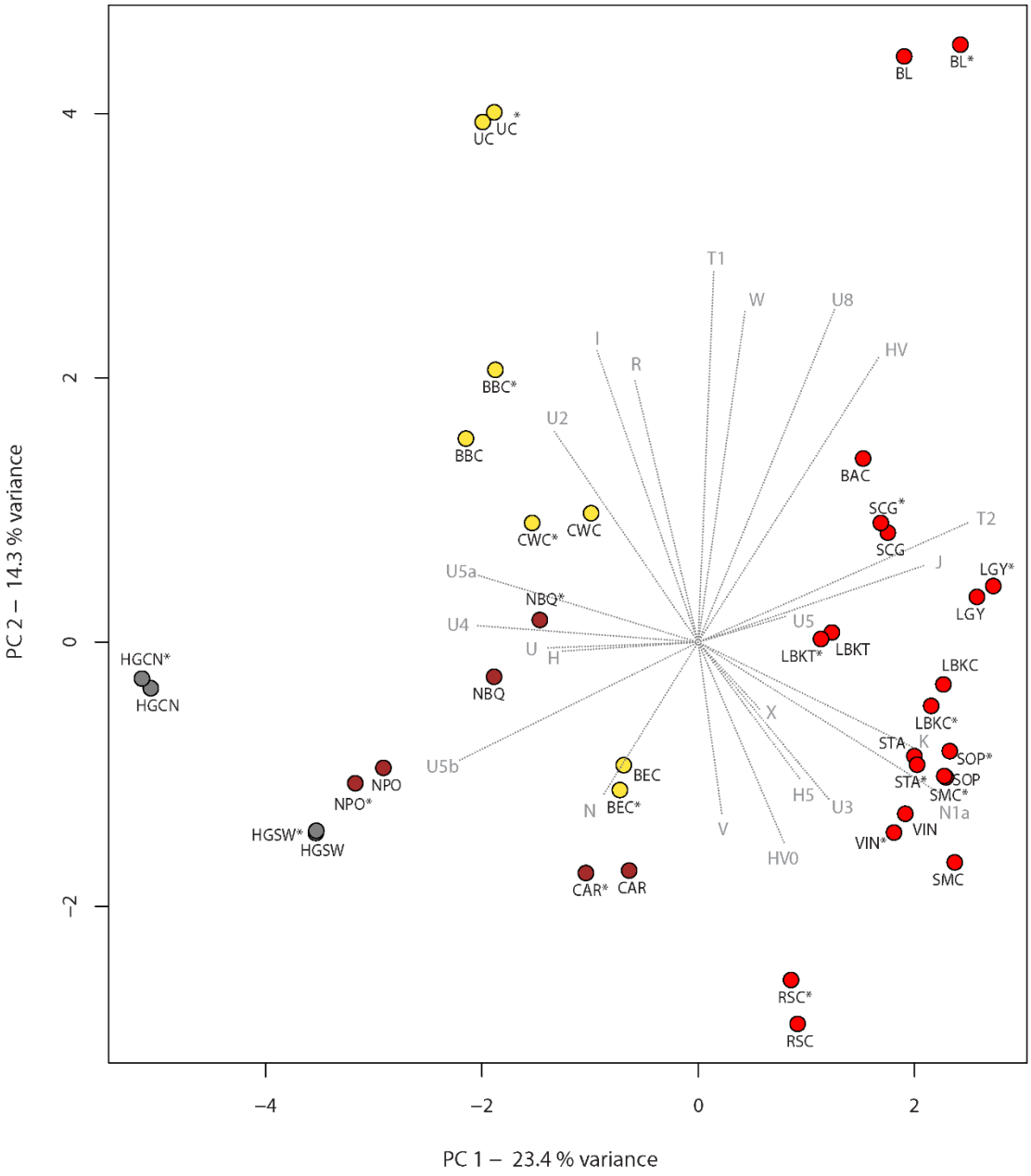


Figure 18. PCA with 20 prehistoric populations (20POP) from Europe, represented in 39 datasets.

The PCA represents 37.7% of the haplogroup variation. Two sets of data were used for each population: a complete dataset and a reduced one (*), which does not contain potential maternally related individuals. The sixth-fourth millennia BC Central European and Carpathian Basin populations are signalised with red-filled circles, the third-second millennia BC Late-Neolithic/Early Bronze Age datasets in Central Europe are yellow. The hunter-gatherers are in grey, and the Iberian Neolithic is in brown. Culture/population abbreviations: hunter-gatherers in Central and North Europe (HGCN), hunter-gatherers in South-West Europe (HGSW), LBK in Central Europe (LBK), Starčevo culture (STA), LBK in Transdanubia (LBKT), Vinča culture (VIN), Sopot culture (SOP), Lengyel culture (LGY), Balaton-Lasinja (BL), Rössen culture (RSC), Schöningen group (SCG), Baalberge culture (BAC), Salzmünde culture

(SMC), Bernburg culture (BEC), Corded Ware culture (CWC), Bell Beaker culture (BBC), Únětice culture (UC), Cardial and Epicardial culture (CAR), Neolithic Basque Country and Navarre (NBQ), Neolithic Portugal (NPO).

It is noteworthy that the Bell Beaker dataset is the closest to the Corded Ware on the 29POP and 20POP PCAs (Figure 14Figure 18). On the PCA plot of Brandt et al., BBC has shown an affinity to the Iberian Neolithic (especially to the Neolithic Portugal) (Brandt et al., 2013), which result is not supported by my data. The difference between the two analyses is that I involved the Carpathian Basin data, the recent north Pontic study (Wilde et al., 2014), and further six Bell Beaker samples (Lee et al., 2012) into the here presented evaluation.

On the MDS plot with a 20POP assembly (Figure 19, Supplementary Table 11), the two parallel datasets of each population are slightly more separated in some cases: the genetic picture of the groups from the Neolithic Portugal, Basque Country and Navarre, and Salzmünde dataset are also influenced by maternal relatives. At these cases, the reduced datasets in the populations' comparison should be rather considered than the complete datasets. Nevertheless, the reduced sets still group into their previously described clusters. LBK in Central Europe is close to the LBK in Transdanubia and the Starčevo, Lengyel, and Vinča datasets are also very similar to each other, surrounded by the Schöningen, Salzmünde, Baalberge and Rössen populations. It is seen on Figure 19 that the separation of the yellow (third-second millennium BC) and red (sixth-fourth millennium BC) clusters becomes more apparent than on the 29POP MDS plot. Since the number of datasets is smaller in the 20POP, the genetic variation is simpler to visualize, which is apparent on the decreasing stress value too.

The Bernburg population attributed to the Late Neolithic/Early Bronze Age cluster (third-second millennium BC) by Brandt et al., shows different affinities on the 29POP and 20POP PCAs than on the two MDS plots. The Bernburg culture has been archaeologically associated with the Funnel Beaker complex, and the increasing North European/Funnel Beaker elements in its maternal gene pool (containing hunter-gatherer lineages) has been demonstrated as well (Brandt et al., 2013). This connection is seen on the 29POP PCA plot and on the F_{st} analysis. The Bernburg dataset is the only one, besides the eastern hunter-gatherers, which does not have significant genetic distances from the hunter-gatherer Pitted Ware population from the Scandinavian Neolithic period. However, this pattern of connections does

not prevail on the 29POP MDS plot and on the Ward cluster dendogram, probably due to the hardly reducible complexity of the total genetic variation.

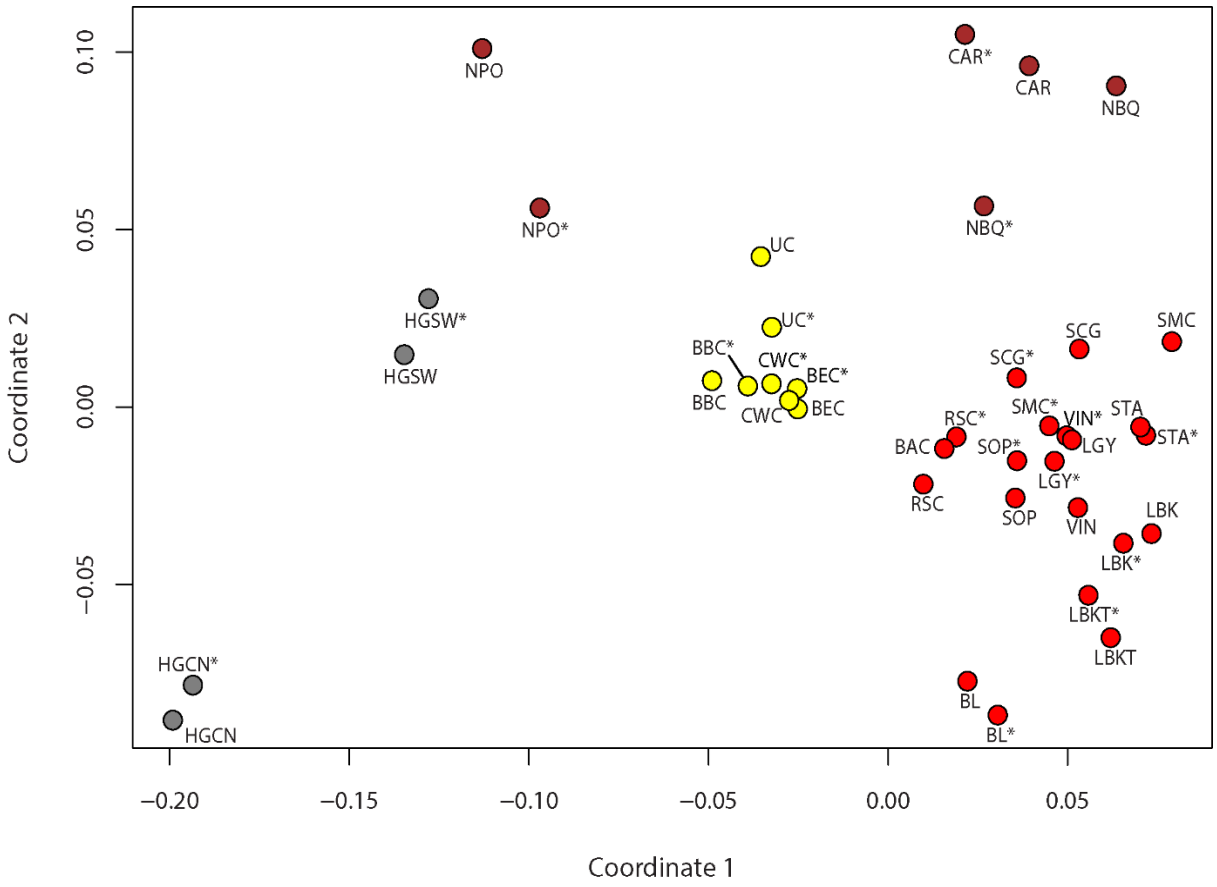


Figure 19. MDS with 20 prehistoric populations (20POP) from Europe represented in 39 datasets.

Stress value is 0.1255. Population abbreviations, colour shadings, and dataset compositions are according to Figure 18.

5.1.4.5 Analysis of molecular variance

AMOVA analysis was performed in order to assure the separation of the sixth-fourth millennia BC Carpathian Basin and Central European populations from the third-second millennia BC Central European populations, which was previously assumed from the 20POP MDS plot (Figure 19).

An analysis of molecular variance has been reported with the STA, LBKT and nine Central European populations (Szécsényi-Nagy et al., 2014a). In this previous analysis, we have tested 82 arrangements. The highest among-group variance was observed when STA and LBKT were arranged in one group with the Central European LBK and with all fifth-fourth millennia BC populations, while the third-second millennia BC populations (Bernburg, Corded Ware, Bell

Beaker, Únětice cultures) were arranged in a second group (among-group variation=3.50%, F_{st}=0.03501, p=0.00396; within-group variation=0.20%, F_{st}=0.00203, p=0.31139) (Szécsényi-Nagy et al., 2014a). Based on these 82 tests and the clustering observed on the 20POP and 29POP MDS plots (Figure 17, Figure 19), I tested the ten most plausible arrangements of the 15 populations (STA, VIN, LBKT, SOP, LGY, BL, LBK, RSC, SCG, BAC, SMC, BEC, CWC, BBC, UC, see Supplementary Table 12). The highest among-group variance was observed again, when BEC, UC, BBC, CWC are in one group, and all the sixth-fourth millennium BC populations in the second group (among-group variation=3.21%, F_{st}=0.03207, p=0.00050; within-group variation=0.1%, F_{st}=0.00108, p=0.3704). Although the Rössen and Baalberge datasets are close to the third-second millennia BC cluster on the 20POP MDS plot, they arrangement with the sixth-fourth millennia BC populations obtains a low within-group and high among-group variance.

5.1.4.6 Ancestral shared haplotype analysis (ASHA)

I used shared haplotype analysis (Excoffier and Lischer, 2010) and also a modified approach by considering the temporal succession of the cultures/populations (the method is called ancestral shared haplotype analysis, for a description see chapter 0, and Szécsényi-Nagy et al., 2014a). This enabled me to ascribe mtDNA lineages to particular cultures/populations or time periods according to their first appearance in a defined chronological order (Figure 20, Supplementary Table 13), and to estimate the amount of ancestral lineages in each population, potentially derived from hunter-gatherers, STA, VIN, LBKT, LBK, SOP, LGY, BL, and other subsequent populations.

The hunter-gatherer lineages were scarce in the Early and Middle Neolithic (sixth-fourth millennia BC) populations, their appearance did not reach the 10% level. However, they reappeared in the Late Neolithic Bernburg dataset in Central Europe, which is in accordance with the previous statements about the Northern European influx affecting this culture's people (Brandt et al., 2013). Focusing on the impact of the Carpathian Basin populations on Central Europe Neolithic, it can be assumed that the genetic variability of the Starčevo culture's people had the strongest effect on the maternal gene pool of the succeeding populations in both regions. This strong "Starčevo" or Early Neolithic genetic effect lasted even

in the Bronze Age Únětice period, when 36.17% of the mtDNA lineages are still deducible from the Starčevo culture.

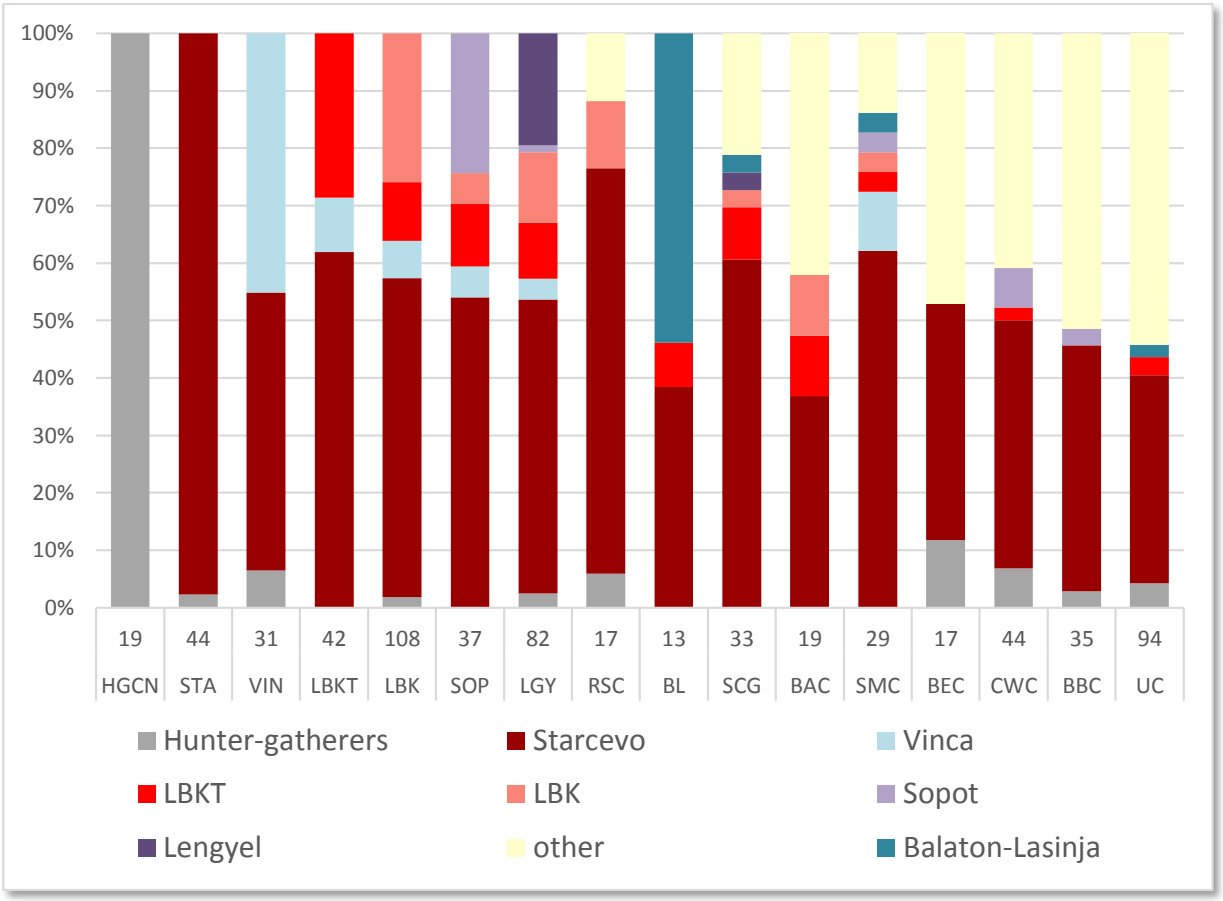


Figure 20. ASHA with 16 populations from the Carpathian Basin and Central Europe. Ancestral lineages were differentiated with colour shadings. Hunter-gatherer ancestral lineages are grey, Starčevo lineages are dark red, etc. The numbers above the culture abbreviations indicate the studied sample sizes. Culture/population abbreviations: hunter-gatherers in Central and North Europe (HGCN), LBK in Central Europe (LBK), Starčevo culture (STA), LBK in Transdanubia (LBKT), Vinča culture (VIN), Sopot culture (SOP), Lengyel culture (LGY), Balaton-Lasinja (BL), Rössen culture (RSC), Schöningen group (SCG), Baalberge culture (BAC), Salzmünde culture (SMC), Bernburg culture (BEC), Corded Ware culture (CWC), Bell Beaker culture (BBC), Únětice culture (UC). The ancestral lineages younger than the LBK in the Middle-Elbe Saale region were not differentiated.

Further Carpathian Basin populations also show some matches in the Central European succeeding datasets, but they never exceed a level of 11%. It raises the question, whether the VIN and SOP lineages in Central Europe mean novel influx(es) from the Carpathian Basin. In my opinion, it could rather indicate the legacy of the Early Neolithic substrate as well, which

was coincidentally not detected in the STA-LBKT period of the Carpathian Basin. On the other hand, the small amount of LGY and BL matches in the Schöningen (4,100-3,950 cal BC), and the ancestral BL haplotypes in the Salzmünde population (3,400-3,100/3,000 cal BC) could have represented new influxes in Central Germany from the south.

5.1.5 Population genetic analyses, focusing on the western Carpathian Basin

5.1.5.1 Test of population continuity

I performed a test of population continuity (TPC), using a Marcov chain Monte Carlo simulation, as described in a study of Brandt et al., 2013. I compared the absolute mtDNA haplogroup frequencies of the hunter-gatherers, the six Transdanubian populations, and the

A	**Ne= 500**							
	HGCN	STA	VIN	LBKT	LBK	SOP	LGY	BL
HGCN	*	7.79E-07	7.79E-07	1.50E-06	7.79E-07	7.79E-07	7.79E-07	3.55E-06
STA		*	0.448192	0.285787	0.741119	0.276345	0.234067	0.3016185
VIN			*	0.093455	0.230838	0.15959	0.127303	0.0200295
LBKT				*	0.54612	0.264595	0.206091	0.3964887
LBK					*	*	0.045311	0.0375817
SOP						*	0.751276	0.3463248
LGY							*	0.3076714
BL								*
B	**Ne= 5000**							
	HGCN	STA	VIN	LBKT	LBK	SOP	LGY	BL
HGCN	*	7.79E-07	1.50E-06	1.50E-06	7.79E-07	7.79E-07	7.79E-07	1.50E-06
STA		*	0.467734	0.284519	0.73551	0.277111	0.241386	0.2973407
VIN			*	0.096081	0.237004	0.14872	0.121791	0.018019
LBKT				*	0.563599	0.261269	0.232716	0.4013401
LBK					*	*	0.043637	0.0332718
SOP						*	0.756707	0.3213022
LGY							*	0.3114789
BL								*
C	**Ne= 30000**							
	HGCN	STA	VIN	LBKT	LBK	SOP	LGY	BL
HGCN	*	7.79E-07	2.88E-06	7.79E-07	7.79E-07	7.79E-07	7.79E-07	1.50E-06
STA		*	0.456832	0.272396	0.731391	0.283275	0.266206	0.304034
VIN			*	0.083829	0.239976	0.164966	0.114606	0.0185849
LBKT				*	0.578721	0.257999	0.24005	0.3877343
LBK					*	*	0.039987	0.0316754
SOP						*	0.745858	0.3463248
LGY							*	0.3076714
BL								*

Table 7. Results of the test of population continuity.

The table are divided into three sections, representing the p values calculated with effective population sizes (Ne) 500, 5,000 and 30,000 (4A-4B-4C table parts). Light green shade signalizes p values smaller than 0.05, dark green shade indicates p values smaller than 0.001. Culture/population abbreviations: hunter-gatherers in Central and North Europe (HGCN), LBK in Central Europe (LBK), Starčevo culture (STA), LBK in Transdanubia (LBKT), Vinča culture (VIN), Sopot culture (SOP), Lengyel culture (LGY), Balaton-Lasinja (BL).

Central European LBK. I tested the continuity with three hypothetic population sizes (Ne=500, 5,000, 30,000), using differences between the terminal BC dates of the cultures in the simulation. Since the LBK and the SOP have almost the same terminal dates, they could not be compared by this analysis.

For all three effective population sizes, comparable p values were obtained (Table 7). Highly significant p values were found comparing the hunter-gatherers in North/Central Europe to all Neolithic-Chalcolithic populations. It means a population discontinuity at the turn of the Mesolithic/Neolithic transition, which cannot be explained by genetic drift alone.
Further significant p values were detected between the VIN-BL, LBK-LGY, LBK-BL pairs of populations. Since VIN and BL are not direct descendants of each other, and STA-VIN, VIN-LBKT, and VIN-SOP do not show significant p values, the genetic difference between VIN and BL can rather be interpreted as marker for a small rate of infiltration, which accumulated into a significant genetic difference over the elapsed ~1,000 years between the two periods. The LBK-LGY and LBK-BL differences are either a consequence of the different pre-Neolithic substrate of Central Europe from that of the Carpathian Basin, or new influx(es), arriving into the western Carpathian Basin after the dispersal of the LBK(T).

5.1.5.2 F_{st} analysis and AMOVA of the western Carpathian Basin cultures

The F_{st} analysis of the Transdanubian populations, comparing to the two most important reference groups, the hunter-gatherers and the LBK in Central Europe gave a comparable result to the previous F_{st} analysis with the 29POP dataset (Figure 16, Supplementary Table 14). Because the F_{st} values slightly vary with the extent of involved groups or populations in the calculation, it was necessary to recalculate the genetic distances focusing specifically on the Transdanubian dataset. The correction of the p values (with the Benjamin Hochberg method) did not resulted in p value shifts crossing the significance level of 0.05 (Supplementary Table 14b).

Summarizing the F_{st} analysis, there is no significant genetic distance between any pair of Neolithic-Chalcolithic Transdanubian populations, but the distances are significant between the hunter-gatherers and the farmers. The F_{st} values between the Transdanubian datasets are remarkably low. They are even below zero in the following comparisons: LBKT compared VIN, and SOP; SOP compared VIN, and LGY; LBK compared VIN, LBKT, SOP, and LGY (Supplementary

Table 14a). These negative F_{st} values result in zero values on the Slatkin matrix, which fact inhibits a MDS visualisation of the VIN-SOP-LGY populations' genetic connections.

Although BL is slightly detached on the 20POP and 29POP PCA and MDS plots showing positive F_{st} values from the other Transdanubian datasets, the Transdanubian populations do not show a consequent placement pattern on the plots. If an MDS with the Transdanubian cultures is forced after all, the positions of the reduced (*) and complete datasets of the SOP, and LGY populations become so divergent that the plot does not carry meaningful information. Although an MDS plot of the Transdanubian groups would not result in a reliable outcome, I performed an AMOVA with the Transdanubian datasets and the LBK in Central Europe (7POP_2 dataset). I tested 58 arrangements and groupings of the STA, VIN, LBKT (n=39, without Nitra site), LBK, SOP, LGY and BL data (Supplementary Table 16), in order to reveal possible differences among the datasets, which might only be observable in certain arrangements of the groups. HGCN was not involved in this analysis, since its difference from the other groups is so large that it would unbalance the outcome.

Significant p values for the among-group genetic distances were obtained only in two cases, when STA, LGY, BL are in one group and LBK, LBKT, VIN, and SOP in another (among-group percent of variation=0.5, Fct=0.00504, p=0.02861+-0.00144). Clustering the datasets into three groups, the only significant arrangement was when STA, BL were in the first, SOP and LBKT in the second, and VIN-LBK-LGY were in the third group (among-group percent of variation=0.57, Fct=0.00575, p=0.02673). For both constellations, a negative value was obtained for the within-group and among-population percent of variation. These results can either be the consequence of the small sample sizes, or rather the genetic differences are too small for a valid sequence based differentiation.

Since other analyses do not support the separation of the STA from the other Transdanubian datasets, and the genetic distances are generally very small between each Transdanubian population, we should interpret the AMOVA results as stochastic outcomes.

5.1.5.3 Shared haplotype analyses, focusing on the Transdanubian populations

A shared haplotype analysis and an ancestral shared haplotype analysis were performed, focusing on the people of the Transdanubian Neolithic. The HGCN dataset was used as a proxy for the pre-Neolithic genetic substrate in the Carpathian Basin.

The classic shared haplotypes analyses (Excoffier and Lischer, 2010) shows that Transdanubian Neolithic populations share maternal lineages in 40.5-60.9% with each other (on an average of 52.48%) (Supplementary Table 15a). Shared lineages decrease in the Balaton-Lasinja population, but the Starčevo affinity is still high (38.5%).

In the ancestral shared haplotype analysis (ASHA), the populations were considered in a chronological order. The Vinča culture precedes the LKBT in this hypothetical chronological order, because the archaeological records support the Vinča cultural effects on the early LBKT (Marton and Oross, 2010). In reality, the analysed VIN and LBKT samples have roughly the same age, because the aDNA result are from the Late LBK phases, roughly contemporaneous with the studied two Vinča sites.

A general trend can be observed on the ASHA diagram (Figure 21, Supplementary Table 15b). The Starčevo genetic legacy was between 46.3-61.9% in the Middle and Late Neolithic Transdanubia. Additionally, the LBKT, SOP, and LGY populations had new mtDNA lineages in 28.5-31.7%. The Vinča population contained most of the identifiable Central/North European hunter-gatherer lineages (6.45%), but some elements of the Mesolithic substrate also occurred in the STA and in the LGY. Interestingly, hunter-gatherer haplotypes were not found in the LBKT, although it contains hunter-gatherer type haplogroups (U2, U5a). The hunter-gatherer sample set is probably still too small for a reliable estimation of the forager genetic components on haplotype level.

The new genetic components are up to 45.2% of the Vinča lineages, which includes 12.9% identical haplotypes with the LBKT. These LBKT haplotypes are seen, if the VIN-LBKT chronological order is reversed.

The LBKT has a similar composition to the Vinča gene pool, though the identifiable hunter-gatherer components are missing. New LBKT lineages appear in 28.6% of the variation, additionally to the ancestral Starčevo and Vinča lineages. The LBKT has 7.7-10.8% directly observable genetic legacy in the succeeding three Transdanubian cultures' population. This ASHA analysis remarkably reveals the genetic stratification of the Neolithic populations. The descendants of all precursor Transdanubian populations were in some extent incorporated into both the SOP, and the LGY populations. In the meanwhile, each Middle and Late Neolithic population has its own, unique lineages representing about 30% of their maternal genetic variation (Figure 21).

Besides the slightly reduced Starčevo genetic signature, the Balaton-Lasinja carried ancestral LBKT lineages as well. The BL dataset contains neither ancestral VIN, SOP nor LGY

lineages, which can indicate population a genetic event, or can be caused by the biasing effect of the small sample set.

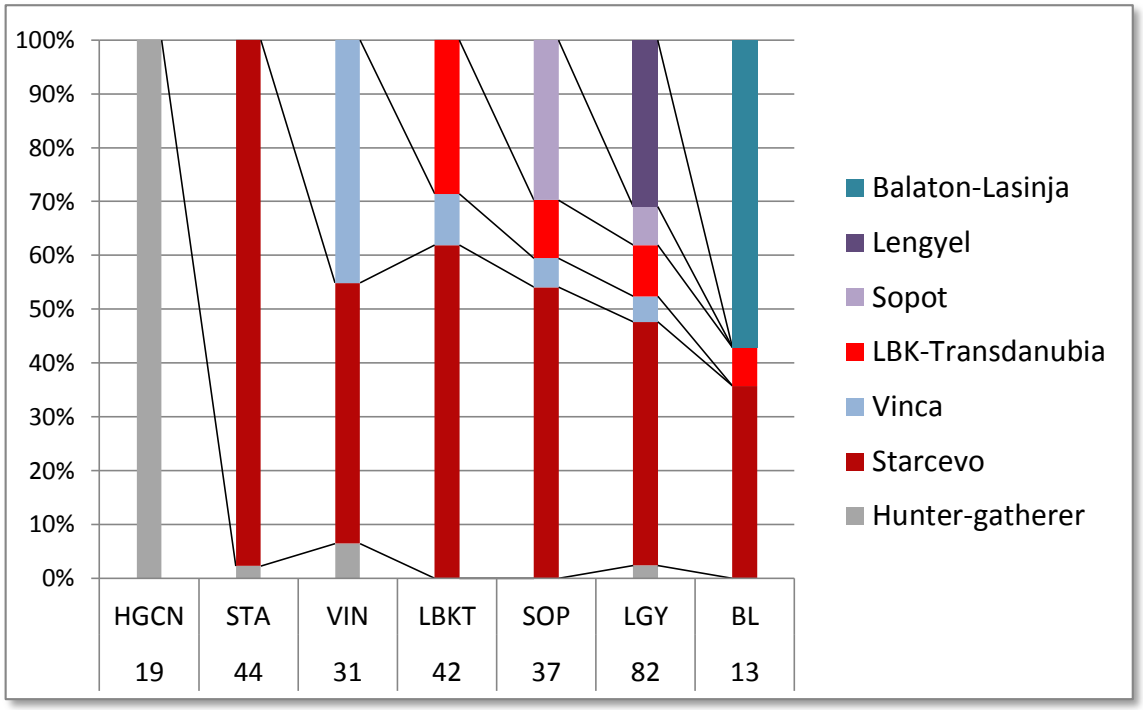


Figure 21. ASHA with the Transdanubian Neolithic and Chalcolithic datasets, comparing to the Central/North European hunter-gatherers.

For abbreviations, see the legend of Table 7. The numbers below the culture abbreviations indicate the sample sizes.

5.1.6 Testing regional genetic differences in the Transdanubian Neolithic datasets

Since the above presented analyses showed predominant homogeneity in the maternal gene pool of the Transdanubian Neolithic populations, I looked for other aspects, which could affect genetic structure in the studied area. One of such causes can be the sedentariness of the farmers, living on a more or less steady place over several generations. Regional differences can be formed in such sedentary societies as well, even in such a small but diverse geographic region as Transdanubia. Regional differences could occur for example, if the populations at the onset of the Neolithic were originated from different substrates or regions, and they did not admixed entirely, or if later infiltration/immigration events did not affected the whole region of Transdanubia.

In order to test this hypothesis, I divided Transdanubia into two sections, following the line of the Lake Balaton (Figure 22). This theoretical line was the approximate boundary of the northern extension of the Starčevo culture. This transect is concordant to the so called "Early Neolithic Agro-Ecological Barrier", which has been localised by P. Sümegi and R. Kertész [a summary of the theory was given by E. Bánffy and P. Sümegi, 2012)]. According to their theory, a barrier existed in the middle zone of the Carpathian Basin between two types of climate variants. These two climate and environment types were differenced in eastern continental and western oceanic influences (Sümegi et al., 2002).

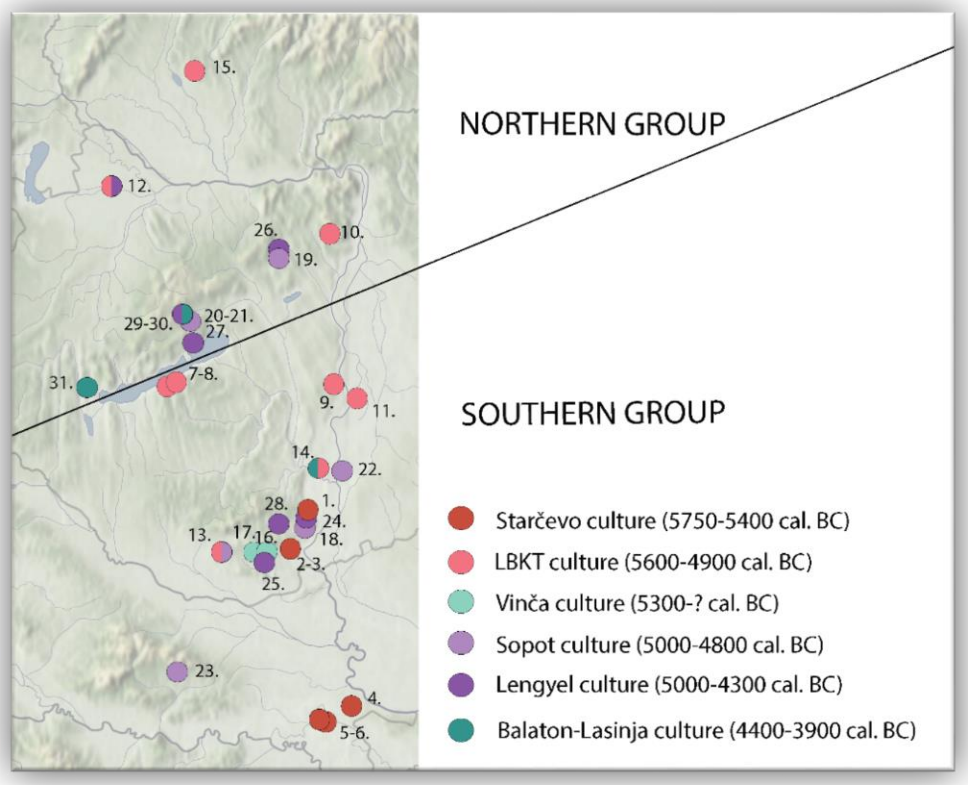


Figure 22. Division of the studied sites according to two regions of Transdanubia: the northern and the southern Transdanubian groups.

Consequently there was a cultural and ecological section in the Early/Middle Neolithic of Transdanubia, which has recently been rather considered as a zone of interaction than a barrier (Bánffy and Sümegi, 2012). Although the early LBKT shows uniform characteristics in Transdanubia, the material culture of the classical LBKT had a section (or overlapping zone) in Middle Transdanubia, dividing the Keszthely and Notenkopf/Zselíz groups. The question is, whether these sections in Transdanubia are reflected in the genetic pattern of the Middle/Late Neolithic population as well.

Based on the suggestion of K. Oross, I assigned the two LBKT sites, Balatonszárszó and Balatonszemes in to the LBKT-south group. They show a mixed characteristic of the southern and northern archaeological traditions, being in the overlapping zone of the Keszthely and Zselíz groups (K. Oross, personal communication). However, the small sample size did not validate a section of third "middle group" of the LBKT.

Looking at the haplogroup distribution of the LBK samples, differences are seen in some extent. Haplogroups J and N1a were detected only in south of the Lake Balaton, whereas the haplogroup U occurred only in north Transdanubia (Figure 23).

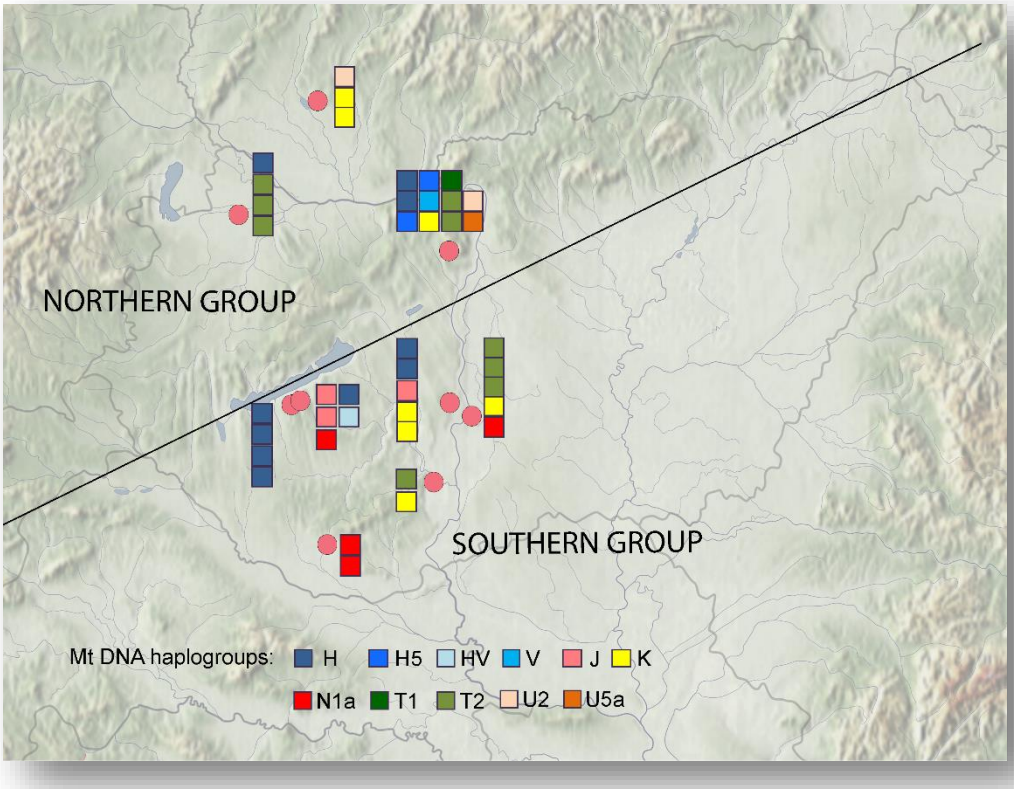


Figure 23. In the LBKT detected mtDNA haplogroups and their distribution in the Carpathian Basin.

In case of the Sopot culture, a sampling bias would occur, if the unequally distributed samples were divided into a south and north Transdanubian group. This imbalance is a state of archaeological research, since our team sampled almost every skeleton assigned to the Sopot culture in Transdanubia. The number of known Sopot burials was doubled by our team, when archaeologist colleagues identified the Alsónyék Sopot cemetery (with 18 graves) in the course of our sampling campaign (T. Marton, A. Osztás, personal comm.). On Figure 24, the geographic distribution of the Sopot mtDNA haplogroups can be observe. N1a haplogroups

were found only in southern Transdanubia. On the other hand, the small northern sample set and the large variety of the Alsónyék haplogroups do not justify any further conclusions about the possible regional differences within the Sopot culture.

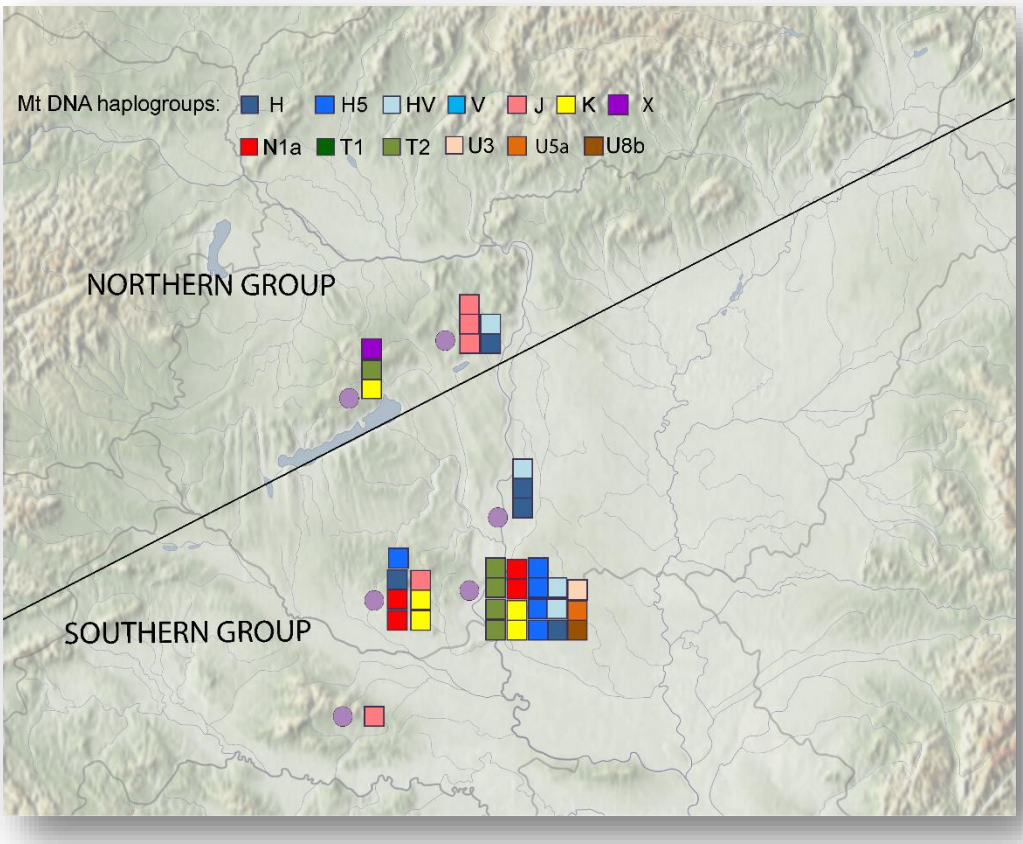


Figure 24. In the Sopot dataset detected mtDNA haplogroups and their distribution in the Carpathian Basin.

According to the archaeological records, the Sopot culture rather forms western/eastern then southern/northern groups in Transdanubia (Regenye, 2002). This means that one should assemble site Nemesvámos with Szemely and Alsónyék with Bicske. Though it would not fit into the tested hypothesis, and the sample number would be still unbalanced, so the Sopot culture remained undivided in the further analyses.

The regionality is more apparent if we look at the haplogroup frequencies of the north-south divided Lengyel culture (Figure 25). I detected comparatively more haplogroups U, H, T2 in the north, and more haplogroups J, K, N1a on the southern Transdanubian sites. Both LGY groups contain equal size (n= 41) of individuals, which distribution lays stress on the genetic differences.

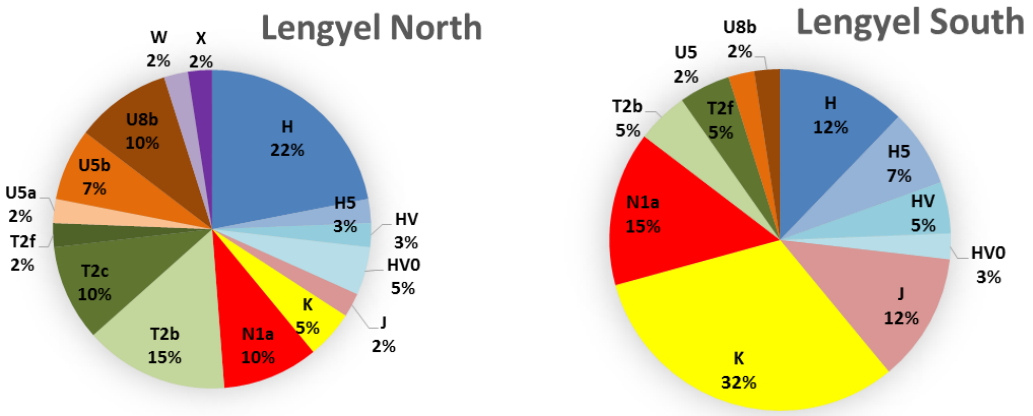


Figure 25. MtDNA haplogroup frequencies of the southern (n=41) and northern (n=41) Transdanubian Lengyel datasets.

In order to verify these primary assumptions, PCA, MDS, F_{st}, and AMOVA analyses were performed with the divided LBK, LGY datasets and the unsectioned STA, VIN, SOP, and BL data. The reason behind this method was that STA and VIN disseminated only to south Transdanubia, and in the BL dataset only three samples originated from south Transdanubia. Therefore, datasets STA, VIN, and SOP were hypothesised as southern groups and BL as northern, based on the studied site locations. Furthermore, I involved the Neolithic mtDNA results of V. Keerl from the Alföld region (eastern Hungary) into the second step of the analyses (Keerl, 2014).

I carried out a PCA with the southern-northern regional groups of the LGY and LBKT datasets, the STA, VIN, SOP, BL groups and the HGCN. I tested the complete datasets, and the reduced sets (*) as well, following similar concept presented in the case of the 20POP dataset. The PCA plot shows (Figure 26), that the maternal relations have no biasing effect on the results; the difference in location is minimal between the complete datasets and the reduced ones.

If the Central/North European hunter-gatherers are used as an “outgroup”, the Transdanubian populations are located in the minus range of the PC1 on the PCA plot (Figure 26, Supplementary Table 17a). The south Transdanubian datasets (oriented by the J, K, N1a, X, and U3 vectors), are close to each other, placed along the positive range of the PC2. The north Transdanubian LBKT and Lengyel groups are in negative ranges of the diagram, along the T1, W, U8, HV, T2, and H haplogroup vectors.

The Sopot sample set shows rather southern characteristic in its haplogroup composition, and it is located in the southern cluster on the PCA plot (Figure 26). The BL dataset fits into the northern cluster, congruently with the starting hypothesis.

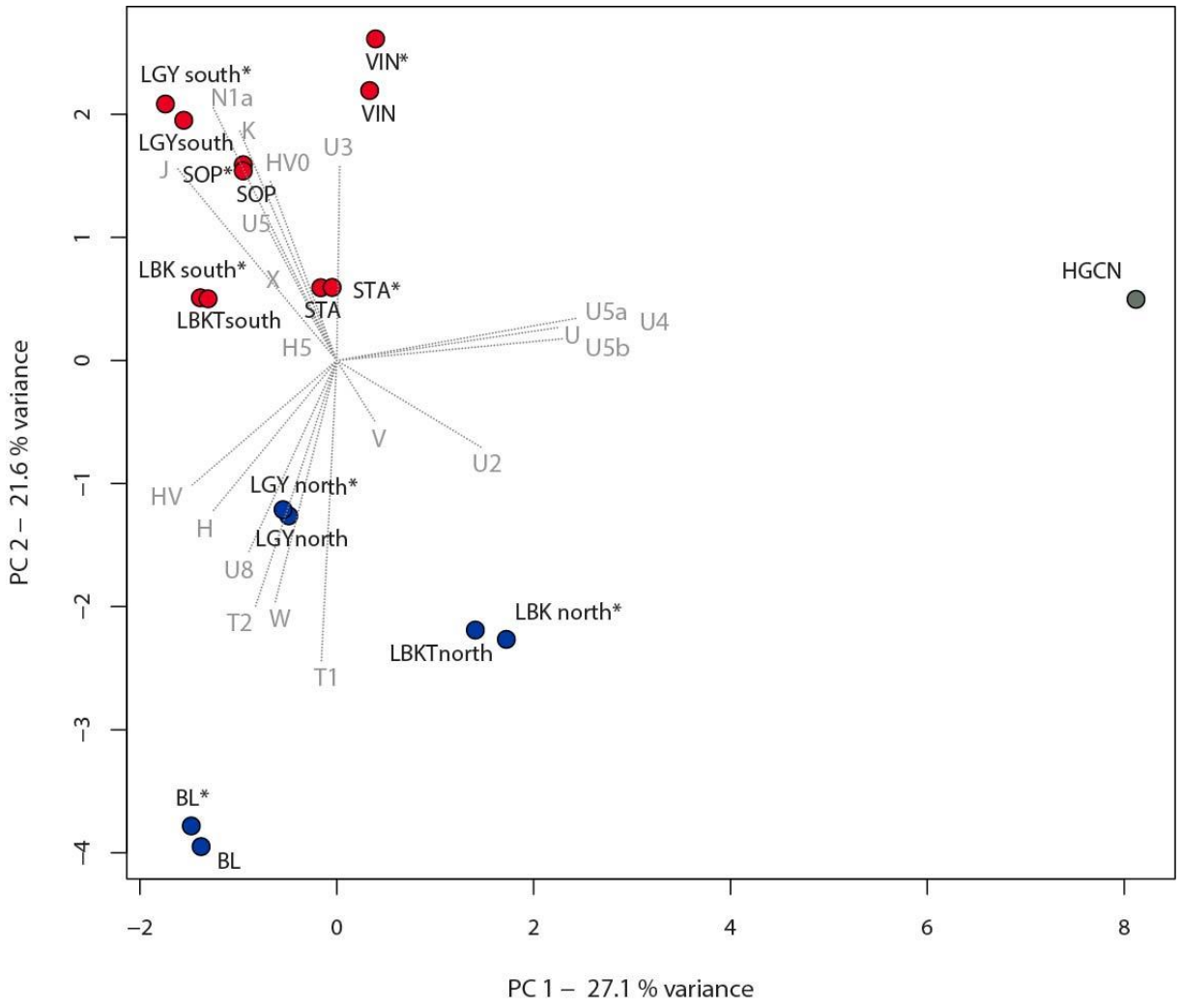


Figure 26. PCA with the Transdanubian populations and the hunter-gatherers from Central and North Europe.

The north Transdanubian groups are marked with blue circles and the south Transdanubian groups are marked with red circles. The reduced datasets (free from maternally related individuals) are marked by an asterisk. The PCA represents 48.7% of the total haplogroup variation.

The MDS analysis was not possible with the hunter-gatherers as an "outgroup", because the stress became nearly zero from this combination of datasets. This could be a result of the fact that the Transdanubian groups are all very similar to each other. In contrast, they are all very different from the hunter-gatherers. In previous MDS plots, the arrangements

were not so dualistic, since the third-second millennia BC cultures from Central Europe formed a third pole on the plots. The north Transdanubian groups line up on the minus range of the first coordinate on the MDS plot focusing on Transdanubia (Figure 28, Supplementary Table 18a). A clustering is not remarkable on the plot, but some of the groups are located in noteworthy distances from each other. The comparison of the south and north Lengyel and south and north LBKT group's results in significant F_{st} values respectively, as the Lengyel-south is in significant distance from the Balaton-Lasinja and from the LBKT-north (Supplementary Table 19).

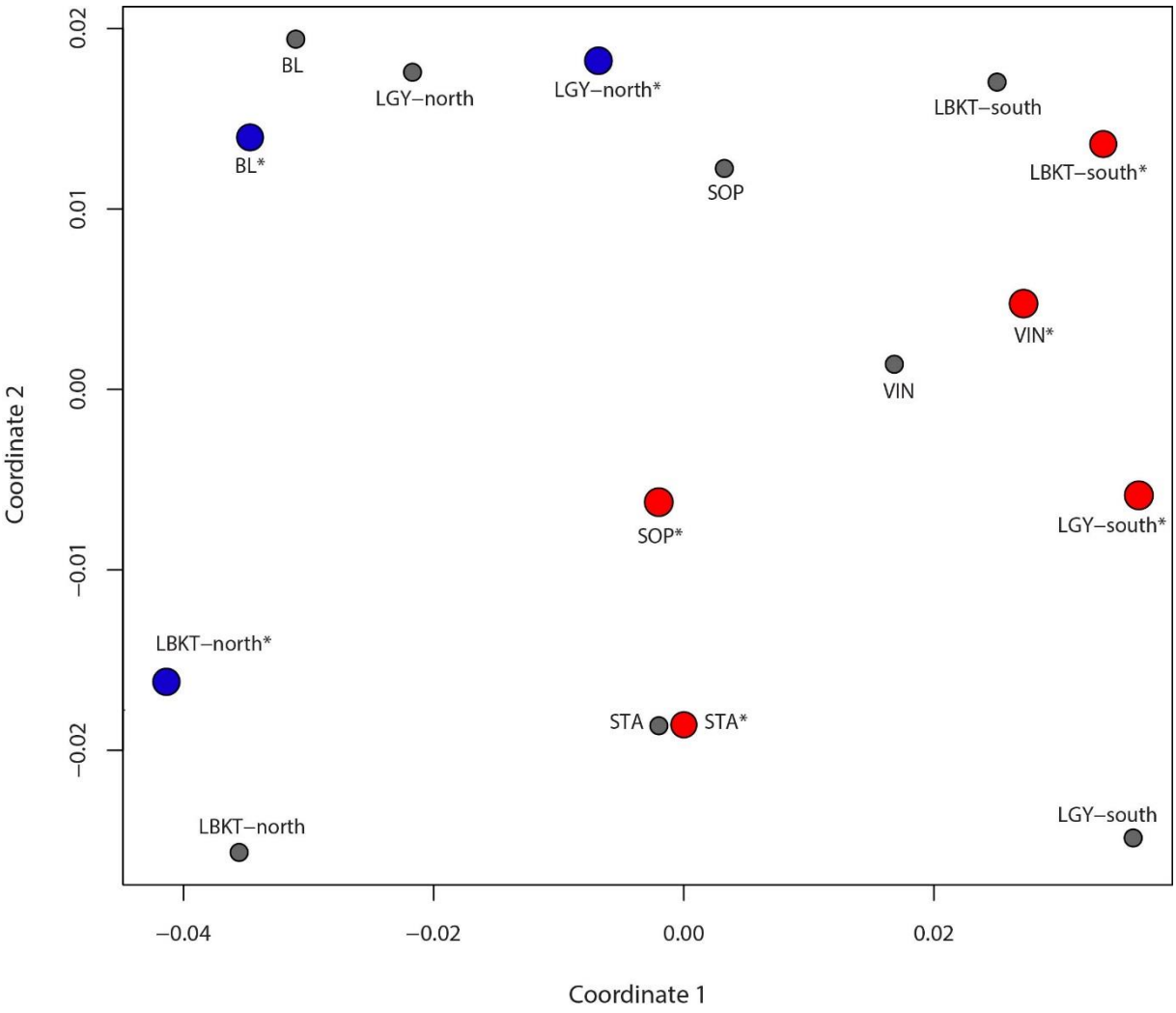


Figure 27. MDS plot with the Transdanubian Neolithic and Chalcolithic regional groups.

Stress value is 0.0999, and non-metric fit (R^2) is 0.99, which means a good fit between the two-dimensional graph and the original distance matrix. The reduced datasets (free from maternally related individuals) are marked by an asterisk and colours: north Transdanubian groups are marked with blue, and the south Transdanubian groups are marked with red colour.

I accomplished an AMOVA analysis, in order to verify the clustering of the Transdanubian groups, which was not distinct on the MDS plot (Figure 27). Nevertheless, I could use this MDS plot as a starting point, because AMOVA and MDS are strongly related analyses, as both are based on the genetic distances, calculated from HVS-I sequences. The best arrangement, with the highest among-group variance, was the clustering of the LGY-north, BL, and LBKT-north into one group and the STA, VIN, LBKT-south, SOP and the LGY-south into a second group (among-group variation is 1.54%, F_{st}=0.01544, p= 0.01802; among population and within groups variation is 0.29%, F_{sc}= 0.00292, p= 0.29515 [see Supplementary Table 20]). Consequently, we can assume that even the MDS plot did not reveal a north-south genetic difference on its own, it did become assured by the AMOVA analysis.

In the next step, I involved the Alföld cultures from eastern Hungary to the analyses of the Carpathian Basin populations/regional groups. The Alföld LBK (ALBK) datasets could be divided into further regional groups, such as ALBK1- Szatmár (which is the earliest ALBK in north Alföld), ALBK-Esztár (eastern Alföld), ALBK-Tiszadob/Bükk (northern eastern Hungary) and ALBK-Szakálhát. This latter one had its formative phase in the southern Alföld, hereafter it disseminated to the central parts of the Great Hungarian Plain as well. The Szakálhát sample set originated mainly from the central part of the Alföld. The dataset of the Tisza culture could not be divided into north and south groups either. The only northern Alföld site was Pusztataskony-Ledence within the Tisza dataset.

The north-south division of the LBK and Lengyel cultures becomes apparent on the PCA with the Carpathian Basin populations too (Figure 28, Supplementary Table 17b). The red cluster in the negative range of PC2 contains mostly southern groups, characterised by haplogroups K, N1a, H5, V, J, X, U3, and U5. The two Early Neolithic cultures, Körös and Starčevo are in close proximity to each other (similarly to the 3D PCA plots, data not shown), which is remarkable, regarding that only 16 individuals from the Körös culture were compared to 44 from the Starčevo. From the Alföld region the ALBK-Esztár and the ALBK-Szatmár joins to this cluster.

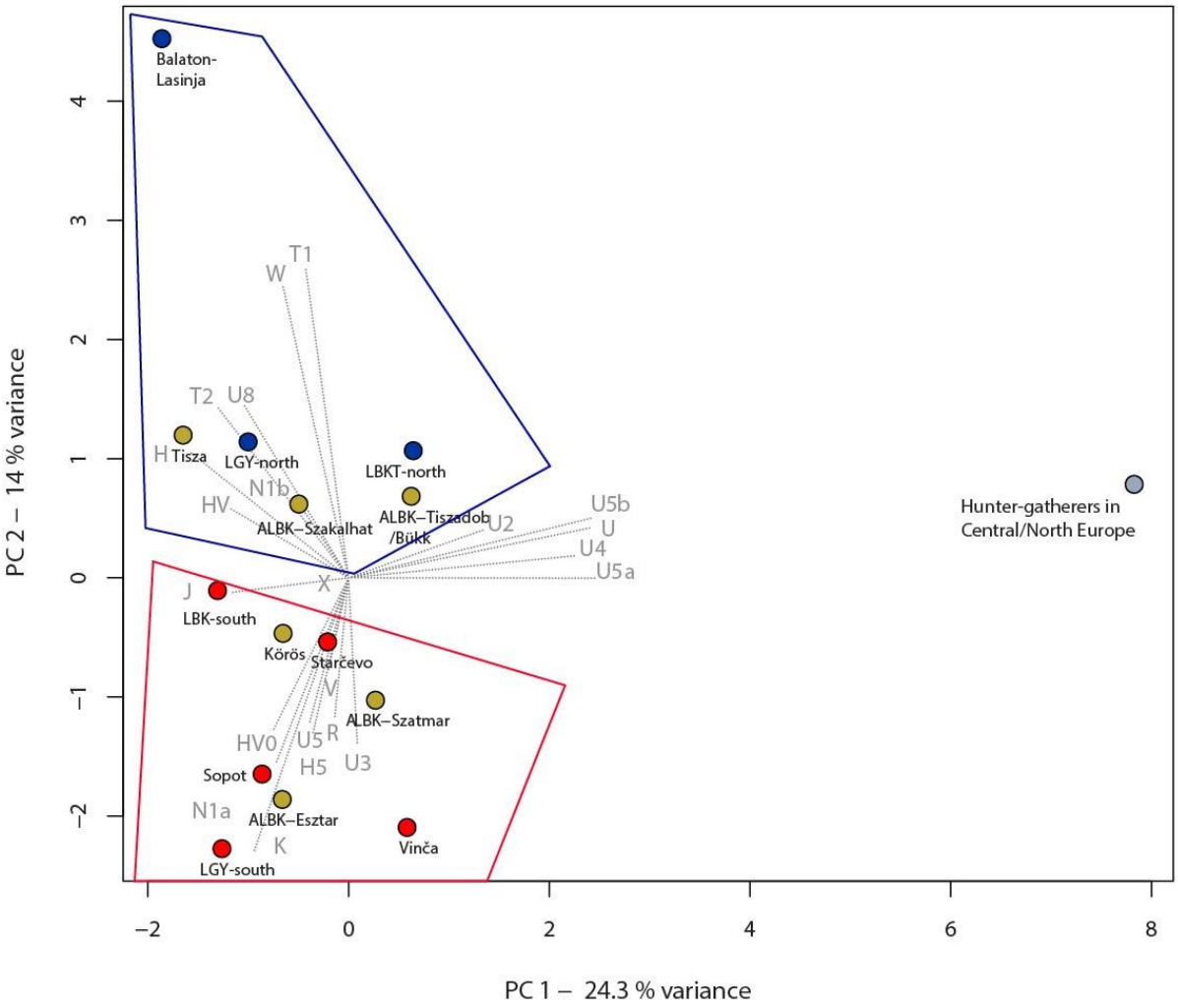


Figure 28. PCA with the regional groups of Transdanubia and the cultural groups from the Alföld.

Alföld mtDNA data were taken from the PhD thesis of V. Keerl, 2014.

On the positive range along the PC2, ALBK-Szakálhát, ALBK-Tiszadob/Bükk, Tisza, Lengyel-north, and LBK-north are situated towards the BL, which is separated from all other

populations of the plot. The Tisza samples, Szakálhát samples, and the ALBK-Tiszadob/Bükk data are located within the blue barrier. Haplogroups W, T1, T2, U8, and H characterise this cluster in general. The three Alföld datasets have highly similar frequency of haplogroups T2, K, H, U5b, but major differences are also observable, considering the X and J haplogroup rates. Due to an elevated frequency of haplogroups U, the LBKT-north and the ALBK-Tiszadob/Bükk are slightly shifted toward the hunter-gatherers.

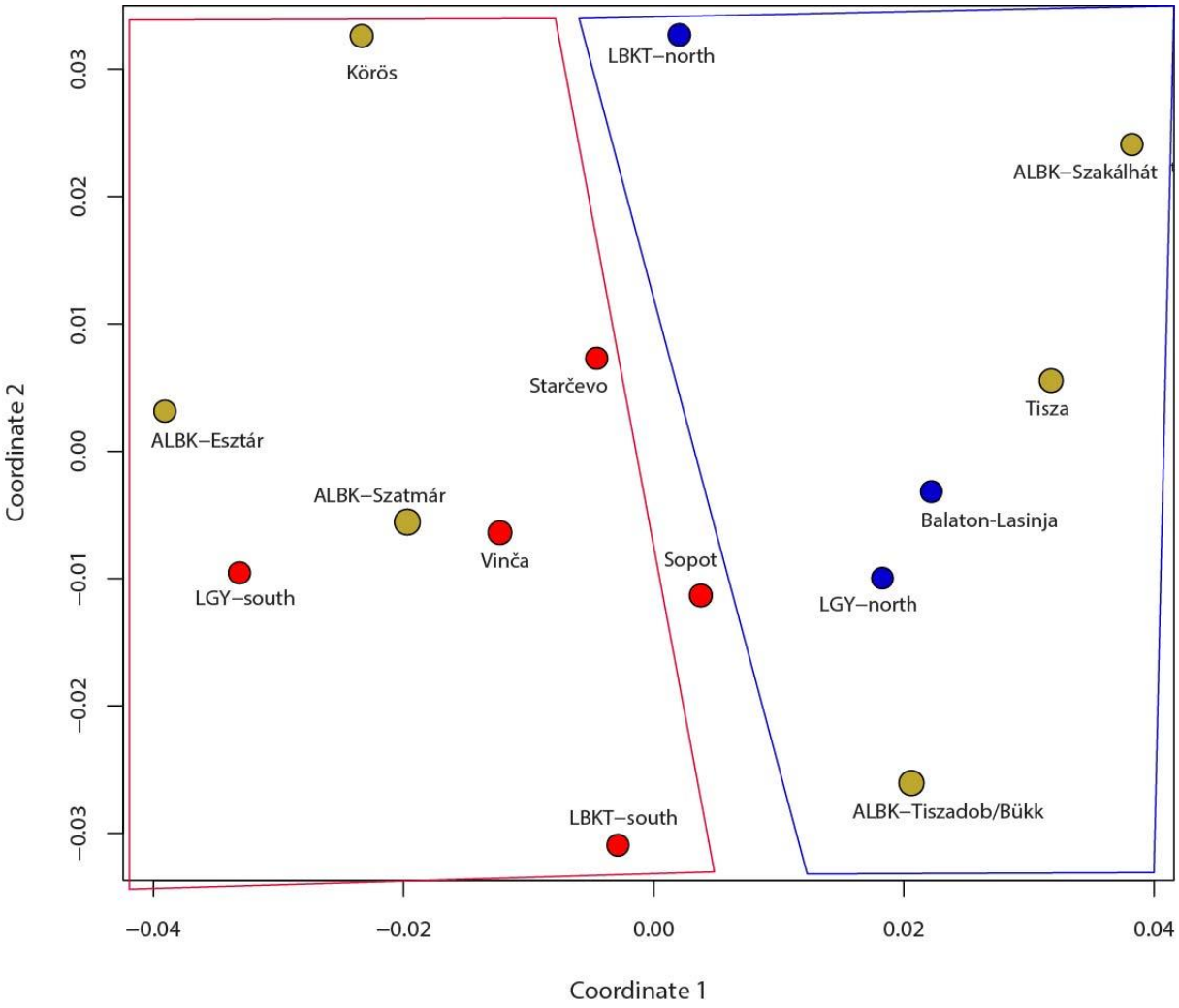


Figure 29. MDS plot with the regional groups of the Transdanubian Neolithic and the Alföld populations (Alföld mtDNA results are taken from Keerl 2014).

Stress value is 0.1362, and non-metric fit (R^2) is 0.981, which means a good fit between the two-dimensional graph and the original distance matrix. The F_{st} and p values of this dataset are presented in Supplementary Table 21, Slatkin matrix in Supplementary Table 18b.

The hunter-gatherers were served as a reference point on the PCA, giving again an orientation point for the other datasets presented on the diagram (Figure 28). The similar use of the hunter-gatherer dataset in the MDS analysis was not possible, since it gave again a

special constellation of the studied datasets, which resulted in a nearly zero stress during the distance calculation in R. The MDS presented on Figure 29 had a stress value of 0.1362.

One can also transect the MDS plot, separating two clusters from each other, which are composed of almost the same datasets as on the PCA plot (Figure 28). However, this arrangement is seemingly not obvious, similarly to the clustering on the previous MDS plot, where the hunter-gatherers were also omitted (Figure 27). The hypothesis of this clustering can be rejected or supported by an AMOVA analysis. Testing different arrangements of the groups/cultures along the drawn diagonal transection (SOP, STA, LBKT-north and south), it can be inferred that the best arrangement is similar to that we see on the MDS plot (Figure 29). Grouping the SOP into the blue (northern) section, results in the highest among-group variation of 1.92% [with an F_{st} value of 0.01918, p value=0.005 (for details see Supplementary Table 22)]. This result shows the uncertain position of the regionally mixed Sopot dataset, and the north-south genetic difference within the LGY and LBKT.

Summarizing the regionality tests, the samples are not equally distributed from the different regions of Transdanubia, which issue biases the analyses. The single balanced dataset is the LGY, with 41-41 samples from the north and south Transdanubia respectively. The haplogroup and haplotype composition of the two LGY groups shows significant differences, signalling a genetic structure among the different regions of Transdanubia. The LBKT shows as well a north-south genetic section or structure, but a more robust dataset should be needed for its reassurance. The SOP and BL datasets were not adequate for a north–south division, due to the unequally distributed origins of the samples. In the PCA and MDS with a Transdanubian focus, both datasets were considered per regional group/population. The exclusion of possible maternal relatives did not influence the PCA (Figure 27), though it did shifted the positions of the SOP and LGY-south populations on the MDS plot remarkably (Figure 28).

5.1.7 Population genetic analyses, comparing the prehistoric datasets to modern Eurasian and African mtDNA data

5.1.7.1 Principal component analyses

MtDNA PCAs were performed with the prehistoric datasets and 73 modern populations, compiled from literature data by G. Brandt and our team (see chapter 4.3.6, Brandt et al., 2013). I carried out six PCAs, comparing each culture individually, because the six Transdanubian cultures have a unique haplogroup composition in relation to the modern populations. Consequently, a simultaneous analysis would result in a separate prehistoric cluster on the PCA plot.

The frequencies of haplogroups N1a, T1, T2, K, J, U3, X, HV, and the absence of Asian and African lineages in the STA, locate the early farmers close to populations of the modern Near East and the Caucasus (Supplementary Figure 3a). The LBKT falls closer to South and southeast Europe (Greeks, Bulgarians and Italians) than the STA, which is caused probably by a higher frequency of haplogroup H, the occurrence of U2, and the absence of U3 and X in the LBKT (Supplementary Figure 3b). The Vinča dataset falls between the Caucasus and South Europe, but somewhat further from the modern populations, toward the minus range along the first and second components (Supplementary Figure 3c). The Sopot and the Lengyel cultures show more similarity to South and southeastern Europe, they are the closest to the Greeks, Bulgarians, and Italians (Supplementary Figure 3d-e). The Balaton-Lasinja culture on the other hand shows an elevated affinity to the Near East and Caucasus again, it falls virtually to the same place as the STA dataset (Supplementary Figure 3f).

Three dimension PCAs were also performed with the STA and LBKT (see in Szécsényi-Nagy et al., 2014a). The high frequencies of haplogroup N1a, T2, and K are the most characteristic for the Carpathian Basin cultures, which all have high loadings on the third component, shifting the prehistoric datasets away along PC3 from all modern populations.

5.1.7.2 Genetic distance mapping

The HVS-I sequence-based genetic distance maps are largely consistent with PCAs and reveal the greatest similarities of the STA to populations of the modern Near East (e.g. Iraq, Syria) and the Caucasus (Azerbaijan, Georgia, Armenia), as well as some European populations, such as Italy, Austria, Romania, and Macedonia (Figure 31).

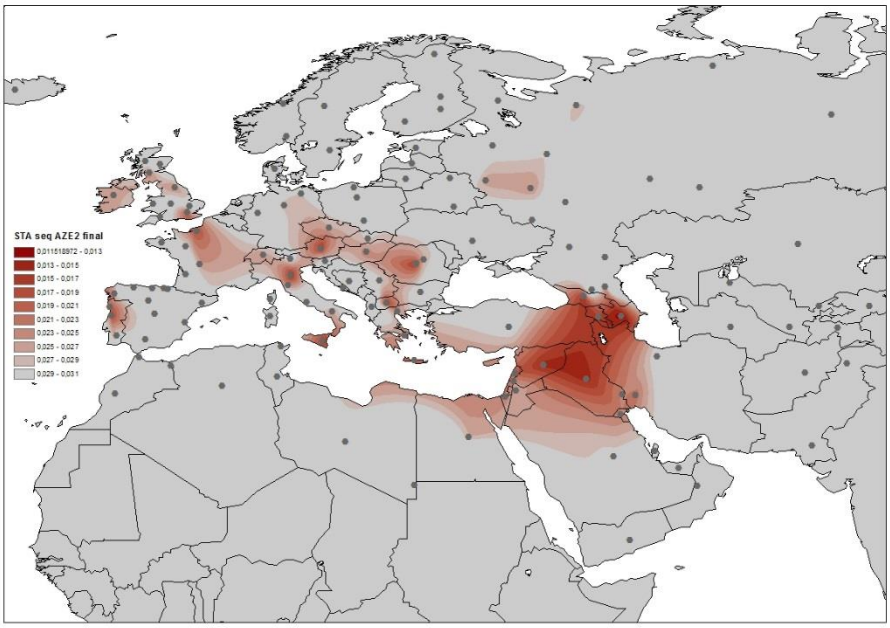


Figure 30. Genetic distance mapping with the Starčevo dataset (STA).

Genetic distances (F_{st}) between the STA and 130 present-day populations of Eurasia and North Africa were computed based on HVS-I sequences and visualized on a geographic map. Grey dots denote the location of modern populations. Colour shadings indicate the degree of similarity or dissimilarity of the Neolithic cultures to these populations. Short distances and great similarities are marked by dark red areas. F_{st} values are scaled by an interval range of 0.002. F_{st} values higher than 0.029 were not differentiated (grey areas). Population information and F_{st} values are listed in Supplementary Table 23.

The distance map of the LBKT displays affinities that are overall similar to the STA, which includes populations from Azerbaijan, Syria, and Iraq. Similarities to present-day Europeans are also observable on the GDM, such as the populations of Great Britain, Portugal, Romania, Crete, and Russia (Figure 31). These similarity peaks are likely explained by elevated frequencies of shared lineages due to shared genetic drift in modern-day populations.

The genetic distance maps of the Sopot and Lengyel culture (Figure 33, Figure 32) show a general and overall similarity to the present day Europeans. The smallest distances between the Sopot dataset and the modern populations were found in case of the Hungarians, Portuguese, Romanians, British populations, and the Scots. The most similar modern populations to the Lengyel dataset are from Syria, France, Italy, Portugal, and Great Britain.

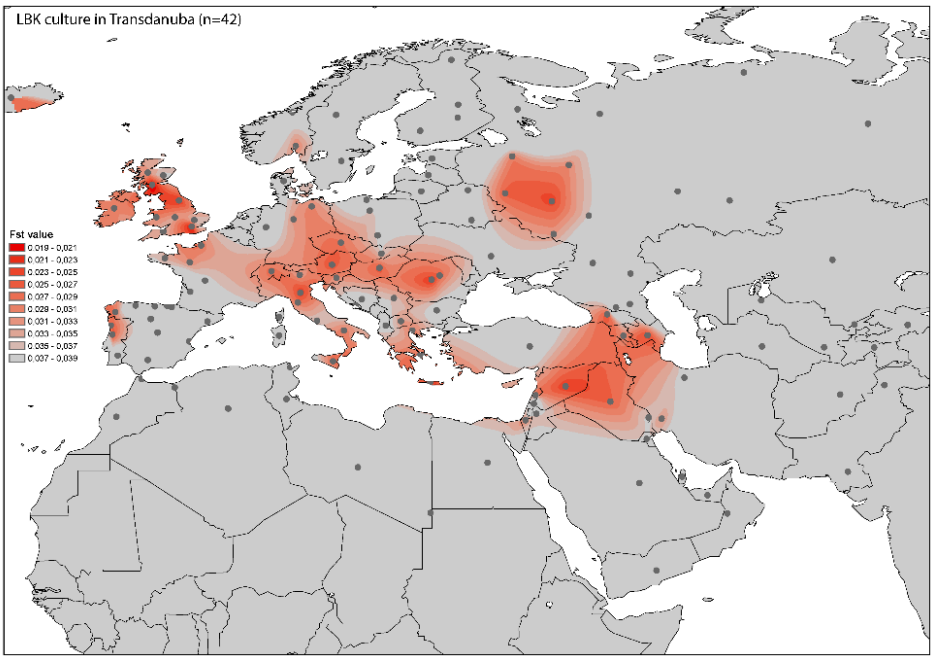


Figure 31. Genetic distance mapping with the LBK dataset in Transdanubia (LBKT).

Short distances and great similarities are marked by orange areas. F_{st} values are scaled by an interval range of 0.002. F_{st} values higher than 0.039 were not differentiated (grey areas). Population information and F_{st} values are listed in Supplementary Table 23**.**

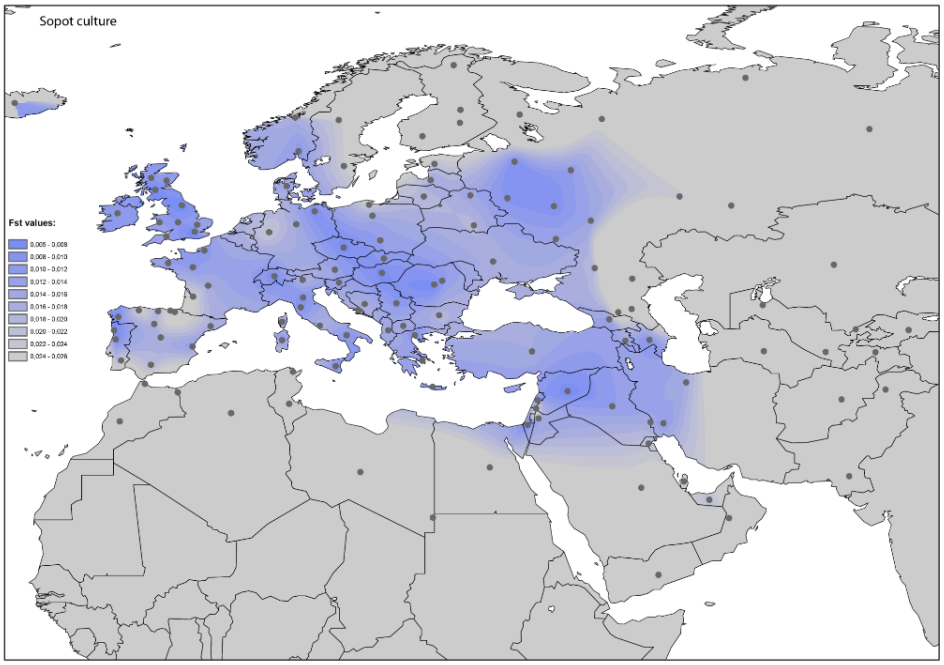


Figure 33. Genetic distance mapping of the Sopot dataset (SOP).

Short distances and great similarities are marked by blue areas. F_{st} values are scaled by an interval range of 0.002. F_{st} values higher than 0.026 were not differentiated (grey areas). Population information and F_{st} values are listed in Supplementary Table 23.

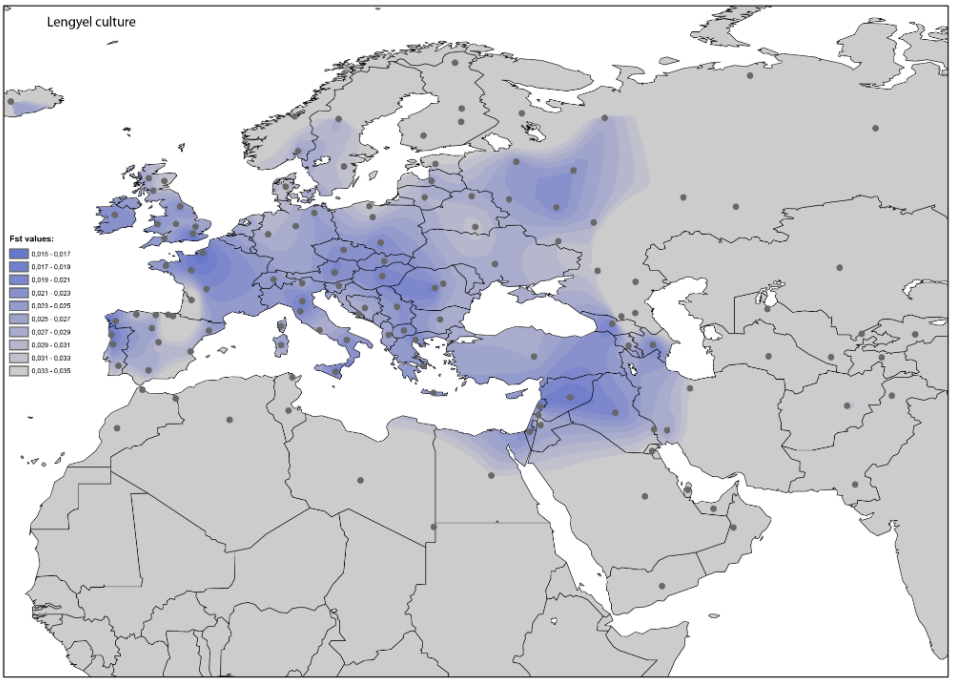


Figure 32. Genetic distance mapping with the Lengyel dataset (LGY).

Short distances and great similarities are marked by violet areas. F_{st} values are scaled by an interval range of 0.002. F_{st} values higher than 0.035 were not differentiated (grey areas). Population information and F_{st} values are listed in Supplementary Table 23**.**

The people of the Early Chalcolithic Balaton-Lasinja culture show an elevated affinity to the modern Near East (Figure 34), similarly to the PCA result (Supplementary Figure 3f). Populations of Iraq, Azerbaijan, Kuwait, Armenia, and the Chechen Republic are the closest to the Balaton-Lasinja dataset. These similarities can indicate a new immigration wave originating from the Near East. Nevertheless, we have to be cautious about the small sample size of the Balaton-Lasinja culture, because it can cause accidental affinities.

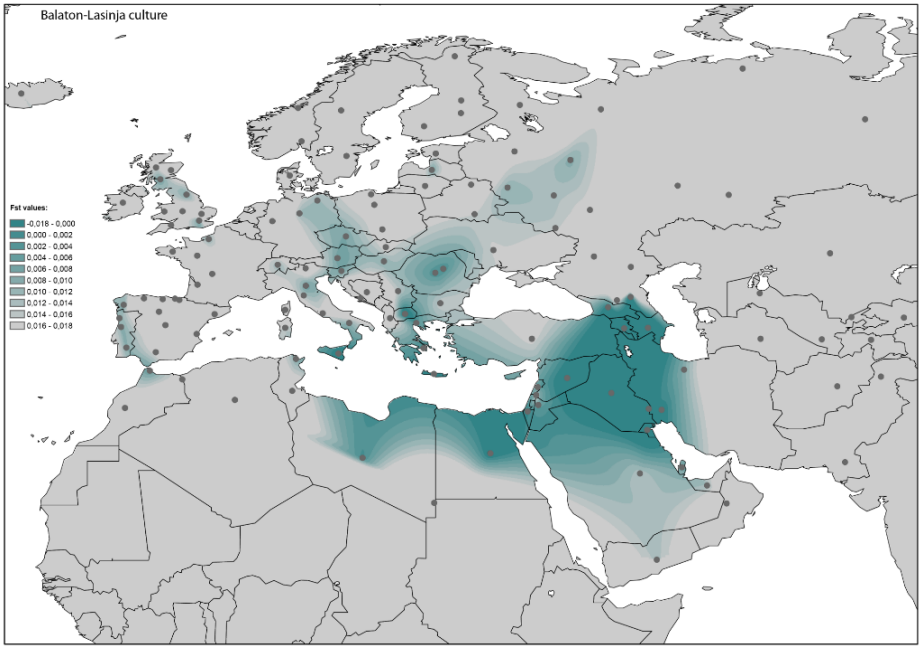


Figure 34. Genetic distance mapping of the Balaton-Lasinja culture (BL).

Short distances and great similarities are marked by dark green areas. F_{st} values are scaled by an interval range of 0.002. F_{st} values higher than 0.018 were not differentiated (grey areas). Population information and F_{st} values are listed in Supplementary Table 23**.**

5.1.8 Shared haplotype evidence, comparing the prehistoric Carpathian Basin with modern age mtDNA data

The SHA method (Excoffier and Lischer, 2010; Haak et al., 2010) was used, in order to compare the ancient STA and LBKT samples with mtDNA haplotype variation of modern populations. The aim of this investigation was to reveal the HVS-I based haplotype level background of the affinities, observed on the genetic distance maps. I compared the STA and LBKT haplotypes with the very same 130 populations, which were used for the F_{st} calculation in the GDM analyses. Out of the 32 different STA haplotypes, 14 were found to be Eurasia-wide common today. Another ten haplotypes can be regarded as "informative", since they occur only scarcely in some of the North-African and Eurasian populations. These haplotypes belong to the haplogroups T2e, T2c, K, V6, X2, N1a1, W, and HV0 (Table 9). The informative haplotypes are shared in a relatively high frequency (3-5%) with populations from Libya, Russia (Arkhangelsk Oblast), Spain (Valencia), Caucasus (Tatars) and Romania (Seklers) (Supplementary Figure 4). Counting the percentage of informative matches among all shared haplotypes (as it was done by Haak et al., 2010), I observed over 10% of shared haplotypes in the populations of north Libya, Russia (Arkhangelsk Oblast), Romania (Seklers), north Pakistan, Syria, Tatars, Valencia and Ossetia.

Further eight haplotypes are unique Starčevo ones, which do not have matches with the 36,840 analysed modern HVS-I haplotypes. These haplotypes belong to the haplogroups T2f, K, J, X2, and N1a1.

In the LBKT, I differentiate 23 haplotypes or lineages. Eight are "informative" (belonging to haplogroups H, T2, N1a, and U5a) and three are unique (haplogroups U2, N1a, and K) among them. The informative haplotypes have a high frequency (4.5-3%) among populations from Russia (Arkhangelsk Oblast), Spain (Valencia), Romania (Seklers), Libya and the Caucasus (Tatars).

Summarizing, this method can help interpreting the Neolithic haplotypes, but it does not prove to be useful when we are looking for the origin or modern Eurasian genetic legacy of these early farmers.

Culture		Samples	haplotype (reference is rCRS)	haplogroup
STA	informative sequences	BAM5	16126c 16153a 16294t 16296t	T2e
		M6-116.9	16126c 16292t 16294t	T2c
		BAM2	16172c 16224c 16311c	K
		VINK5, BAM9. LGCS4	16093c 16189c 16224c 16311c	K
		VINJ2	16093c 16224c 16261t 16311c	K1a
		VINJ4	16162g 16298c	V6
		BAM15	16189c 16223t 16278t 16362c	X2
		BAM22	16086c 16147A 16172c 16223t 16248t 16320t 16355t	N1a1a
		LGCS1	16093c 16223t 16292t	W
		VINK3	16298c	HV0
	unique sequences	BAM8	16093c 16126c 16189c 16294t 16296t 16304c	T2f
		M6-116.4	16166g 16224c 16311c	K
		BAM7	16093c 16224c 16225.1c 16311c	K1a
		VINK2	16069t 16126c 16224c 16261t	J1c
		VUKG1,VUKG3	16069t 16126c 16302g	J
		BAM23	16179t 16189c 16223t 16278t 16362c	X2
		BAM25	16147A 16172c 16193t 16223t 16248t 16355t	N1a1
		LGCS2	16147A 16172c 16189c 16223t 16248t 16274a 16355t	N1a1
LBKT	informative sequences	BUD 15	16192t 16304c	H
		BAB 4	16320t	H
		HARG3	16126c 16292t 16294t 16296t	T2b
		BUD 3, TOLM3	16126c 16153a 16294t 16296t	T2e
		KON 4, KON5	16126c 16147t 16294t 16296t 16297c 16304c	T2b
		HARG2	16086c 16147A 16172c 16223t 16248t 16320t 16355t	N1a1a
		BSz5, SZEH9	16147A 16154c 16172c 16223t 16248t 16320t 16355t	N1a1a3
		BUD 1	16192t 16249c 16256t 16270t 16399g	U5a1b
	unique sequences	BUD 9	16051g 16092c 16179t 16274a	U2
		SZEH4	16147A 16154c 16172c 16223t 16248t 16300g 16320t 16355t	N1a1a3
		HARG4	16224c 16311c 16398a	K

Table 9. Informative and unique haplotypes in the STA and LBKT datasets.

The shared haplotypes were searched in the maternal gene pool of 130 modern populations.

	n	STA	VIN	LBKT	SOP	LGY	BL	Lom-bards	Con-quest HUN	Cuma-nians	modern HUN
Lombards	28	25	0	3.57	0	0	3.57	67.86			
Conquest HUN	25	28	0	0	0	0	0	0	72		
Cumanian	11	54.5	0	0	0	0	0	0	0	45.45	
modern HUN	284	31.7	0.7	2.11	0	2.11	1.76	1.41	0	0.35	59.86

Table 8. Ancestral shared haplotypes with the Lombards, Conquest period Hungarians, Cumanians and modern Hungarians.

The numbers are given in percent (%).

A further possibility is to study ASHA in a diachronic way up to modern ages. For this analysis I used the recent Lombard mtDNA results from the Transdanubian Szólád site (Alt et al., 2014), the 25 thus far published Hungarian Conquest/early Árpád period (10-11th century) data (Tömöry et al., 2007), the 11 studied Cumanians from the Alföld (Bogácsi-Szabó et al., 2005), and the two available modern Hungarian datasets (Irwin et al., 2007; Tömöry et al., 2007). The ASHA reveals that HVS-I variability in 32.14% of the Lombards, 28% of the Conquest

period Hungarians, 54.5% of the Cumanians, and 38.38% of the modern Hungarians can be derived from the Neolithic (Table 9). These matches are mainly basal haplotypes of K, J, H and T2 haplogroups, which cannot be further resolved into sub-clusters through the HVS-I region. The Cumanians are represented in a small number, which can cause bias in the ASHA. Nevertheless, calculating an average, 34.8% of the (early) medieval haplotypes match with the haplotypes of the Early Neolithic people (STA).

5.2 Y-chromosomal results

After screening the samples with the amplification of the mtDNA, I also tested the preservation of the Y-chromosome on the complete sample set. Because the anthropological sex data were not complete for the whole series, and the sex of infants (children) are not surely determinable by osteological analyses, I involved all individuals (n=298) into the analyses. I used the GenoY25 SNP typing multiplex system (Haak et al., 2010), and completed it by several singleplex PCRs (see in Methods, chapter 4.1.3.2).

I managed to define Y-chromosomal haplogroups of 32 individuals (Table 10). NRY haplogroups F*, G2a, and I2a are present in the Early Neolithic Starčevo dataset. Only one G2a2b haplogroup was reproduced from the succeeding Vinča population. In the LBKT population, one G2a2b and one I1 were detected.

	n	G2a (P15)	G2a2b (L126)	F* (M89)	I2 (M438)	I2a1 (P37.2)	I2a2 (M223)	I1 (M253)	E1b1b1a 1 (M78)	J2 (M172)	C (M216)	R1b (M343)
Starčevo	7	1	2	3	-	1	-	-	-	-	-	-
Vinča	1	1	-	-	-	-	-	-	-	-	-	-
LBKT	2	-	1	-	-	-	-	1	-	-	-	-
Sopot	5	-	-	-	1	1	-	-	1	1	1	-
Lengyel	13	3	-	6	1	-	-	-	1	1	1	-
Balaton-Lasinja	2	-	-	1	-	-	-	-	-	-	1	-
Bronze Age	3	-	-	-	-	-	1	-	-	-	-	2

Table 10. Summary of the detected Y-chromosomal haplogroups.

Absolute haplogroup frequencies are given in this table. Each haplogroup, defined by the terminal SNP in bracket, is named after ISOGG Y-haplogroup tree version 9.

The G and F haplogroups were not found in the Sopot dataset, but I2-I2a1 reappeared in the Late Neolithic, and three new haplogroups occurred in this cultures well: E1b1b1a1, J2 and C. These three novel haplogroups were also present in the Lengyel dataset, besides the persisting Early Neolithic haplogroups, such as F* and G2a. In the small Balaton-Lasinja sample set, I detected one F* and one C haplogroup. I also had some Bronze Age samples, which were processed as Neolithic samples, before dating by ^{14}C analysis. New haplogroups, one R1b (M343), one R1b (M269), and one I2a2 (M223) were found among them (Supplementary Table 5).

Furthermore, incomplete SNP profiles of nine specimens potentially belong to the same haplogroups: three G2a2b (S126), two G2a (P15), one I (M170) were detected in the STA dataset, one G2a2b (S126), one F* (M89) in the LBKT, and one G2a in the LGY were reproduced insufficiently. These partial results further support the dominance of the haplogroup G2a in the Early Neolithic Carpathian Basin. Nevertheless, these haplogroups were excluded from further population genetic analyses because of their unreproduced profile.

5.2.1 Detected Y-chromosomal haplogroups and their frequencies in the prehistoric and modern Europe

The observed similarities of the STA to LBKT are founded on the presence of the haplogroup G2a2b, which is noteworthy rare in present-day Europe. The modern prevalence of haplogroup G and its subgroups slightly increase towards the Near East and reaches the highest frequency in populations of the South and Northwest Caucasus (Balanovsky et al., 2011; Yunusbayev et al., 2012). Population genetic studies have discussed haplogroup G (Behar et al., 2004; Semino et al., 2000) and its subgroup G2a as potential representatives of the Neolithic demic diffusion from the Near East to Europe (Battaglia et al., 2009). This theory has been supported recently by Neolithic data from northern Spain (Lacan et al., 2011b) and southern France (Lacan et al., 2011a), which attested G2a a pivotal role in the expansion of farming over the Mediterranean route. Intriguingly, haplogroup G2a2b has also been described from the LBK in central Germany (Haak et al., 2010). Furthermore, G2a (G2a4-L91) has also been reported from the Tyrolean Iceman (Keller et al., 2012). All these findings suggest that subhaplogroups of G2a were frequent in Neolithic populations of the sixth-fourth millennia BC across Europe, presumably indicating a common Y-chromosome signature of

early farming communities covering a vast territory including the Carpathian Basin, Central Europe, and the Iberian Peninsula.

The second early Neolithic haplogroup was paragroup F*, found in the STA and in the LBK (Haak et al., 2010). It shows an infrequent and diffuse dissemination pattern in Western Eurasia, which might be the result of insufficient haplogroup resolution in some early NRY studies (Nasidze et al., 2004; Wells et al., 2001). Paragroup F* was observed primarily on the Indian subcontinent and Indonesia at low frequency (Karafet et al., 2005; Kivisild et al., 2003; Sengupta et al., 2006; Zerjal et al., 2007). F* has a particular role in studying the peopling of South Asia. Based on the diversification of paragroup F*, theories have been set up about the coastal early human migration out of Africa toward India (Kivisild et al., 2003). The LBK data of W. Haak et al. and my results show that F* might have originated in Southwestern Asia, or that the F* paragroup has diversified shortly after the out of Africa event, and it is not an indigenous Indian paragroup as it was claimed to be (Zhao et al., 2009).

The prehistoric presence of F* is difficult to interpret through the modern NRY data. It might represent a type that has become extinct or significantly decreased during the last seven millennia. Although this haplogroup is not very diverse, further analyses would be needed, in order to test for the subhaplogroups of F (F1-2) in the ancient samples. The subgroups of F have not been well studied, F1 is mainly known from the Indian subcontinent [Sri Lanka, (Karafet et al., 2008)], and F2 is from China [Lahu tribe, (Sengupta et al., 2006)]. Other subgroups (F-P96 [F3], and F-P254 [F4]) have an uncertain phylogenetic position, and have been recently reordered (www.isogg.org). The NRY haplogroup H belongs to the F* paragroup, which was not screened by the GenoY25 multiplex assay. I designed a singleplex PCR for the typing of the M69 position, which defined the H haplogroup on the ISOGG Y-tree in 2012. The Y tree has been recently restructured, and M69 has become a marker of H1. One Starčevo sample (BAM25) has been ordered to the H2 haplogroup recently, based on deep sequencing data (Haak et al., 2015). Consequently, the Neolithic NRY diversity of the F* haplogroups is difficult to presume, further SNP typing or rather whole Y-chromosomal analyses would help in this issue.

Haplogroups I1 (M253) and I2a1 (P37.2) are most prevalent in present-day populations of Europe with highest frequencies in Scandinavian (Karlsson et al., 2006; Lappalainen et al., 2008; Rootsi et al., 2004) and southeastern European populations (Rootsi et al., 2004).

Population genetic studies of modern-day European's have proposed that both I haplogroups were present in Europe since the Late Upper Palaeolithic (Pericić et al., 2005; Rootsi et al., 2004) and expanded after the Last Glacial Maximum from the Franco-Cantabrian (I1) and a southeast European glacial refuge (I2a1) (Rootsi et al., 2004).

Interestingly, haplogroup I1 (M253) has been hitherto undetected in aDNA studies, while haplogroup I2a1 has been reported from Central and North European Mesolithic (Lazaridis et al., 2014), and Neolithic of southern France and northern Spain (Lacan et al., 2011a, 2011b). A recent study has detected I2a in the Alföld region from the Early Neolithic Körös culture and the Late Neolithic Lengyel culture (Gamba et al., 2014).

Three new haplogroups appeared at the turn of the Middle/ Late Neolithic in Transdanubia, the E1b1b1a1 (M78), the C (M216) and the J2 (M172). E1b1b1a1 has only one prehistoric parallel from the Iberian Epicardial culture (Lacan et al., 2011b). Modern genetic studies reveal a northeastern African origin of this subclade, which has a moderate frequency in Europe, Western Asia and a higher frequency (up to 50%) in North and East Africa. Its occurrence is the highest in the Mediterranean (up to 32% in Albania) in Europe (Cruciani et al., 2007).

Haplogroup C has been characterized in modern population genetic studies as a very diverse clade. It mainly occurs in South Asia, but it also has branches in East Asia and America. The recently described sub-branch C1a2 (or formerly C7 with a terminal SNP of V20) occurs mainly in Europe (Scozzari et al., 2012). Due to the scarce occurrence of this group, further studies are necessary to clarify whether C1a2 chromosomes are the relics of an ancient European gene pool. Some archaeogenetic hints for a solution have been published recently. Haplogroup C has been described from a Mesolithic individual, found in La Brana-Arintero site in northern Spain (Olalde et al., 2014). The low genome coverage does not allow an exact phylogenetic determination. Based on one read with a derived allele, it possibly belonged to the haplogroup C1a2 (V20). Another recent study describes C1a2 in two specimens from the Neolithic Alföld LBK (with 5,300-4,950 cal BC dates) as well (Gamba et al., 2014).

Haplogroup J2 (M172) has today its highest frequency in the Caucasus and Iraq (Mesopotamia), and in the geographic region of Levant. In early modern genetic studies, J2 (together with F and G) was claimed to be an indicator of the Neolithic expansion (Semino et al., 2000), based on the clinal pattern of its frequency among the modern European and Western Asian populations. The theory has been further specified since the early 2000ies, and frequency distribution plots and surface distribution maps have revealed the J2a (M410)

(Sengupta et al., 2006) as a possible marker for early farmers' eastward migration in Central Asia. Furthermore, the subgroup J2b (M12) has also been suggested as a marker for the European Neolithic expansion (King et al., 2008). Its less frequent occurrence in modern west Turkey (Cinnioğlu et al., 2004), but more frequent appearance in Greece has been even interpreted as an indication for a maritime route of Neolithic colonisation in South Europe (King et al., 2008).

It is interesting, that J2 (M172) has not been detected in Neolithic context yet, and it is not present in the western Carpathian Early/Middle Neolithic dataset either. It might have come first with the people of the Late Neolithic cultures into Transdanubia, which means either that it is not the marker for the earliest dispersal of farmers, or that it halted in southeastern Europe for about millennium, before reaching the Carpathian Basin.

It is noteworthy that the R1b occurred first after the Middle Chalcolithic in Transdanubia. (Late Chalcolithic has not been not examined yet, and so a hiatus remains between the Middle Chalcolithic and the Early Bronze Age data.) The two R1b samples are dated to the Vučedol period (~2,870-2,580 cal BC) and to the Gáta/Wieslburg culture (~1,950-1,760 cal BC). R1b is the most frequent haplogroup in today's Europe, with a frequency peak in Western Europe (Balaresque et al., 2010). From prehistoric context, this haplogroup is known from the Late Neolithic Central Germany (Bell Beaker culture, Lee et al., 2012). The theory that R1b reached Central Europe (and possibly the Carpathian Basin as well) with the Bell Beaker migration, starting from southwestern Europe (Brandt et al., 2014) seems to be collapsing, as R1b (M269) has recently been found in Yamnaya (3,300-2,700 cal BC) population on the Russian steppe as well (Haak et al., 2015).

5.2.2 Population genetic analyses, comparing the Neolithic Y-chromosomal results to modern populations

5.2.2.1 Principal component analyses

I assembled a large number of sample sets from literature (49,516 Y-chromosomes with SNP profile), in order to make a comparative PCA and GDM analysis with the Neolithic NRY data. Due to the insufficient resolution of several, important haplogroups, almost half of the collated data had to be ignore (see Methods, chapter 4.4.2). The prehistoric cultures were arranged into Early/Middle and Late Neolithic groups, in order to enlarge the prehistoric NRY datasets.

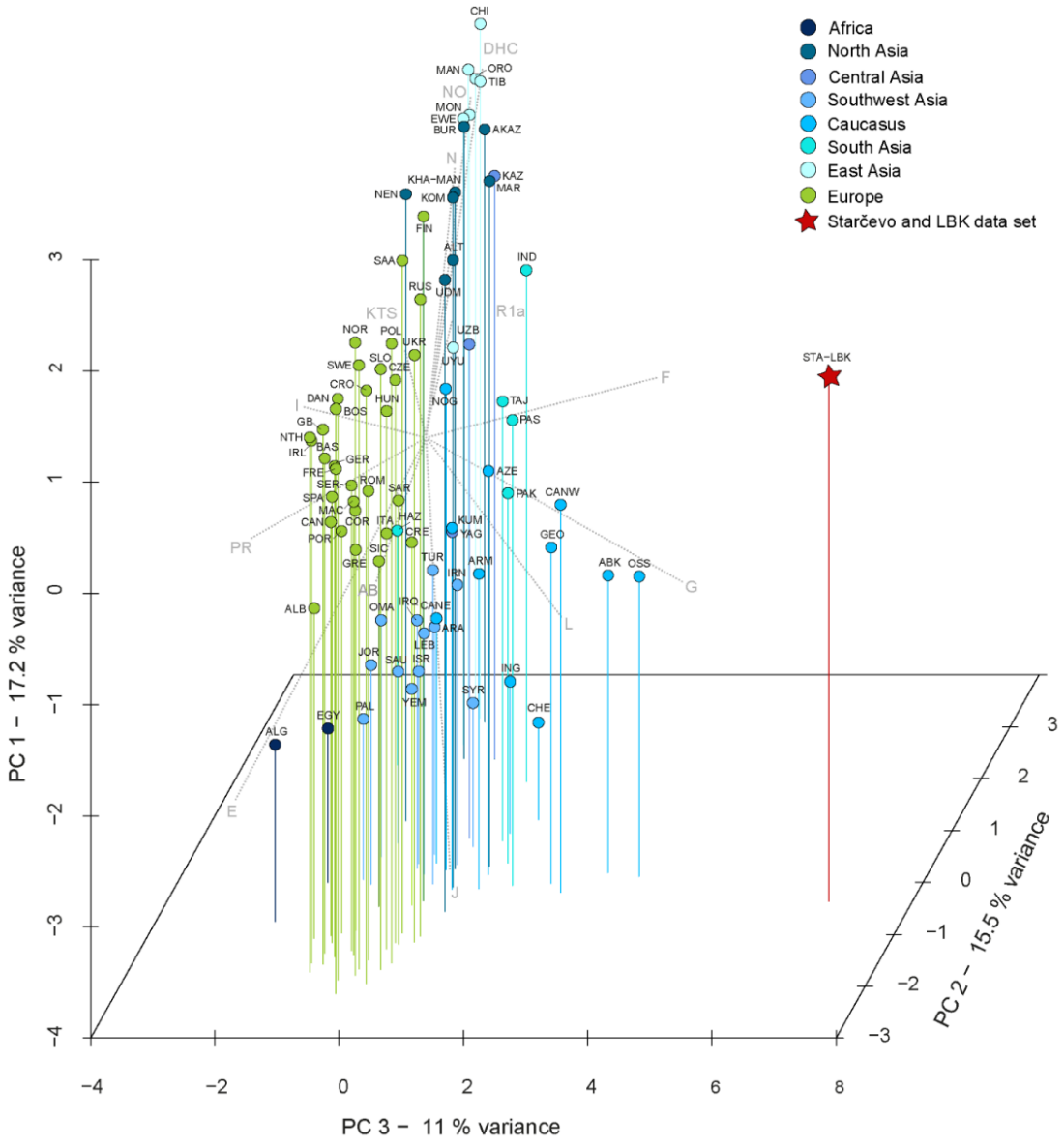


Figure 35. PCA with the pooled STA-LBKT-LBK NRY dataset and 80 modern populations. Colours of data points indicate populations from different Eurasian and African regions. The contribution of each haplogroup is superimposed as grey component loading vector. The first three principal components of the PCA display 43.7% of the total genetic variation Population codes: Starčevo and Linearbandkeramik culture (STA-LBK), Abkhazians (ABK), Albanians (ALB), Algerians (ALG), Altai (ALT), Altai Kazakhs (AKAZ), Arabs in UAE, Qatar, Kuwait (ARA), Armenians (ARM), Azeri (AZE), Basques (BAS), Bosnians (BOS), British (GB), Buryats (BUR), Spaniards in Canary Islands (CAN), north-east Caucasus (CANE), north-west Caucasus (CANW), Chechens (CHE), Chinese (CHI), Corsicans (COR), Crete (CRE), Croatians (CRO), Czech-Slovakians (CZE), Danish (DAN), Dutch (NTH), Egyptians (EGY), Ewenki (EWE), Finns (FIN), French (FRE), Georgians (GEO), Germans (GER), Greek (GRE), Hazara (HAZ), Hungarians (HUN), Indian (IND), Ingush (ING), Iranians (IRN), Iraqis (IRQ), Irish (IRL), Israeli (ISR), Italians (ITA), Jordanians (JOR), Kazakhs (KAZ), Komi (KOM), Kumyks (KUM), Lebanese (LEB), Macedonians (MAC), Manchu (MAN), Mansi & Khanti (KHA-MAN), Mari (MAR), Mongolians (MON), Nenets (NEN), Nogays & Kara Nogays (NOG), Norwegians (NOR), Omani (OMA), Oroqen (ORO), south-north Ossetians (OSS), Pakistani (PAK), Palestinian (PAL), Pashtun (PAS), Poles (POL), Portuguese (POR), Romanians (ROM), Russians (RUS), Sami (SAA), Sardinians (SAR), Saudi Arabians (SAU), Serbians (SER), Sicilians (SIC), Slovenians (SLO), Spaniards (SPA), Swedes (SWE), Syrians (SYR), Tajik (TAJ), Tibetans (TIB), Turks (TUR), Udmurts (UDM), Ukrainians (UKR), Uyghurs (UYU), Uzbeks (UZB), Yagnobi (YAG), Yemeni (YEM).

This clustering was warranted by the similar haplogroup composition of the STA-LBKT-LBK and the SOP-LGY datasets respectively. The LBK dataset is represented by three published Y-chromosomes from the LBK in central Germany (Haak et al., 2010). I performed the PCA analyses separately for the STA-LBKT-LBK and SOP-LGY merged datasets, because their affinities to modern populations are better to study if they don't build a separate Neolithic cluster on the PCA plot [for similar reasons as it was in the case of the mtDNA PCAs (Supplementary Figure 3)].

On Figure 36, we see the PCA scatterplot with pooled STA-LBKT-LBK dataset in relation to modern Eurasian and North African populations in three dimensions. Thirteen haplogroups or haplogroup clusters were used for this analysis, which give the populations a geographic-like distribution. The 3D PCA covers 42.7% of the total variation of thee 13 haplogroups. On the PC1, haplogroups J, L, E have the highest loading (as component loading vectors) along the first coordinate toward the negative, and haplogroups NO, N, I, R1a toward the positive direction. Along these vectors North European and North Asian populations (NO, N, I, R1a) are

divided from Southwestern Asia (J, L). The second PC is dominated by the DHC, NO and I, PR haplogroup pairs, dividing East Asia (DHC, NO) from Europe (I, PR). I plotted the PCA in 3D (scatterplot), because G and F* haplogroups, which are the most characteristic for the early farming NRY dataset, do have their highest loading on the third PC (for PC loadings see Supplementary Figure 5).

The STA-LBKT-LBK group is the closest to the Ossetians, Abkhazians and to populations of the North-West Caucasus and Georgia in three-dimensional space. This affinity became warranted by a Ward type cluster analysis with Manhattan distance calculation method and with paired group cluster analysis as well (Supplementary Figure 6).

The SOP-LGY dataset has different haplogroup composition from that of the STA-LBKT-LBK (Figure 36). Two new haplogroups (C and J2) cause Central Asian affinity of the Late Neolithic dataset. Along the PC1, haplogroup loading vectors PR, H, L (characteristic for Central and South Asia) point toward the positive direction and I, R1 (characteristic for Europe) toward the negative. PC2 separates North and East Asia (with high frequency of NO and C) from Southwestern Asia and Africa (NRY haplogroups E, J, and G). On the third PC, R1a, I, NO, and C are the two counterpart dominant haplogroup pairs (see Supplementary Figure 7 for the PCA loadings).

On a three dimensional PCA plot, the Neolithic group is the closest to the Nogays and Azeris (east and south Caucasus). This affinity is also observable on the scatterplot, but the arrangement is not so easy to see as on a real 3D image that can be rotated in 360°. The population of the Hazaras in Pakistan, Uyghurs, and Uzbeks are twice as far from the SOP-LGY dataset, along with a set of European populations, which are connected through the haplogroup E and I2 to the Neolithic cultures.

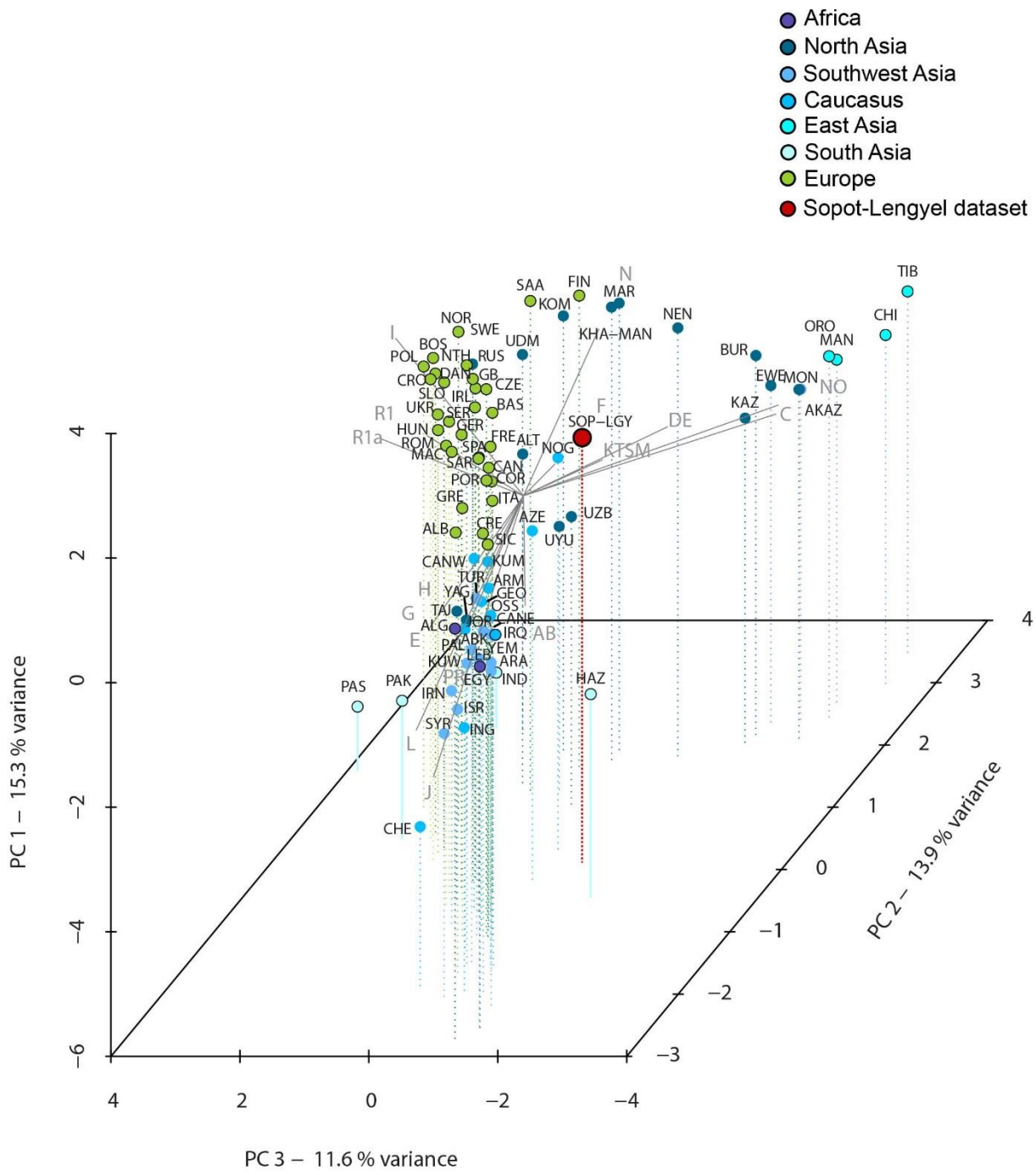


Figure 36. PCA with the pooled SOP-LGY dataset and with 79 modern populations.

The population codes are the same as at Figure 36. However, this PCA does not contain the population of Saudi Arabia and Oman separately; they are merged into the Arabs (ARA) group. The only additional population is the Kuwaitis (KUW).

The cluster analysis did not give a congruent result in this case Supplementary Figure 8a-b). The ward clustering method groups SOP-LGY dataset with modern Indians Hazaras, Tibetans, and Nenets, which is not deducible from the 3D PCA results or from the haplogroup

frequency tables. Tibetans do not share haplogroups G, I, E and they harbour just a very small proportion of J and C. In the Nenets occur no C, G, J an E at all, which questions the correctness of the dendogram built by Ward type algorithm (a). Indians and Hazaras (people in Afghanistan) have more common with the SOP-LGY dataset; both share C, G, J haplogroups with the Neolithic group. The pair group clustering gave a similar dendogram to the Ward clustering of the STA-LBK data (Supplementary Figure 6a), regarding the arrangements of the modern populations. However, it separates the SOP-LGY group from the Caucasus, Near East and from Europe. The SOP-LGY dataset is at the junction of the previous groups from North European and North, East Asian clusters (b).These inconsistent results can be the consequence of a relative small resolution of NRY haplogroups, used for a genetically heterogeneous cluster of populations. In order to understand the genetic makeup the SOP and LGY datasets better, more precise haplogroup classification is necessary.

5.2.2.2 Genetic distance maps with the Y-chromosomal data

Genetic distance calculations were performed based on 16 NRY haplogroup frequencies, comparing the two Neolithic datasets with 103 modern populations. The genetic distances were mapped in a similar way to the mtDNA GDM analyses. According to the F_{st} values, the ten closest populations to the STA-LBKT-LBK group are the Adyghes, Kabardins, Sardinian, Balkarians, Abkhazians, Azeris, Georgians, Cherkessian, Armenians, and Turks around the Black Sea in this order (Supplementary Table 25).

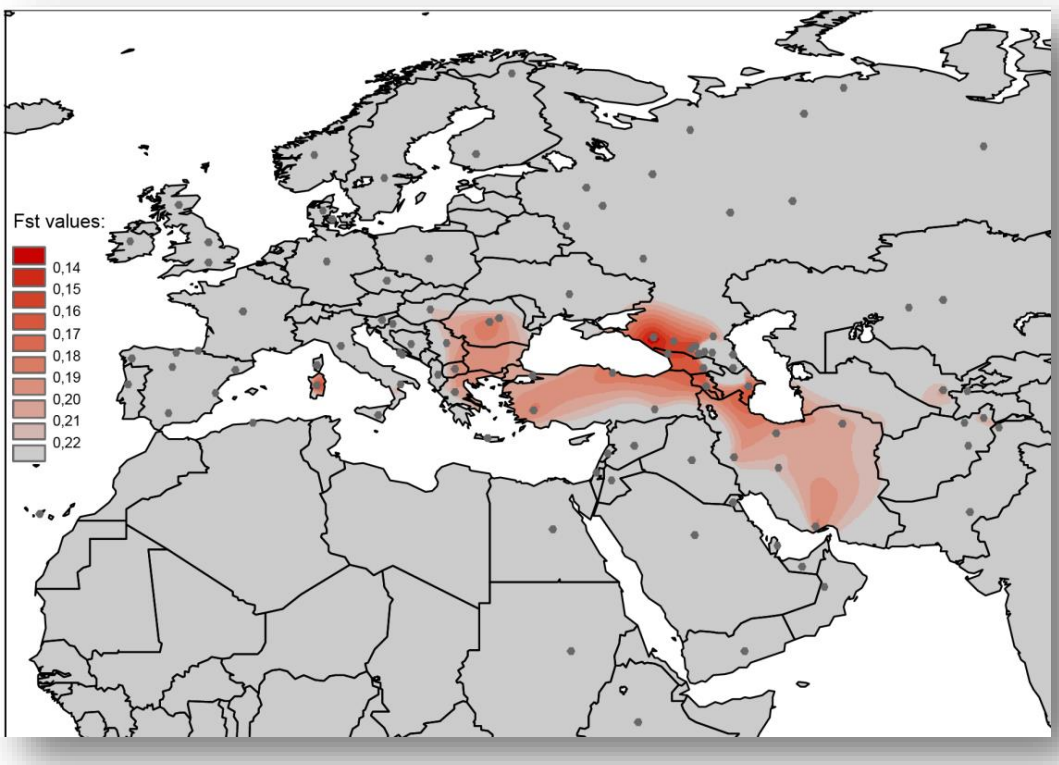


Figure 37. Distance of the NRY STA-LBKT- LBK dataset from 103 modern populations.

Y-chromosomal genetic distances (F_{st}) were computed between the STA-LBKT-LBK samples and 103 present-day populations of Eurasia and North Africa, and visualized on a geographic map. Grey dots denote the location of present-day populations. Colour shadings indicate the degree of similarity or dissimilarity of Neolithic samples to the modern-day populations. Short distances and great similarities to present-day populations are marked by red areas. F_{st} values were scaled by an interval range of 0.01. F_{st} values higher than 0.21 were not differentiated (grey areas).

The SOP-LGY dataset has the highest affinity to the Sardinians, Azeris, Romanians, Macedonians, Greeks, and Turks in the southern Black Sea coastal area, Turks in Istanbul, Armenians, Seklers, and Turks in Mediterranean Region, Central Anatolia Region, and Aegean Region, in this order (Supplementary Table 25). Because I harmonised the interval ranges of the two NRY maps (Figure 37, 38), more than 10 populations were displayed in colour on the SOP-LGY map. Populations of Crete, Afghanistan, Iran, and Uzbekistan are also highlighted, which are within the 20 most similar populations to the SOP-LGY dataset.

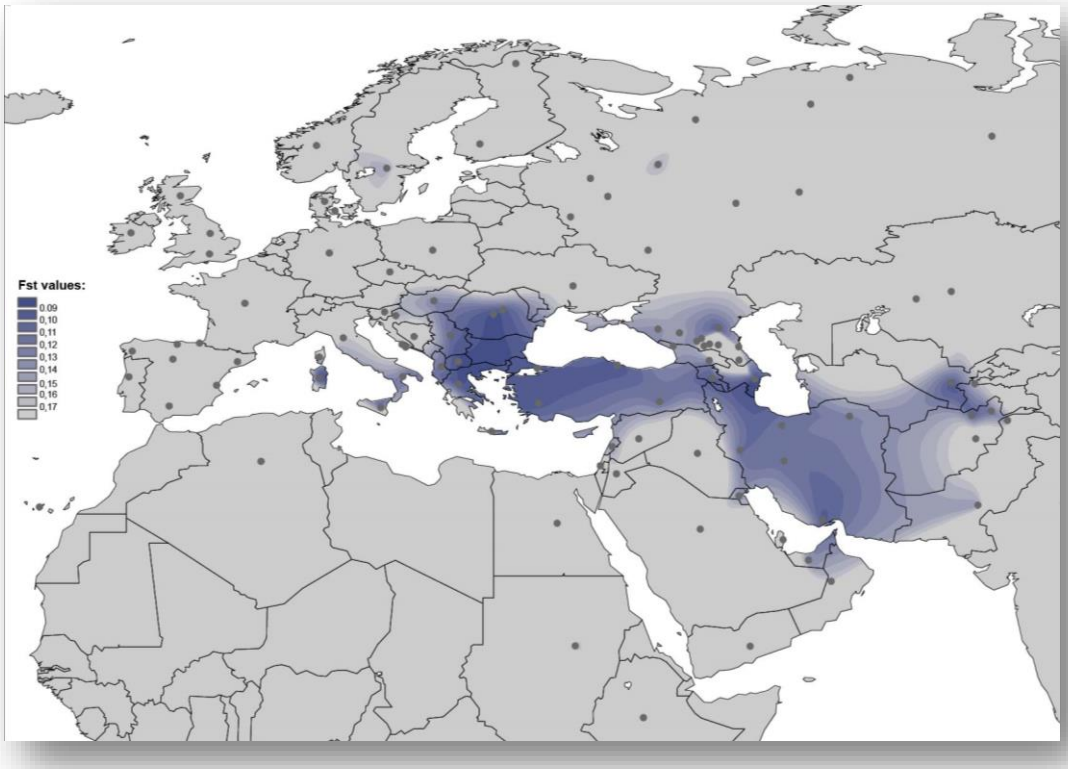


Figure 38. Genetic distances of the NRY Sopot & Lengyel dataset from 102 modern populations.

Y-chromosomal genetic distances (F_{st}) were computed between the SOP-LGY NRY results and 102 present-day populations of Eurasia and North Africa and visualized on a geographic map. Short distances and great similarities to present-day populations are marked by purple areas. F_{st} values were scaled by an interval range of 0.01. F_{st} values higher than 0.18 were not differentiated (grey areas).

Generally, the two NRY distance maps (Figure 37, 38) are comparable. The major difference is that the SOP-LGY group has a greater affinity to Turkey and Central Asia (Uzbekistan, Afghanistan). Comparing the methods of PCA, hierarchical cluster analysis to the GDM, it is interesting that the Sardinians are not affiliated with the Neolithic groups on the PCA plots and on the dendograms. However, they show small genetic distances from both Neolithic groups on the GDMs. The PCA, GDM, and cluster analyses with STA-LBKT-LBK dataset are better comparable than the each evaluation approach of the Late Neolithic data. The three new haplogroups, occurring in the Late Neolithic, have divergent affiliations and complex dispersal histories, which make difficult the evaluation the data due to the available small haplogroup resolution information.

6 Discussion

6.1 Authenticity of the results

The preclusion of probable modern human DNA contamination is always an issue in the investigation of prehistoric human remains, which induced the establishment of numerous criteria for endogenous aDNA authenticity, discussed in the introduction chapter 1.7.3 (Gilbert et al., 2005; Pääbo et al., 2004; Willerslev and Cooper, 2005). It is important to evaluate the authenticity of ancient human DNA by completing several independent experiments in a logical framework (Gilbert et al., 2005). In previous publications, the Bioarchaeometry team in Mainz has discussed several criteria compiling a chain of evidence for the authentication of the ancient DNA results (Brandt et al., 2013; Haak et al., 2010, 2005), which I applied analogously in this study (see in method chapter 4.2). Regarding the key criteria of Gilbert and his colleagues (Gilbert et al., 2005), the first five were fulfilled in the case of the 256 reproduced samples. To be specific, the work areas were isolated (I), negative controls were used and sequenced (II), the ancient DNA had an expected molecular behaviour (III), multiple PCRs gave a consistent results (IV), and the samples were at least partially cloned (V). The seventh point, which is the preservation of other biomolecules correlating with ancient DNA, will be checked during the ongoing C/N isotope study of my colleague, Marc Fecher. Only two major points can be held against the authenticity of the results. The DNA quantification and the independent replication of the results by a separate research group were not accomplished.

The DNA preservation corresponded to the morphological look of the samples and to the sample types (Figure 11. Success rate of amplification with different sample types, amplifying different lengths of fragments.). Well amplifiable DNA with only a few deaminated or degraded nucleotides was gained from bones with a promising look and from most of the teeth and petrous parts of the temporal bones. On the other hand, major problems occurred with the samples having a bad morphological preservation, such as samples from the sites Balatonszárszó and Nitra. They were rather porous, or had only thin bone cortical part with unusual colour (e.g. white or dark brown).

The mtDNA sequences of the co-workers (Supplementary Table 27) differ from all reproduced ancient haplotypes connected to a certain researcher. Thus, all investigators can be clearly excluded as potential source of contamination. Furthermore, during the three and a half year lasting palaeogenetic investigations of this study, only 4.14% of the PCR blank

controls showed sequenceable DNA products in general (see the details in chapter 5.1.1). These sequences were compared to the samples that were analysed simultaneously with the respective blank controls. However, in all cases the sequences of the blank controls differ from that of the samples. In this way, a potential influence of the detected contaminations can be clearly ruled out.

Consequently, in the light of the authentication criteria that were applied to these data (chapter 4.2), and in absence of indications of a systematic contamination of one or more skeletons, I am confident that the results presented in this study are authentic and derived from endogenous DNA of the samples.

6.2 Ancient DNA preservation

The DNA preservation of the studied western Carpathian Basin samples was good on average, with an overall success rate of 86% for the typing of the mitochondrial HVS-I region (Table 4, Supplementary Table 3). This success rate is comparable to the Middle-Elbe-Saale aDNA study, which could successfully replicate 84.1% of the individual HVS-I haplotypes (Brandt et al., 2013). The amplification rate was slightly smaller on the Alföld (74.4%, Keerl, 2014), which is explicable with different soil and climatic conditions of the Alföld, coupled by the fact, that several sample sets originated from older (30-60 years before) excavations.

During the laboratory work, I observed that the DNA preservation is often site specific, changing from site to site or sample set to sample set. This is explicable, since the progress of DNA fragmentation can be influenced through the soil conditions and the post excavation history of the series, which affect the skeletons within the same storage (collection) simultaneously.

The analysis strategy was to screen every larger site with five individuals (A and B samples per individual). After the first screening, I decided, whether the DNA preservation justify any further examination. The only exception was Balatonszárszó (BSZ), which was a special case study at the beginning of my work in Mainz, taken into my PhD work additionally. In the case of Nitra (NITR), I only processed the selected first five individuals. The rest of the sample set would have consumed too much time and money to analyse. In the case of the site Mórágy (MORT), I favoured the teeth samples after the initial screening, and those individuals

that only had bone samples were not analysed. Skeletons from the rest of the sites were well preserved on an average, and were processed entirely.

The amplification success was enhanced by the sampling strategy of our team. We preferred to sample freshly excavated skeletons that have not been washed before. Out of 110 such cases only six were not reproducible for the HVS-I region. This success is due to the short time lapse between the excavation and our sampling, letting only a few experts handle the skeletons. Those few people, who came into contact with the bones, were easy to track down and involve in the control modern DNA sampling.

The different sample types showed divergent DNA preservation (see chapter 5.1.2, Figure 11); thus teeth and petrous part of the temporal bone proved to have better preserved DNA than long bones. Petrous part has been recently revealed as good resource of endogenous aDNA, containing even more genomic DNA than teeth (Gamba et al., 2014). Such differences could not be observed in my study, because I did not consequently quantify the mitochondrial of Y-chromosomal DNA.

Grounding complete teeth, I did not distinguish different parts of them (crone, pulp cavity, dentine), hence within tooth altering DNA survive could not affect the observed variable success of DNA amplification during my laboratory work (Adler et al., 2011).

Short fragments could be amplified easily (Figure 11), which fact was described as a part of the authentication in aDNA studies (Malmström et al., 2007). Furthermore genome studies also reported the abundance of short endogenous DNA fragments (Brotherton et al., 2013).

Summarizing, DNA preservation was better than expected, which is the consequence of the sampling strategy and the well-established protocols of the laboratory work (Brandt et al., 2013; Haak et al., 2010, 2005). Further improvement of the analyses would be the quantification of the DNA extracts (real-time or Q-PCR), which would help estimating the parameters of aDNA preservation prior, to the establishment of an amplification strategy.

6.3 Methodological discussion of the population genetic analyses

One of the crucial methodical questions is, whether the generated sets of prehistoric individuals can be regarded as populations. The word population is generally used for all inhabitants of a particular place (Oxford Dictionaries), but its meaning alters in different

disciplines. In population genetics, the term population is referred to a group of organisms, in which any pair of members can breed together. This definition implicates, that these organisms belong to the same species and live near to each other. Even if these two criteria were fulfilled for the Carpathian Basin datasets, the contemporaneousness would also be necessary for possible interbreeding. It is however not guaranteed, especially in the LGY "population" of the long lasting Lengyel period.

The populations were differentiated based on archaeological culture affiliations, which implicate methodological challenges. Disregarding from the inconsistent use of term culture in the literature (Zalai-Gaál, 2010), the cultural affiliation of a certain assemblage is not always evident to define. More and more mixed cultural assemblages are discovered, and the solid determination of cultures cannot be maintained in every period of the Transdanubian Neolithic (Jakucs and Voicsek, 2014). Even if the material features (finding assemblages, archaeological characteristics, burial rites an settlement or house types) are characteristic for one specific group of sites or region and period, the question remains whether archaeological cultures represent self-identifying societies or ethnic groups (idea originates from G. Childe), or whether they were differentiating themselves by own languages or dialects. Certain answers cannot be given to these questions regarding the Neolithic period, even if some attempts have been made, for example in connecting Indo-European language to the early farmers (Renfrew, 1987). It is far from my intention, to identify the archaeological cultures with ethnicities, language families, or tribes. The work hypothesis and strategy was rather to define Neolithic populations, using the sets of individuals assigned into archaeological cultures.

I used the term metapopulation for the hunter-gatherers, whose representatives are spatially separated from each other. The interaction of these hunter-gatherer groups are not excluded, but they certainly did not form a population, in the proper meaning of the expression.

During the performance of population genetic analyses in a various constellations of datasets, I noted the challenges and favours of the used statistic and evaluation methods. For example, both the PCA and the MDS analyses have their advantages and disadvantages. Whereas the PCA takes account of phylogenetic structure through the assessment of haplogroups, it also reduces the complete genetic variation to the frequencies of only a bunch of haplogroups. Especially the insufficient resolution of the mtDNA H haplogroups can cause

false clustering, resulting in false positive affinities of groups with divergent H haplogroup types or subgroups (Figure 14). For example, more detailed H resolution (obtained through coding region analysis or whole mitochondrial genome sequencing) would probably better separate the Iberian Neolithic from the Central European and Carpathian Basin Neolithic populations. The F_{st} and MDS analyses on the other hand, take each HVS-I positions independently into account. The considered information is not reduced by haplogroup assignments, evaluating each haplotype or sequence equally and individually. On the other hand, the phylogenetic relationships of different haplotypes are not perceived in these analyses. Even if two haplotypes have only one mutation difference on the range of the HVS-I (for e.g. rCRS H haplotype and a 16298C HV0 haplotype), they can be on different branches of the mtDNA phylotree, being in a greater genetic distance than other haplotypes within the same haplogroup (such as rCRS H haplotype and a 16093C H haplotype).

I completed the suite of analyses with ASHA, which involves advantageous features of both haplogroup and haplotype based analyses, since during the haplotype comparisons the haplogroup assignments are also considered. This method also has boundaries, due to the resolution limitations of the HVS-I sequences. Especially the basal H (rCRS haplotypes) and the basal J (16069C, 16126C haplotype) haplogroups can incorporate several different types in the coding region, which can cause false positive haplotype matches among the prehistoric datasets.

These examples demonstrate that there is no best approach among ASHA, PCA and F_{st} analysis or MDS. Several haplogroup and haplotype based analyses have to be considered simultaneously in order to reveal the genetic composition and connections of the studied populations in such depth as the performed mtDNA and NRY SNP typing enable it.

6.4 Population history through ancient DNA in the Neolithic Transdanubia

6.4.1 The Mesolithic/Neolithic transition in the western Carpathian Basin

The lack of Mesolithic burials and skeletal remains in western Hungary inhibits studying the question of transition concentrating on Transdanubia alone. In order to estimate the Mesolithic substrate in the Carpathian Basin, I processed one individual from the Adriatic Croatia (Korčula Island), and used published Upper Palaeolithic (Magdalenian) and Mesolithic mtDNA data (Fu et al., 2013; Bollongino et al., 2014; Bramanti et al., 2009), constructing a

Central and North European hunter-gatherer dataset (referred as HGCN). I used this group as a projection for the pre-Neolithic mtDNA variation in Transdanubia.

All these pre-Neolithic individuals belong to the mtDNA haplogroup U (U5a-b, U4, and U2 occur among them), and even the new Adriatic sample (called STANKO) was assigned to the sub-haplogroup U5b2a5. Furthermore, regarding the earlier Palaeolithic Dolní Věstonice from the pre-ice age Gravettian period, the haplogroup U8 also joins to the pre-Neolithic variety of U. In contrast to this homogenous hunter-gatherer substrate in Central Europe (Figure 10), I detected a sudden increase of haplogroup and haplotype diversity in the first Transdanubian farming populations (Table 5, 6). The transition was not continuous from a genetic point of view; the early Neolithic discontinuity is proven through TPC and F_{st} analyses on the haplogroup and haplotype level as well (Table 7, 21). The genetic variation of the STA suggests a massive migration of the first farmers into the Carpathian Basin.

On the other hand, migration and acculturation represent only a small fraction of the total possible transition forms (Robb and Miracle, 2007; Zvelebil, 2001) and Neolithic transition might have to be explained by more complex models (especially on a regional scale) than the two mentioned types. For example, it has been suggested that on the western fringe of the LBK dissemination area forager-farming interactions resulted in a greater portion of hunter-gatherer influx in the Neolithic population (Gronenborn, 2007, 1999). A similar scenario has been proposed for the emergence of the LBKT, accordingly, a contact zone of local hunter-gatherers and immigrant STA framers around the Lake Balaton in Transdanubia potentially triggered the formation of the LBKT (Bánffy, 2006, 2004; Bánffy et al., 2007). The hunter-gatherer contribution in the early farming populations has been underlined in the light of the newly described Mesolithic archaeological sites in western Hungary as well (Eichmann et al., 2010). The STA reached the north shore of Lake Balaton, but did not disseminate to the northern parts of Transdanubia. Nevertheless, the Starčevo lithic assemblages show contact to the northern territories, especially through the Szentgál radiolarite material (Regenye, 2010). A potential Late Mesolithic population still does not have any direct traces in northern Transdanubia; however, changes in pollen profiles around Zalavár indicate pre-Neolithic (seventh millennium BC) human effects on the flora in the western Balaton region (Juhász, 2004).

Haplogroup U4 is the single mtDNA haplogroup in the STA dataset, which can probably indicate admixture between the hunter-gatherers and farmers. From the NRY side, the I2a1

haplogroup has to be mentioned, as possible representative of hunter-gatherer legacy. Whereas the "Neolithic genetic package" (Brandt et al., 2013) dominates the haplogroup composition of the STA (86.36%), the genetic difference among the STA and LBKT leaves some space for a small scale hunter-gatherer infiltration, which probably occurred within the Carpathian Basin, parallel to the late STA- formative LBK period.

Despite the fact that the LBKT mtDNA gene pool shows no significant variation compared to the STA (Table 5), but differs significantly from the currently known mtDNA diversity of the Central/North European hunter-gatherers, introgression of local forager populations in early LBKT communities cannot be excluded. A considerable difference between the mtDNA compositions of the LBKT and STA cultures is the elevated frequency of haplogroup H in the LBKT (Table 5). H has been reported with high frequencies (42.9 %) from Mesolithic forager communities of the Iberian Peninsula (Chandler et al., 2005; Hervella et al., 2012) and was also prevalent in hunter-gatherers from north-eastern Europe (Der Sarkissian et al., 2013) indicating the pre-Neolithic presence of haplogroup H in a vast region of Europe. Despite the fact that H is so far undetected among Central/North European foragers (Bollongino et al., 2013; Bramanti et al., 2009; Fu et al., 2013), its presence in this region during the Mesolithic cannot be precluded. Some of the H sub-groups could represent a part of the forager genetic substrate, while other sub-clusters might have arrived with the farmers at Transdanubia (based on the assumptions of Brotherton et al., 2013). Hence, it could be speculated that a noteworthy amount of H variability in our LBKT samples potentially represents a part of the Mesolithic substratum, which would be in accordance with results of modern population genetic studies (Achilli et al., 2004; Behar et al., 2012a; Richards et al., 2000; Roostalu et al., 2007). However, the majority of the markers relevant for the haplogroup H variation is located outside the here studied control region (van Oven and Kayser, 2009). Detailed analyses of significant coding region polymorphisms will be necessary in the future, in order to assign H lineages to their potential Mesolithic or Neolithic origins.

Further candidate markers of a possible hunter-gatherer genetic legacy are the mtDNA haplogroups U2 and U5a, which have been reported from hunter-gatherers of distant European regions (Bollongino et al., 2013; Bramanti et al., 2009; Chandler et al., 2005; Der Sarkissian et al., 2013; Fu et al., 2013; Hervella et al., 2012; Krause et al., 2010; Sánchez-Quinto et al., 2012). Furthermore, the Y-chromosome haplogroup I1-M253, whose expansion in Europe has been postulated in the Late Upper Palaeolithic (Rootsi et al., 2004), was identified

in the LBKT. This NRY sub-haplogroup has not been described yet in European pre-Neolithic context, but the dominance of NRY I* haplogroup among the hunter-gatherers (Lazaridis et al., 2014) suggests that I1 belonged to the Mesolithic substrate too.

The change in paternal genetic substrate at the Mesolithic/Neolithic transition is even more hypothetical. If haplogroups I1 and I2a1 could be interpreted as a persisted Mesolithic substratum in the early farming communities, it was diminished by incoming farmer lineages such as G2a, G2a2b, and F* (Table 10). Even if the here presented STA and LBKT Y-chromosomal data means a significant increase of to date available Y-chromosomal information from the sixth millennium BC, it is still scarce to define the proportion of male hunter-gatherers' admixture in farming populations.

6.4.2 The genetic succession of the six studied cultures in the western Carpathian Basin

I performed a suite of population genetic analyses (e.g. calculation of molecular diversity indices, F_{st} analysis, AMOVA, TPC, and ASHA) in order to reveal the genetic connections of the studied Neolithic cultures' people (Table 6- 7, Figure 21-22). I found predominant genetic continuity through the Neolithic of Transdanubia. The STA mtDNA gene pool determined the succeeding Neolithic periods. It contributed to 61.5% of the mtDNA variability of the Middle Neolithic LBKT, 48.4% of the VIN, and 54%-46.3% to the genetic diversity of the Late Neolithic SOP and LGY populations respectively (Supplementary Table 15). Even with a gradual dilution of the STA lineages, still 38.5% of the BL lineages had a STA legacy in the Middle Chalcolithic Transdanubia. These results stand in contrast with the classic osteological evidence, which suggest discontinuity between the STA and the LBKT (K. Zoffmann, 2012). Nevertheless, the large Starčevo series from Alsónyék-Bátaszék have not yet been involved in the Penrose analyses. These will probably change the connection network of the Neolithic cranial series on some extent.

The mtDNA haplogroup N1a has to be highlighted from the other "Neolithic package" haplogroups. N1a was quite frequent in the Neolithic of the Carpathian Basin (Figure 39) and in the sixth-fourth millennia BC central Germany. It disappeared in the Middle-Elbe Saale region from the Bernburg period and the local Late Neolithic (third millennium BC, Brandt et al., 2013). N1a frequency raised after the Starčevo period in Transdanubia, and had a quite

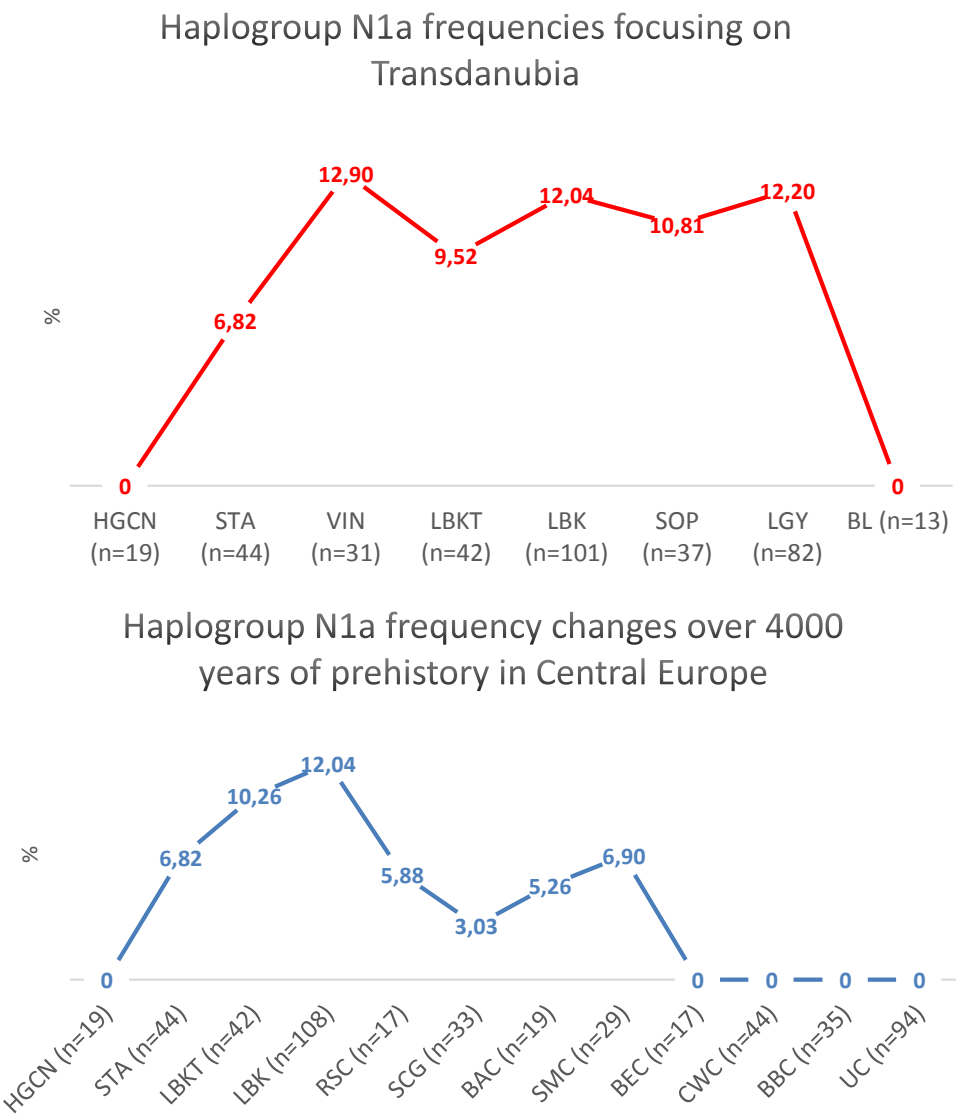


Figure 39. Alteration of the N1a haplogroup frequencies in Central Europe and in the western Carpathian Basin during the studied Mesolithic, Neolithic, Chalcolithic and Bronze Age periods.

Abbreviations of cultures/populations: HGCN: Hunter-gatherers in Central/North Europe, STA: Starčevo, VIN: Vinča, LBKT: LBK in Transdanubia, LBK: LBK in Central Europe, SOP: Sopot, LGY: Lengyel, BL: Balaton-Lasinja, RSC: Rössen, SCG: Schöningen, BAC: Baalberge, SMC: Salzmünde, BEC: Bernburg, CWC: Corded Ware, BBC: Bell Beaker, UC: Únětice.

constant rate during the Transdanubian Middle and Late Neolithic (9.5-12.9%). Although, the N1a was not detected in the Balaton-Lasinja, this phenomenon can rather mean a research gap, caused by the small number of samples from this culture. Balaton Lasinja is almost 1,000 years earlier than the Bernburg culture in central Germany, which lets presume that the dilution of N1a haplogroup could rather have happened in the Late Chalcolithic period of the Carpathian Basin.

Besides this continuity, spanning across ca. 1,500 years of Transdanubian prehistory, I also observed smaller scale infiltration or migration events. This is reflected in the fact that each genetic dataset has some new components. Such examples are the elevated number of hunter-gatherer mtDNA haplogroups in the Vinča population, the increase of H haplogroup frequency in the LBKT, new U8b sub-haplogroups from the SOP period on, the mtDNA haplogroup T2f in the LGY, and the retrieval of the T1 in the BL population. Further indices for population genetic events are the significant Tajima D value of the SOP population (Table 6) and the significant difference between the Vinča and the BL in the TPC analysis (Table 7) This latter can be explained with a suite of small-scale population genetic events, which resulted in a detectable difference over this period of ~1,000 years.

Among the above-mentioned mtDNA evidence, the U8b might signalize genetic influx from communities that were isolated before the Late Neolithic within/from Transdanubia, which populations might have had a hunter-gatherer ancestry. Haplogroup U8 (based on the finding from the Gravettian site Dolní Věstonice) has been defined as a part of the hunter-gatherer mtDNA substrate by G. Brandt and our colleagues (Brandt et al., 2013; Fu et al., 2013). Congruently to this theory, V. Keerl has found the U8b in ALBK Szatmár and Tiszadob groups on the north-Alföld as well. On the site Mezőkövesd-Mocsolyás of the Szatmár cultural group, several other haplogroups (U4, U5) of possible hunter-gatherer origin have been detected by V. Keerl (Keerl, 2014). Contrasting hypotheses exist about the origin of the Szatmár group, which lived partially overlapping with the southern Körös communities. One of the early assumptions connected it to a Mesolithic substrate (Kalicz and Makkay, 1977, p. 29). Recently, it has been regarded as the northward movement of the Körös culture, with some eastern (Criş) connections (Domboróczki, 2009). The following ALBK-Tiszadob group, containing also haplogroups U8, has been described as having hunter-gatherer legacy in the material culture (Domboróczki, 2003). Even if the archaeological assessment of the Mesolithic contribution is not coherent regarding the evaluation of Alföld Neolithisation, recent genomic evidence have

been demonstrating hunter-gatherer components in farming populations as early as the Körös period (Gamba et al., 2014).

Looking at the Y-chromosomal variation, the turn of the Middle/Late Neolithic bears more changes than it is observable in the mtDNA gene pool. The three new NRY haplogroups J2, C, and E1b1b1a (Table 10, Supplementary Table 5) signalize new population elements in the Sopot communities, which subsisted during the Lengyel period of the region as well. Without any Y-chromosomal comparison data from the adjacent Neolithic populations (Alföld region or central Germany), it is difficult to define the origin of these new components. The first haplogroup J2 probably means influx from southeastern directions, according to the modern occurrence of this group (see in chapter 5.2.1). The question is, why this group cannot be seen in the Early Neolithic of Transdanubia. Had it been there and remained undetected by chance? Alternatively, did J2 arrived with the first farmers, and halted in southeastern Europe for a while (~1,000 years), from where it dispersed in the Sopot period? J2 has been described as a potential marker for the first farming migration toward Central Europe (King et al., 2008; Semino et al., 2000; Sengupta et al., 2006), and therefore its lack in the STA and LBKT is an extremely interesting question, which remains open for now.

The other new NRY haplogroup is C, which might have a different origin than the J2. Interpreting its first occurrence in the SOP, we have to reach back to the La Brana Mesolithic specimen, which is a NRY haplogroup C too (Olalde et al., 2014). Haplogroup C can be regarded as a candidate for representing the pre-Neolithic NRY substrate in the Carpathian Basin too, along with NRY haplogroup I. It is an interesting phenomenon, with several examples from Europe that hunter-gatherer communities might not have admixed with the first farming communities, but rather persisted isolated over a longer period of time (see for genetic and isotopic evidence Bollongino et al., 2013, Haak et al., 2015, archaeological hypothesis from Hungary in Domboróczki, 2003; Bánffy, 2006). Forager groups might have withdrawn into some recluse territories (e.g. highlands) in a farming milieu, and established contacts that were more intensive only after a time lapse of several generations. It could be the case in in the Earliest/Early LBKT (Bánffy, 2006), although we have no dated archaeological traces of foragers from this time. A recent study by Gamba et al. has described two Y-chromosomal haplogroups C in the Neolithic of the Alföld region. The haplogroup C in the ALBK and Alföld distribution of the LBKT gives a context to the appearance of this haplogroup in the Sopot culture's population (Gamba et al., 2014). Summing up, the NRY haplogroup C could either

came with a new wave of farmers (in the SOP period) from southeastern Europe or it could persisted during the sixth millennium BC in the Carpathian-Basin (e.g. highlands of northeastern Hungary) and reflects increasing contacts between the Alföld and Transdanubia at the turn of the sixth-fifth millennia BC. In this context, the NRY haplogroup C can be considered in parallel with the mtDNA haplogroup U8, inasmuch both of them signalise the new connection system of the Sopot population.

During the Lengyel period, the people of the Sopot and LBKT communities in Transdanubia lived further with a virtually unchanged genetic structure (Table 5). The influx, which is detectable in the Sopot population, imprinted its signs on the paternal and maternal genetic variation of the LGY population as well. The continuity is especially marked in south-Transdanubia, where the Starčevo genetic legacy was the strongest, dominating even in the Late Neolithic period.

The small Balaton-Lasinja sample set shows different haplogroup composition from the previous periods. However, on a local scale, its genetic continuity is assured on the archaeological site of Veszprém Jutasi út. Here, the specimens from the Lengyel II, III periods and the Lasinja skeletons carry shared haplotypes and haplogroups in an elaborate nexus (Supplementary Table 3). On the other hand, the re-emergence of the mtDNA haplogroup T1a in the BL, after a hiatus in the SOP and LGY periods (sum of all typed individuals =119) is noteworthy. It can mean a new migration wave in the BL period, which reached Transdanubia probably from the southeastern directions again, if we consider the archaeological assumptions (Bánffy, 2002). Nevertheless, we have to keep in mind that Transdanubia is a quite large geographic region, and not all part of them were covered in all cultural periods (Figure 6). Especially the western fringe of Transdanubia could not been sampled and studied sufficiently, and therefore we cannot compare the BL data from Keszthely (where the T1a haplogroups were detected) with gene pools of local antecedent populations. This site is near to an assumed “green corridor”, which might have served as the “western route” through the Carpathian Basin (Bánffy, 2002). Movements along this path could have caused reoccurrence of haplogroups of southern origins, remaining undetected in the eastern part of Transdanubia.

To conclude the genetic population history of the Transdanubian Neolithic and Early/Middle Chalcolithic periods, I observed a pronounced genetic relation between each of the six studied Transdanubian cultures’ population. This relation is best shown on the ASHA diagram, which demonstrates that on average 70% ancestral lineages [coming from the previous cultur(es)] mix with 30% own, new lineages in the VIN, LBKT, SOP and LGY

populations (Figure 21). The prevailing genetic continuity is coloured by small changes and new components in each population. Whether these smaller migration events were connected to the changes of culture or happened more gradually, remains an open question. Genome wide analyses with comprehensive radiocarbon measurements could enlighten this issue.

6.4.3 Genetic regionality in Transdanubia

Studying the genetic data of two significant site-complexes (the LGY-BL Veszprém and the STA-SOP-LGY Alsónyék sites), I detected considerable population continuity over a time span of several hundred years and over more than thousand years respectively. The question has to be raised however, whether local continuity was accompanied by regional genetic differentiation in the Middle and Late Neolithic Transdanubia. The mosaic-like environment of Transdanubia, contrasting the ecological uniformity of the Alföld region (Great Hungarian Plain) could have promoted separation or isolation of farming communities. The LBK dispersal is shown to have been a massive migration event, starting from the western Carpathian Basin. It is highly likely that the north, northwestward migration of a part of the LBKT population did not sweep over the whole region, affecting a genetic mixture and homogeneity of the remaining LBKT population. This hypothesis is prompted by the fact that the late LBKT cultural traits were differentiated within Transdanubia, which was either the result of previous connection systems or new influences from southern and northern adjacent regions. Archaeological regional variations have been identified in the cultural characteristics of the SOP and LGY as well (Regenye, 2011; Bánffy, 2002). However, the regions' or groups' expanses do not seem to be conserved over the whole Neolithic period. The north-south section, observed in the Early Neolithic period, retained with a slightly shifted section (or interim) zone in the late LBKT period in the Middle Neolithic as well. Nevertheless, the Sopot and the Lengyel material culture shows rather east-west differences in Transdanubia (Regenye, 2011, 1996a).

From the several ways of regional sectioning, I chose the north-south one, proceeding from the north-south division of Transdanubia observed in the Early Neolithic period. It can be regarded as a work hypothesis, since I did not (could not) consider all archaeological culture and group divisions, due to the unbalanced distribution of the studied site's locations. The LBKT and the LGY datasets have proven to be large enough to study this question, however,

the SOP and BL sets of data are too small for regionality studies, and therefore they were regarded on the whole.

On the PCAs and in the AMOVAs revealed north-south genetic difference in the LBKT and LGY datasets are the first signs of a genetic structure in Neolithic Transdanubia (Figure 26, Supplementary Table 17, 20). Regionality in the material culture, and possibly in the genetic compositions were a dynamic phenomenon in the Carpathian Basin, induced by contact networks, geographic, climatic capabilities, and accesses to food resources. In certain regions, which were distant from the major contact and migration routes, settlements of autochthon populations could have subsisted, preserving the genetic composition of the forbears. For a more exact consideration of this question, further systematic genetic analyses would be needed from the yet unsampled regions of the western Carpathian Basin, accompanying with intensive radiocarbon dating of the samples.

The Neolithic individuals studied by V. Keerl from the Alföld region were geographically not evenly distributed either. However, the Alföld-LBK (ALBK) regional groups of Bükk and Tiszadob represent doubtless a supposed northern genetic makeup. The Carpathian Basin scale PCA and MDS did show that the Alföld groups and populations build two clusters; one involves the Körös, ALBK-Szatmár and ALBK-Esztár with southern genetic characteristics (Figure 28, 30), and the second group is built by the Tisza, ALBK-Szakálhát and the ALBK-Tiszadob-Bükk. Tisza and Szakálhát are the two latest cultures under study from the Alföld, which connections to the Tiszadob/Bükk and to the northern LBKT and northern LGY are remarkable. This dualistic genetic affinity could be a result of a substantial Early Neolithic maternal genetic substrate (Starčevo-Körös-Criş) along with intensified contact systems between east and west part of the Carpathian Basin, affecting especially the northern part of Transdanubia. The common northern specific maternal lineages, which also dominated the Late Neolithic Tisza Region seemingly arrived northern Transdanubia from the east or northeast part of the Carpathian Basin. Nevertheless, without studying the adjacent regions (especially Slovakia, Croatia, Serbia, Romania), the assumptions about diffusion directions remain research hypotheses for future investigations.

6.5 Evaluation of the Transdanubian ancient DNA results in the light of the available Neolithic and Bronze Age Eurasian aDNA datasets

Considering the large-scale prehistoric genetic structures, the prehistoric populations cluster by geographic location and chronological position on the mtDNA PCA and MDS plots (Figure 14, 17). The mtDNA haplogroup variability, identified through the HVS-I region, obtains sufficient information to resolve the genetic differences between geographically distant populations, such as the separation of the inhabitants of the Yamnaya culture from the people of the Bell Beaker culture, or the hunter-gatherers in East and southwest Europe on the PCA plot (Figure 14). However, the picture became less clear, regarding the sequence level analysis, visualized by MDS plot (Figure 17). Either the F_{st} calculation (based on solely the HVS-I) does not sufficiently separate all of the geographically distant groups (such as hunter-gatherers in Eastern and in southwest Europe), or the PCA gives too much emphasis to less frequent haplogroups (e.g. T*, U* or N*). The actual mtDNA structure can only be revealed regarding all analyses simultaneously.

The hunter-gatherer groups from East and Southwest Europe show a genetic differentiation from each other, and from the Neolithic cultures, regarding their mtDNA haplogroup frequencies, which separation is parallel to geographic locations (Figure 14). Hunter-gatherers from Eastern Europe show affinity to Eastern European and West Asian Bronze Age populations, whereas hunter-gatherers from southwest Europe are the closest to the Iberian Neolithic datasets. The group of hunter-gatherers from Central/North Europe are somewhere between the two other foraging metapopulations, it's position can only be anchored by the Neolithic forager Pitted Ware population from the Island of Gotland (Figure 14).

The affinity of the Carpathian Basin Neolithic groups to the populations of the Central European populations in the sixth-fourth millennia BC is remarkable in the result of each performed analysis (Figure 14, 15, 18). Besides the five Transdanubian Neolithic datasets, the Central European distribution of the LBK, Rössen, Schöningen, Baalberge, and Salzmünde populations belong to this Early/Middle Neolithic cluster.

Along the Continental Route of the Neolithic dispersal, the common characteristics of the STA, LBKT and LBK datasets support the migration of the LBK populations from the western Carpathian Basin, as it has been described previously by our team (Szécsényi-Nagy et al.,

2014a). The migration of the first farmers through Transdanubia to Central Europe is suggested by the PCA, MDS, F_{st}, TPC, and ASHA analyses of this study as well (Figure 18-21, Table 7). Consequently to the swiftness of the dispersal between the Carpathian Basin and the region of today's Germany, I suggest to call the LBK dispersal a "folk migration" [term of Zvelebil, (Zvelebil, 2001)], rather than a demic diffusion [term of Ammerman and Cavalli-Sforza, (Ammerman and Cavalli-Sforza, 1984)]. The demonstrated genetic homogeneity of the LBK across different geographic regions also support its spread from one source region, through migration of people, who took the technological innovations and domesticates along to the new territories (Szécsényi-Nagy et al., 2014b). The people of the LBK(T) either migrated to northwards along the Danube River (Marton and Oross, 2010) or from the Drava valley along the western margin of Transdanubia to the Vienna Basin (Bánffy, 1996b). Interestingly, this latter route could be subsequently used during the Late Neolithic, Early Chalcolithic period as well (Bánffy, 2002). Nevertheless, defining the source region of the LBK(T) migration in the western Carpathian Basin is still an unresolved demographic issue, because the to date described formative and earlier LBKT sites in the western Carpathian Basin do not represent such large LBKT population, which would support the colonisation of Central Europe. Even the model of "leapfrog colonisation" [term of Zvelebil, (Zvelebil, 2001)] is a good explanation to the spread of the LBK by small groups of people, the genetic data in Central Europe speaks for a reversed hunter-gatherers and farmers relation. Recent genome studies estimate the hunter-gatherer ancestry of one LBK farmer from Stuttgart to be between 0-45% (Lazaridis et al., 2014), which is a quite inexact proportion, but still speaks for a massive migration of the early farmers. The famers built the dominant communities in the sixth millennium BC Central Europe, and the signs of the forager inhabitants, probably lived on in small enclaves) resurrected only in the Middle/Late Neolithic period of the region (Brandt et al., 2013; Haak et al., 2015).

From the Transdanubian populations, only the Chalcolithic Balaton-Lasinja shows inconsistent genetic relatedness to the contemporaneous and neighbouring populations. This can be due to the small sample size, but the re-emergence of the haplogroup T1 in the Middle Chalcolithic Transdanubia, together with the continuity of the Late Neolithic U8 haplogroup can cause the Neolithic-Chalcolithic differences as well. The second closest cluster to the sixth-fourth millennia BC datasets is a group of populations from the third-second millennia BC Central Europe. The Bernburg, Corded Ware, Bell Beaker, and Únětice mtDNA datasets

build a separate cluster in the AMOVA analyses as well (Supplementary Table 16). These cultures witnessed three migration events in the Late Neolithic-Early Bronze Age of Central Europe, affected by influx from the north (event B2 during the Bernburg culture), east (event C in the Corded Ware), and the southwest (event D in the Bell Beaker) (Brandt et al., 2013). Therefore, the third-second millennia BC Central European cluster can hardly be characterised as uniform. They have some common mtDNA haplogroups (T1, U2, U4, and I) with the Eastern European and Western Asian Late Neolithic and Bronze Age populations. Consequently, emergence of these haplogroups in Central Europe has been connected to the eastern origin of the Corded Ware culture (Brandt et al., 2014; 2013). The southern French Treilles population, bearing a mixed mtDNA characteristic, joins rather to the sixth-fourth millennia BC cluster than to the Iberian Neolithic groups (Figure 14, 15, 17). These latters have a clearly different mtDNA gene pool from that of the Transdanubian populations (bearing southwestern hunter-gatherer types such as U5b, H and N*). Nevertheless, they also show haplogroups that are uniformly characteristic for the early farmers, such as T2, X and K. Both the mtDNA and the Y-chromosome data (i.e. NRY haplogroup G2a, discussed in chapter 4.4.1) suggest the common origin of the west Mediterranean and the Carpathian Basin-Central European early farmers, who reached and superimposed different Mesolithic substrates (a summary of these coherences is given by our team in (Brandt et al., 2014).

Summarizing the prehistoric comparative aDNA analyses, the western Carpathian datasets show affinity to the sixth-fourth millennia BC populations of the closest studied region, Central Europe. The Neolithic populations from the Alföld were not discussed here, but they also form the part of the Carpathian Basin-Central-European sixth-fourth millennia BC cluster (detailed in Keerl, 2014).

6.6 Genetic evidence for the origin of the first farmers in the Carpathian Basin

To date, there are only a few and partial Near Eastern ancient HVS-I data published, which predate the Starčevo samples in my study (Fernández et al., 2014). Due lack of available information, comparative haplotype analyses were not possible with them. Looking at the haplogroups, the Near Eastern farmers from the Pre-Pottery Neolithic B (PPNB) show more affinity with the western Mediterranean Neolithic datasets than with the genetic package of the Continental Route. The east-west connection in the Mediterranean is mainly based on the

shared haplogroup N* (Supplementary Figure 2). One PPNB K haplotype (16224C 16311C) occurs in the Neolithic Carpathian Basin as well, but it is one of the most common haplotypes in our complete prehistoric dataset. This K appears both in Central Europe and in the Iberian Peninsula, and so this haplotype match is rather uninformative. The Syrian PPNB farmer population contains further mtDNA haplogroups (L3, R0), which probably did not dispersed to Europe, or have not been detected in the study region so far.

Lacking appropriate comparative ancient DNA datasets, I used modern genetic data as comparison to the Neolithic genetic variation, and calculated F_{st} values and performed PCAs with them. Such populations were searched, which can serve as proxy to the hypothetical source populations of the earliest farmers. I used comprehensive population genetic analyses (PCA, hierarchical clustering, F_{st} calculation and distance mapping) with modern North African and Eurasian populations for tracing the genetic origins (and observing genetic legacy) of the farming Transdanubian populations. Comparing the Starčevo mitochondrial DNA and Y-chromosomal dataset to the modern populations, a southwestern Asian origin of the first farmers become supported. The Starčevo data shows high affinities to modern Near East and South Caucasus (Figure 30) from the maternal side, and connections to the West Caucasus and southeastern Europe from the paternal viewpoint. The subsequent populations reveal more affinity to modern Europe; especially the people of the two Late Neolithic cultures have small F_{st} values from the extant Europeans. The genetic affinities of the Middle Chalcolithic Balaton-Lasinja indicate a new genetic input to the Carpathian Basin from southeastern Europe, through a freshly emerged Near Eastern affinity.

Comparing my results to recent genome-wide studies, the Southwestern Asian connection of the early farmers are more expressed by my mitochondrial and Y-chromosomal DNA results, than by the genome-wide autosomal markers (Gamba et al., 2014; Lazaridis et al., 2014). Ancient genome studies have repeatedly stressed the modern southeastern European affinities of the first farmers, noting the Sardinians as an especially similar population to the Neolithic farmers. An explanation for that phenomenon is that they probably remained in an ancient genetic composition by isolation during the last millennia. Only the NRY distance map of the merged STA-LBKT-LBK dataset shows a possible connection to the Sardinians (Figure 37), none of the mtDNA distance maps indicate such connection. The Y-chromosomal haplogroup G is frequent in the modern Sardinian population, which can cause its Neolithic affiliation. Nevertheless, detailed analysis of the G2a haplotypes (STR or

complete chromosome) would be necessary to compare the Neolithic Carpathian Basin to modern Sardinians or to the prehistoric West European, Iberian Neolithic data (Lacan et al., 2011a, 2011b). By a shared haplotype analysis of a Spanish Neolithic G2a haplotype, ancient G2a lineage has been found to be rare in modern populations (Lacan et al., 2011b). It raises the question, on what extend are the Carpathian Basin G2a haplogroups related to the modern Sardinian or Caucasian G2a Y-chromosome types?

Uniparental markers are more conservative, being not affected by admixture as much as autosomal markers, which can cause other affinity patterns to modern populations as it has been shown by genome studies (Gamba et al., 2014; Haak et al., 2015; Lazaridis et al., 2014). It means that a small extent of admixture between farmers and local European hunter-gatherers might not resurge in the mtDNA variation of the early farmers, but it certainly had a genetic imprint on their autosomes. This can be a reason, why genome studies have not detected pronounced Near Eastern affinities of the first farmers' gene pools. On the other hand, the Near East was witness of several population replacements and major migration events during the last millennia, which could affect the autosomal gene pool more than the uniparentally inherited systems. Ancient DNA genomes from the Near East will certainly unravel this issue in the future.

The Caucasus, especially the West Caucasus was stressed by the affinity of the early Neolithic NRY data. However, is not considered to be a part of the primary Neolithisation zone (Özdoğan, 2008). The Neolithic genetic connection between the Carpathian Basin and the Caucasus might reflect genetic drift, caused by isolation and small effective population size after a direct gene flow from the Near East, which lead to a fixation of NRY haplogroup G2a (Balanovsky et al., 2011).

As shown in the previous paragraphs, the genome wide analyses cannot reveal the Near East as a source region, when the farmers are compared to modern datasets (Haak et al., 2015). The ascertainment in this question can only be expected from archaic genomic data from southeastern European and Near Eastern Early Neolithic (or also PPN) populations.

6.7 Genetic legacy of the Neolithic populations

Recent study of our team has shown that the Late Neolithic and Early Bronze Age population genetic events had predominant role in the formation of the extant Central

European gene pool (Brandt et al., 2013). Bayesian simulation has defined continuity from the Early Bronze Age to the modern age. From the modern Central European haplogroups a proportion of 31.2% can be assigned to the Early Neolithic, 16% to the pre-Neolithic, and 5.8% to the Late Neolithic. Nevertheless, a significant proportion (47%) needs further resolution, and could not be assigned to any period with certainty (Brandt et al., 2013).

Genome study from the Alföld region has defined genomic differences between the sequenced Neolithic/Copper Age and Bronze Age individuals. The studied two Bronze Age individuals (from the Makó and Kyatice cultures) fall into modern Central European genotypes, contrastingly to the southeastern affinities of the earlier individuals (Gamba et al., 2014). However, the formation of the extant Central European or Carpathian Basin gene pool certainly did not finish in the Bronze Age. For example, one Iron Age (Scythian) sample shows new influx from the Eurasian steppe, reflected also in the mtDNA and Y-chromosomal haplogroups (Gamba et al., 2014). Furthermore, X-XIth century mtDNA data from Hungary show new haplogroups as well (e.g. haplogroups B, M, D), which were unseen in the prehistoric datasets of the region (Bogácsi-Szabó et al., 2005; Tömöry et al., 2007). Thinking of the Roman or Migration period of the Carpathian Basin's prehistory, the genetic influxes from various regions could have been uncountable, sequentially contributing to the formation of the Carpathian Basin's modern gene pool.

Regarding the shared haplotype analyses in this study, certain conclusions can be drawn. Neolithic lineages were found in 38.38% of the modern maternal gene pool of the Carpathian Basin, which is a similar proportion to the Central European ASHA results of G. Brandt and our team (Brandt et al., 2014). However, the exact amount of the Neolithic contribution to the modern gene pool cannot be estimated by mtDNA HVS-I haplotypes alone. A new genome study of our extended team suggests instead that a Neolithic proportion of 30-40% in the modern European gene pool would be an overestimation (Haak et al., 2015).

This PhD work was not extended to the Late Chalcolithic and Early Bronze Age periods of Transdanubia, and so a diachronic comparison with the Brandt et al. study focusing on 4,000 years of prehistory was not entirely possible. On the other hand, some hints of NRY evidence from the Bronze Age Transdanubia presume major population genetic event(s) in the Carpathian Basin Early Bronze Age as well, along with the appearance of the NRY haplogroups R1b in Transdanubia (Table 10). The two detected haplogroups R1b, from two different cultural contexts coincide with the Bronze Age evidences of the Gamba et al. study.

Summarizing the above presented results, the persistence of a portion of the Neolithic genetic substrate (especially the more conservative uniparentally inherited systems) is assumable, though the region of Transdanubia was the stage of several migration events during the last 6,000 years. Further diachronic studies with genome wide focuses could enlighten the population history of the post-Neolithic periods.

6.8 Interdisciplinary outlook: demographic and social implications of the results

When the processes of the Mesolithic/Neolithic transition and LBK dispersal are considered, the implications of Neolithic demographic studies have to be integrated into the line of argumentation. The evidence of rapid LBK migration into Central Europe raises the question, whether this dispersal had a demographic base (large population growth) in the source region. A sufficient population size for a colonisation event had to exist in the western Carpathian Basin, because the small proportion of hunter-gatherer genetic component of the LBK in Central Europe argues for a massive LBK migration.

Demographic simulations have not supported the demic diffusion, or the LBK colonisation of Central Europe (Galeta and Bruzek, 2009; Galeta et al., 2011). Nonetheless, the LBKT settlement research records of the last decade (Oross, 2013, p. 164-187.) have not been involved in these studies. To date, more than 64 LBKT settlements are known or described from Transdanubia, which let a larger LBKT population be presumed than supposed a decade ago. However, detailed chronological analysis of the sites should select the traces of the early LBKT populations. Furthermore, the number of LBKT graves uncovered in Transdanubia, are still not corresponding to the presumed population size. They were all found on settlements, and no formal LBKT cemetery is known from western Hungary (Oross and Marton, 2012). The question, whether these burials represent a special status of the LBKT society, cannot be answered yet. The occurrence of grave goods in the late LBKT period contradicts the theory, that only low prestige individuals would have been buried in settlement pits (Oross and Marton, 2012). Most of the graves (especially in the early LBKT period) do not contain any grave goods, which might indicate that the grade of the social differentiation was not high in the early LBKT phases. Even if archaeological records cannot verify the causes of the above-described phenomena, when investigating the skeletal remains from STA and LBKT settlement

burials, we should not neglect the possibility that these individuals might not represent the community that lived on the surrounding settlement.

The social structure of the cemeteries changed in the Late Neolithic period of Transdanubia. The early hints of social differentiation in the Vinča and Sopot cultures have been summarised by I. Zalai Gaál (Zalai-Gaál, 2010). Formal cemeteries became common in the Lengyel period of Transdanubia, with increased number of graves and signs of social differentiation (Dombay, 1960; Osztás et al., 2012; Zalai-Gaál, 2010, 2002, 1988). The detected number of possible maternal and paternal relationships increased in the Sopot and especially in the Lengyel culture's cemeteries (Supplementary Table 3, 5). According to the here presented data, the burials' distribution within a graveyard still did not follow a genetically observable kinship mediated rule, but the grave groups were seemingly used by one community or group of families (this was especially good observable in the case of Alsónyék-elkerülő, Bátaszék-Lajvér, and Veszprém-Jutasi út)[1].

The detected high mtDNA diversity is remarkable in these Neolithic communities. However, on the largest Starčevo site Bátaszék-Alsónyék, the heterogeneity of the mtDNA haplotypes can also be the consequence of different settlement events, since the people of the Starčevo culture lived there in at least two phases, arching over ~300 years (T. Marton personal comm., see Supplementary Table 1). In the case of the Sopot Alsónyék-elkerülő graveyard in contrast, the five radiocarbon data indicate a rather short lifespan of the cemetery (Supplementary Table 1). Speaking about this latter and the Lengyel formal cemeteries, the mtDNA heterogeneity can be an indicator of residential rules and marital systems. It especially supports the theory of patrilocal residential rules, which has been suggested by aDNA (Lacan et al., 2011a) and isotopic studies (Bentley et al., 2012; Haak et al., 2008) for different Neolithic communities. The Europe-wide homogeneity of the early farmers' paternal gene pool (dominance of G2a haplogroup) along with the observed high mtDNA haplogroup and haplotype diversities suggest the patrilocality and patrilineality as basic rules of the early farming societies (Szécsényi-Nagy, 2014a). The patrilocality was probably coupled with exogamy or wife exchange between farming communities, which is a possible way of keeping the mtDNA diversity long-term on the observed high level (Table 6). These results are concordant to the estimations of model-based inferential techniques. The sedentism of the farmers possibly led to a decrease in male gene flow, whereas female gene flow would either

[1] The detailed analyses of these case studies were performed for conference and workshops papers; however, their presentation here would be apart from the framework of the thesis.

have remained constant or would have increased by exogamy or patrilocality according to R. Rasteiro and L. Chikhi (Rasteiro and Chikhi, 2013). The paternal genetic diversity raised in the Late Neolithic, but the maternal genetic diversity remained on a high level.

The process of “frontier mobility”, described by M. Zvelebil (Zvelebil, 2001) in the farmer-forager interacting zones could have existed to moderate extent at most. MtDNA signals of possible intermarriage of hunter-gatherer women are very scarce, though it has been considered as plausible by M. Zvelebil. However, complete mitochondrial genome analysis would let a deeper insight into partially hidden genetic formation of the LBKT as well. Instead of forager women’s mobility, one NRY haplogroup I2a1 could possibly indicate intermarriage of forager men into farming communities. Similarly, from the restricted heterozygosity and hunter-gatherer affinity of a male Körös individual, Gamba et al. have inferred that KO1 was an exogenous forager individual in the farming community (Gamba et al., 2014). This individual lived on a Körös site in the supposed forager-farmer northern Alföld interaction zone, on the margin of the Körös culture’s dissemination area (Domboróczki et al., 2010). Intriguingly, to date archaeologically undefinable (and certainly not dense) Late Mesolithic/Early Neolithic forager population becomes little by little indirectly explored through the genetic evidence. It can be expected from further whole genome studies of the Starčevo and LBKT specimens to shed new light on the genetic legacy and connections of the early farmers of Transdanubia.

7 Concluding remarks

This population genetic study, joining to the Neolithic archaeogenetic project of the Central German Middle-Elbe-Saale region (dissertation of Guido Brandt, 2014) and to the molecular genetic analyses of the Neolithic eastern Hungary (dissertation of Victoria Keerl, 2014) opens a new era in the research history of palaeogenetics. The enormous quantity of prehistoric mtDNA results (n=364 by G. Brandt and W. Haak, n=245 by V. Keerl, and n=298 by myself), enables meaningful performance of prehistoric comparative population genetic analyses. With the so far published other ancient mtDNA data (n= ca. 400), there are more than 1,300 available prehistoric HVS-I haplotypes from across Europe by now, which allow archaeogenetics independent data evaluation in a European-wide prehistoric context. However, comparative analyses with modern mtDNA data are still needed, when examining the southern European, Near Eastern origin of the first farmers, due to the lack of southeastern aDNA data. Furthermore, ancient Y-chromosome results are still too scarce (n= ca. 50) for relevant comparative aDNA analysis.

During my PhD work, I often encountered the problem of archaeological definition of a certain group or community, which are mostly named after the dominating material culture traits. These assemblages of artefacts mostly vary and transform gradually over space and time, splitting up the traditional conception of Neolithic relative chronology of successive cultures and their phases. Since archaeologists often face mixed assemblages, which cannot be ordered to a distinct cultural unit, the choice of an operational unit at the population genetic and biostatistical analysis is crucial. In order to avoid the biasing effect of archaeological preconceptions, all studied individuals should have been dated by ^{14}C method. Individual dates combined with geographic locations could provide a solid basis for an independent grouping. Nevertheless, this approach was financially not achievable. However, the thorough evaluation of the archaeological context remained (done mostly by J. Jakucs), combined with occasional radiocarbon dating (Supplementary Table 1).

The field of archaeogenetics has advanced significantly In the last few months, and the first complete genomes from the European Mesolithic and Neolithic period have been published as well (Gamba et al., 2014; Haak et al., 2015; Lazaridis et al., 2014). Whole-genome data of an individual contains information about thousands of that individual's ancestors, and not only information about a single ancestral (maternal or paternal) lineage, as it is the case

with the mtDNA and Y-chromosome. Therefore, it is important to combine the mtDNA and Y-chromosome studies with analysis of autosomal DNA. Continuing with the autosomal analyses on a selection of the presented samples, a more precise definition of the ancestral genetic components of the studied populations will become possible. Although the typing of the uniparentally inherited genetic systems is an important step reconstructing population history of the Carpathian Basin, several open questions remain for further whole genome analyses to answer.

8 Summary

This PhD work is a part of a large interdisciplinary, DFG (German Research Foundation) funded project, entitled as “Population history of the Carpathian Basin in the Neolithic period and its influence on the colonisation of Central Europe.” The focus of my study is the Neolithic in the western Carpathian Basin, more precisely the western part of today’s Hungary, which also called as Transdanubia.

The aims of this work are to study the genetic diversity of the Neolithic and Early/Middle Chalcolithic cultures’ populations (sixth-fifth millennia BC) in Transdanubia from both the mtDNA and Y-chromosome perspectives, and to specify the main population genetic events during this period. Closer observing the Transdanubian sample set, the genetic regional differences are also examined. Further topic is the genetic investigation of the Mesolithic/Neolithic transition and the origin of the southeastern European Starčevo farmers. I also aim to compare the Transdanubian results with prehistoric data, especially with the European pre-Neolithic and the Central European Neolithic datasets, searching for the genetic contribution of the Transdanubian populations to the variability of the LBK and its successor populations in Central Europe. The parallel analyses of the mtDNA and the Y-chromosome raise the question, whether men and women had different migration patterns or Neolithisation histories.

Altogether 32 Mesolithic, Neolithic and Chalcolithic archaeological sites were included in the sampling, which encompassed 323 individuals or skeletal remains from the western Carpathian Basin. Samples from 298 individuals were processed in the archaeogenetic laboratories of the Johannes Gutenberg University of Mainz. Following strict standards of the clean laboratory work and reproducing the results from at least two samples per skeleton, authenticated aDNA haplotype of the mtDNA HVS-I region were obtained in 256 cases. Furthermore, the HVS-II region sequences of 80 individuals were reproduced, detecting potential intra-site maternally kinship relations. Endogenous aDNA sequences were verified through cloning process. The mtDNA haplogroup definitions were ascertained through the analysis of 22 mtDNA coding region polymorphisms. Screening all well-preserved samples, the Y-chromosomal haplogroups were defined in 33 Neolithic and Chalcolithic individuals. For the population genetic analyses, large sets of comparative aDNA, modern mtDNA, and Y-

chromosome data were collated and used. The results were evaluated with a suite of population genetic analyses (Fst analysis, PCA, MDS, ASHA, GDM, TPC).

It can be inferred from the mitochondrial and Y-chromosomal ancient DNA data that at the advent of the Neolithic both farmer men and women, originated from the Near East, migrated into the Carpathian Basin. The first Neolithic Starčevo culture's people had a remarkable mtDNA variability, which were largely transmitted to the succeeding populations. The hunter-gatherers show negligible contribution in the early farmers' maternal and paternal gene pool. The emergence and spread of the Central European LBK can be genetically traced back to the western Carpathian Basin (population of the LBK in Transdanubia or LBKT), corresponding to most archaeological assumptions (Szécsényi-Nagy et al., 2014a). The regional LBK genetic datasets show homogeneity in the maternal gene pool across large distances in Europe (Szécsényi-Nagy et al., 2014b). The genetic effect of the Starčevo population was still significant in the maternal gene pool of the Late Neolithic of Central Germany. The Early Neolithic genetic substrate dominated the Neolithic of Transdanubia as well; only few hints indicate smaller infiltration or immigration events during the Vinča, LBKT, Sopot, and Balaton-Lasinja periods. The complete Transdanubian mtDNA dataset closely affiliated to the Neolithic populations in eastern Hungary (Keerl, 2014), and the Central European sixth-fourth millennia BC populations (published by Brandt et al., 2013).

The close maternal genetic affinity of the Neolithic Transdanubian populations to each other got new shades when the LBKT and Lengyel datasets were split into north and south groups. The here assumed north-south genetic difference in the Middle and Late Neolithic Transdanubia should be further investigated by whole genome studies and more balanced sample distribution.

The observed heterogeneity of the Starčevo and LBKT maternal gene pool coupled with Y-chromosomal homogeneity in the early farming populations suggest the residential rule of patrilocality and patrilineality in these communities. The paternal diversity though raised in the Late Neolithic, the high maternal diversity still support continuous social system from the Early Neolithic onward.

This study presents the first detailed population genetic survey, with an exceptionally large number of ancient DNA data of the sixth-fifth millennium BC western Carpathian Basin, which was a corridor on the Continental Route of the European Neolithisation. My approach integrates biological, cultural and demographic lines of evidence, underlining relations

between social and genetic structures of human populations. Even if the described results in this thesis mark a milestone in the archaeogenetic research of the Carpathian Basin's prehistory, several questions remain open for further ancient genomic analyses.

9 Zusammenfassung der Dissertation

Die vorliegende Dissertation ist Teil des interdisziplinären Forschungsprojektes mit dem Titel *„Bevölkerungsgeschichte des Karpatenbeckens in der Jungsteinzeit und ihr Einfluss auf die Besiedlung Mitteleuropas“,* das zwischen 2010 und 2014 von der DFG gefördert wurde. Der Fokus dieser Studie ist die populationsgenetische Analyse der möglichen neolithischen Bevölkerungswechsel in den 6.-5. Jahrtausende vor Christus, die im westlichen Karpatenbecken, insbesondere im westlichen Teil des heutigen Ungarns (Transdanubien) stattfanden.

Die Zielsetzung der Dissertation war, mittels der Analyse von mitochondrialer und Y-chromosomaler aDNA den Genpool der neolithischen und kupferzeitlichen Populationen des westlichen Karpatenbeckens zu untersuchen und die daraus resultierenden Ergebnisse mit anderen prähistorischen, genetischen Daten [besonders aus der Mittelelbe-Saale Region (Brandt et al., 2013)] zu vergleichen. In diesem Zusammenhang wurden folgende Fragestellungen untersucht: (1) Woher stammen die ersten Bauern Transdanubiens aus genetischer Sicht? (2) Besteht genetische Kontinuität zwischen der mesolithischen und der neolithischen Population in Transdanubien? (3) Wie entwickelte sich die genetische Variabilität der transdanubischen Populationen während den sechs neolithischen und kupferzeitlichen kulturellen Perioden? (4) Können populationsgenetisch nachweisbare Regionalgruppen im neolithischen Transdanubien identifiziert werden? (5) Gab es frühneolithischen (Starčevo) genetischen Einfluss aus der Untersuchungsregion auf die Besiedlung Europas? (6) Unterscheidet sich die genetische Zusammensetzung der neolithischen Bevölkerung in Transdanubien von den benachbarten und entfernteren europäischen, prähistorischen Populationen? (7) Hatten Männer und Frauen eine vergleichbare Neolithisierungsgeschichte? Zeigen sie ähnlichen Migrationsmustern in ihren Genen?

Insgesamt wurden 323 Individuen aus 32 ungarischen, kroatischen und slowakischen Fundplätzen aus der Region des westlichen Karpatenbeckens beprobt und in den archäogenetischen Laboren der Johannes Gutenberg-Universität in Mainz bearbeitet. Entsprechend der standardisierten Protokolle des Labors in Mainz, wurden die Ergebnisse der aDNA-Analyse durch wiederholte DNA-Amplifikationen, Sequenzierungen und Klonierungen authentifiziert. Hypervariable Segmente (HVS-I und teilweise HVS-II) des mitochondrialen

Genoms wurden typisiert, in Kombination mit weiteren multiplexen Analysen von Polymorphismen, die sich in dem kodierenden Bereich des mitochondrialen Genoms befinden. Die mitochondrialen Analysen wurden durch die Typisierung von 25 Y-chromosomalen biallelischen Polymorphismen ergänzt.

Die HVS-I Region und 22 Polymorphismen des kodierenden Bereichs der mitochondrialen DNA konnten bei 256 Individuen reproduziert und authentifiziert werden (mit einer Erfolgsrate von 85.9%). Die Typisierung der HVS-II Region war in 80 Fällen erfolgreich, wodurch potenzielle maternale Verwandte innerhalb der Fundplätzen identifiziert werden konnten. Testend alle gut erhaltene Proben, die Y-chromosomale Haplogruppe konnte in 33 männlichen Individuen typisiert werden.

Die DNA Ergebnisse wurden mit verschiedenen populationsgenetischen Methoden (Hauptkomponenetanalyse, F_{st}-Berechnung, MDS, genetischen Distanzkarten, ASHA, TPC etc.) ausgewertet. Vergleichsdaten von prähistorischen und modernen eurasiatischen Populationen wurden dazu gesammelt.

Die neolithischen, mitochondrialen Haplogruppen deuten auf eine hohe Variabilität des maternalen Genpools hin. Sowohl die mitochondrialen als auch die Y-chromosomalen Daten lassen Rückschlüsse auf eine nah-östliche bzw. südwestasiatische Herkunft der Starčevo Population zu, wobei für Männer und Frauen von einer vergleichbare Migrationsgeschichte während der Neolithisierung auszugehen ist. Im Vergleich zur mittel- und nordeuropäischen Jäger und Sammler-Ursprungsbevölkerung, brachte die erste neolithische Starčevo-Population viele neue mtDNA-Linien nach Transdanubien, die auch in den nachfolgenden Perioden fortbestanden. Die Starčevo- und linearbandkermaischen- Populationen in westlichem Karpatenbecken (letztgenannt abgekürzt als LBKT) und die linearbandkermaischen-Population in Mitteleuropa (LBK) haben so starke Ähnlichkeit auf den mitochondrialen und Y-chromosomalen Ebenen, dass die Verbreitung der LBK nach Mitteleuropa entsprechend der meisten archäologischen Annahmen mit vorangegangenen Wanderungsereignissen zu erklären ist (Szécsényi-Nagy et al., 2014a). Die regionalen LBK-Gruppen weisen europaübergreifend einen homogenen maternalen Genpool auf (Szécsényi-Nagy et al., 2014b). Das noch im Spätneolitikum Mitteleuropas nachweisbare genetische Erbe der LBK-Bevölkerung kann daher mit hoher Wahrscheinlichkeit bis auf die Starčevo–Population zurückgeführt werden. Vergleichend die Transdanubische aDNA Daten mit den

publizierten prähistorischen aDNA Datensätzen von Europa, zeigt das Karpatenbecken (Alföld inbegriffen) hohe Affinität zu den Daten aus 6-4. Millennium vor Chr. in Mitteleuropa.

Die maternal-genetische Variabilität der Starčevo-Population konnte auch innerhalb der nachfolgenden Populationen Transdanubiens festgestellt werden. Nur kleinere Infiltrationen und Immigrationsereignissen konnten während der Vinča-, LBKT-, Sopot- und Balaton-Lasinja-Kultur in Transdanubien identifiziert werden.

Zwischen den transdanubischen Regionen konnten mögliche maternal-genetische Unterschiede nur in der LBKT und in der Lengyel-Periode beobachtet werden, als sich die nördlichen Gruppen von den südlichen Populationen trennten. Diese These sollte im Rahmen zukünftiger Forschungsprojekte mit Genomanalysen und regional ausgegliederter Probenverteilung überprüft werden.

Die festgestellte Heterogenität der mtDNA in Zusammenhang mit der Y-chromosomalen Homogenität in den Starčevo- und LBK-Populationen, weisen auf patrilokale Residenzregeln und patrilineare Abstammungsregeln in den ersten Bauergemeinschaften hin. Obwohl die väterliche genetische Diversität im Verlauf des Spätneolitikums anstieg, stützt die kontinuierliche mütterliche genetische Variabilität eine andauernde Gesellschaftsform im Neolithikum Transdanubiens.

Diese Studie stellt einen exzeptionellen großen mtDNA-Datensatz aus dem 6.-5. Jahrtausend vor Christus des westlichen Karpatenbeckens vor. Während der Neolithisierung war Transdanubien ein wichtiger Teil der kontinentalen Wanderroute, der eine Art Korridor auf dem Weg der ersten Bauern nach Mitteleuropa dargestellt haben könnte. Die Forschungsmethoden integrieren biologische, kulturelle und demographische Belege, und ermöglichen eine integrative Auswertung und Diskussion der Ergebnisse. Obwohl die hier präsentierten Daten einen großen Fortschritt in der Forschung von aDNA und Neolithikum des Karpatenbeckens und Mitteleuropas bedeuten, werfen sie auch mehrere Fragen auf, deren Beantwortung durch zukünftige Genomforschungen erbracht werden könnte.

10 References cited in the text

Abu-Amero, K.K., Larruga, J.M., Cabrera, V.M., González, A.M., 2008. Mitochondrial DNA structure in the Arabian Peninsula. BMC evolutionary biology 8, 45.

Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., Moral, P., Dugoujon, J.-M., Roostalu, U., Loogväli, E.-L., Kivisild, T., Bandelt, H.-J., Richards, M., Villems, R., Santachiara-Benerecetti, A.S., Semino, O., Torroni, A., 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. American Journal of Human Genetics 75, 910–8.

Adler, C.J., Dobney, K., Weyrich, L.S., Kaidonis, J., Walker, A.W., Haak, W., Bradshaw, C.J.A., Townsend, G., Sołtysiak, A., Alt, K.W., Parkhill, J., Cooper, A., 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nature Genetics 45, 450–455.

Adler, C.J., Haak, W., Donlon, D., Cooper, A., 2011. Survival and recovery of DNA from ancient teeth and bones. Journal of Archaeological Science 38, 956–964.

Alt, K.W., Knipper, C., Peters, D., Müller, W., Maurer, A.-F., Kollig, I., Nicklisch, N., Müller, C., Karimnia, S., Brandt, G., Roth, C., Rosner, M., Mende, B., Schöne, B.R., Vida, T., von Freeden, U., 2014. Lombards on the Move - An Integrative Study of the Migration Period Cemetery at Szólád, Hungary. PloS ONE 9, e110793.

Ammerman, A., Cavalli-Sforza, L., 1971. Measuring the rate of spread of early farming in Europe. Man New Series 6, 674–688.

Ammerman, A.J., Cavalli-Sforza, L., 1973. A population model for the diffusion of early farming in Europe. In: The Explanation of Culture Change. Duckworth, London, pp. 343–357.

Ammerman, A.J., Cavalli-Sforza, L.L., 1984. The Neolithic transition and the genetics of populations in Europe. Princeton University Press, Princeton (New Jersey).

Anders, A., Siklósi, Z. (Eds.), 2012. The First Neolithic Sites in Central/South-East European Transect, Volume III. The Körös Culture in Eastern Hungary, BAR-IS 233. ed. Archaeopress, Ofxord.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R., Young, I.G., 1981. Sequence and organization of the human mitochondrial genome. Nature 290, 457–465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nature Genetics 23, 147.

Antoni, J., 1982. Régészeti kutatások a Dunántúlon. Őskori elődeink Csabiban., A Komárom . ed, A Komárom Megyei Múzeumok és a Fejér Megyei Múzeumok Régészeti kiállítása. Tata.

Bagolini, B., Barfield, L., 1991. The European Context of Northern Italy during the Third Millenium. In: Lichardus, J. (Ed.), Die Kupferzeit Als Historische Epoche. Symposium Saarbrücken Und Otzenhausen 6-13. 11. 1988. R. Habelt, Bonn, p. 906.

Bailey, D.W., 2002. Balkan Prehistory, 2. ed. London, New York.

Bailey, G., Spikins, P. (Eds.), 2008. Mesolithic Europe. Cambridge University Press.

Balanovsky, O., Dibirova, K., Dybo, A., Mudrak, O., Frolova, S., Pocheshkhova, E., Haber, M., Platt, D., Schurr, T., Haak, W., Kuznetsova, M., Radzhabov, M., Balaganskaya, O., Romanov, A., Zakharova, T., Soria Hernanz, D.F., Zalloua, P., Koshel, S., Ruhlen, M., Renfrew, C., Wells, R.S., Tyler-Smith, C., Balanovska, E., 2011. Parallel evolution of genes and languages in the Caucasus region. Molecular Biology and Evolution 28, 2905–20.

Balaresque, P., Bowden, G.R., Adams, S.M., Leung, H.-Y., King, T.E., Rosser, Z.H., Goodwin, J., Moisan, J.-P., Richard, C., Millward, A., Demaine, A.G., Barbujani, G., Previderè, C., Wilson, I.J., Tyler-Smith, C., Jobling, M.A., 2010. A predominantly neolithic origin for European paternal lineages. PLoS Biology 8, e1000285.

Balen, J., Potrebica, H., 2006. Nalazište Radovanci u Požeškoj kotlini. Od Sopota do Lengyela- Annales Mediterranea, Koper 21–27.

Ballantyne, K.N., Ralf, A., Aboukhalid, R., Achakzai, N.M., Anjos, M.J., Ayub, Q., Balažic, J., Ballantyne, J., Ballard, D.J., Berger, B., Bobillo, C., Bouabdellah, M., Burri, H., Capal, T., Caratti, S., Cárdenas, J., Cartault, F., Carvalho, E.F., Carvalho, M., Cheng, B., Coble, M.D., Comas, D., Corach, D., D'Amato, M.E., Davison, S., de Knijff, P., De Ungria, M.C.A., Decorte, R., Dobosz, T., Dupuy, B.M., Elmrghni, S., Gliwiński, M., Gomes, S.C., Grol, L., Haas, C., Hanson, E., Henke, J., Henke, L., Herrera-Rodríguez, F., Hill, C.R., Holmlund, G., Honda, K., Immel, U.-D., Inokuchi, S., Jobling, M.A., Kaddura, M., Kim, J.S., Kim, S.H., Kim, W., King, T.E., Klausriegler, E., Kling, D., Kovačević, L., Kovatsi, L., Krajewski, P., Kravchenko, S., Larmuseau, M.H.D., Lee, E.Y., Lessig, R., Livshits, L.A., Marjanović, D., Minarik, M., Mizuno, N., Moreira, H., Morling, N., Mukherjee, M., Munier, P., Nagaraju, J., Neuhuber, F., Nie, S., Nilasitsataporn, P., Nishi, T., Oh, H.H., Olofsson, J., Onofri, V., Palo, J.U., Pamjav, H., Parson, W., Petlach, M., Phillips, C., Ploski, R., Prasad, S.P.R., Primorac, D., Purnomo, G.A., Purps, J., Rangel-Villalobos, H., Rębała, K., Rerkamnuaychoke, B., Gonzalez, D.R., Robino, C., Roewer, L., Rosa, A., Sajantila, A., Sala, A., Salvador, J.M., Sanz, P., Schmitt, C., Sharma, A.K., Silva, D.A., Shin, K.-J., Sijen, T., Sirker, M., Siváková, D., Skaro, V., Solano-Matamoros, C., Souto, L., Stenzl, V., Sudoyo, H., Syndercombe-Court, D., Tagliabracci, A., Taylor, D., Tillmar, A., Tsybovsky, I.S., Tyler-Smith, C., van der Gaag, K.J., Vanek, D., Völgyi, A., Ward, D., Willemse, P., Yap, E.P.H., Yong, R.Y.Y., Pajnič, I.Z., Kayser, M., 2014. Toward male individualization with rapidly mutating y-chromosomal short tandem repeats. Human Mutation 35, 1021–32.

Bánffy, E., 1994. A Balaton—Lasinja-kultúra leletei Balatonmagyaród—Homoki-dűlőről. Zalai Múzeum 5, 234–250.

Bánffy, E., 1996a. Vinča impacts on late neolithic/early Chalcolithic Transdanubia. In: The Vinča Culture. Its Role and Cultural Connections. International Symposium on the Vinca Culture, Its Role and Cultural Connections, Timișoara, Romania, October 1995. Timisoara, pp. 323–334.

Bánffy, E., 1996b. South West Transdanubia as a mediating area. On the cultural history of the Early and Middle Chalcolithic. In : B. M Szőke (Ed) Archaeology and settlement history in

the Hahót Basin. Antaeus Communicationes ex Instituto Archaeologico Academiae Scientarum Hungaricae 22, 157–196.

Bánffy, E., 2002. A uniqe southeastern vessel type from early chalcolithic Transdanubia: data on the “Western route”. Acta Archaeologica Academiae Scientarium Hungaricae 53, 41–60.

Bánffy, E., 2004. The 6th millennium BC boundary in western Transdanubia and its role in the Central European Neolithic transition (the Szentgyörgyvölgy-Pityerdomb settlement). Archaeological Institute of the Hungarian Academy of Science, Budapest.

Bánffy, E., 2006. Eastern , Central and Western Hungary – variations of Neolithisation models. Documenta Praehistorica 33, 125–142.

Bánffy, E., Brandt, G., Alt, K.W., 2012. “Early Neolithic” graves of the Carpathian Basin are in fact 6000 years younger-appeal for real interdisciplinarity between archaeology and ancient DNA research. Journal of human genetics 57, 467–9; author reply 470–1.

Bánffy, E., Eichmann, W.J., Marton, T., 2007. Mesolithic foragers and the spread of agriculture in western Hungary. In: Kozlowski, J.K., Nowak, M. (Eds.), Proceedings of the XV UISPP World Congress (Lisbon, 4–9 September 2006) Vol. 6. BAR IS 1726. Oxford, pp. 53–62.

Bánffy, E., Marton, T., Osztás, A., 2010. Early Neolithic settlement and burials at Alsónyék-Bátaszék. In: Kozlowski, J.K., Raczky, P. (Eds.), Neolithization of the Carpathian Basin: Northernmost Distribution of the Starčevo/Körös Culture. Kraków-Budapest, pp. 37–51.

Bánffy, E., Sümegi, P., 2012. The Early Neolithic Agro-Ecological Barrier in the Carpathian Basin: A Zone for Interaction. In: Anreiter, P., Bánffy, E., Bartosiewicz, L., Meid, W., Metzner-Nebelsick, C. (Eds.), Archaeological, Cultural and Linguistic Heritage. Festschrift for Erzsébet Jerem in Honour of Her 70th Birthday. Archaeolingua Alapítvány, Budapest, pp. 57–69.

Barbujani, G., Bertorelle, G., Chikhi, L., 1998. Evidence for Paleolithic and Neolithic gene flow in Europe. American Journal of Human Genetics 62, 488–92.

Barker, G., 1985. Prehistoric farming in Europe. Cambridge University Press, Cambridge.

Bar-Yosef, O., 1998. The Natufian Culture in the Levant, treshold to the Origins of Agriculture. Evolutionary Anthropology 159–177.

Battaglia, V., Fornarino, S., Al-Zahery, N., Olivieri, A., Pala, M., Myres, N.M., King, R.J., Rootsi, S., Marjanovic, D., Primorac, D., Hadziselimovic, R., Vidovic, S., Drobnic, K., Durmishi, N., Torroni, A., Santachiara-Benerecetti, A.S., Underhill, P.A., Semino, O., 2009. Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. European journal of human genetics : EJHG 17, 820–30.

Behar, D.M., Garrigan, D., Kaplan, M.E., Mobasher, Z., Rosengarten, D., Karafet, T.M., Quintana-Murci, L., Ostrer, H., Skorecki, K., Hammer, M.F., 2004. Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. Human Genetics 114, 354–65.

Behar, D.M., Harmant, C., Manry, J., van Oven, M., Haak, W., Martinez-Cruz, B., Salaberria, J., Oyharçabal, B., Bauduer, F., Comas, D., Quintana-Murci, L., 2012a. The Basque paradigm: genetic evidence of a maternal continuity in the Franco-Cantabrian region since pre-Neolithic times. American Journal of Human Genetics 90, 486–93.

Behar, D.M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, N.M., Kivisild, T., Torroni, A., Villems, R., 2012b. A “Copernican” reassessment of the human mitochondrial DNA tree from its root. American Journal of Human Genetics 90, 675–84.

Behar, D.M., Villems, R., Soodyall, H., Blue-Smith, J., Pereira, L., Metspalu, E., Scozzari, R., Makkan, H., Tzur, S., Comas, D., Bertranpetit, J., Quintana-Murci, L., Tyler-Smith, C., Wells, R.S., Rosset, S., 2008. The dawn of human matrilineal diversity. American Journal of Human Genetics 82, 1130–40.

Bentley, R.A., 2012. Mobility and the diversity of early Neolithic lives: Isotopic evidence from skeletons. Journal of Anthropological Archaeology 32, 303–312.

Bentley, R.A., Bickle, P., Fibiger, L., Nowell, G.M., Dale, C.W., Hedges, R.E.M., Hamilton, J., Wahl, J., Francken, M., Grupe, G., Lenneis, E., Teschler-Nicola, M., Arbogast, R.-M., Hofmann, D., Whittle, A., 2012. Community differentiation and kinship among Europe’s first farmers. Proc Natl Acad Sci U S A 109, 9326–30.

Berzsényi, B., Dálnoki, O., 2006. Plant cultivation and crop processing in the formative LBK settlement of Szentgyörgyvölgy–Pityerdomb in Transdanubia (6th Millennium BC). Antaeus Communicationes ex Instituto Archaeologico Academiae Scientarum Hungaricae 28, 261–270.

Biagi, P., Shennan, S., Spataro, M., 2005. Rapid Rivers and slow seas? New data for the radiocarbon chronology of the Balkan peninsula. In: Nikolova, L., Higgins, J. (Eds.), Prehistoric Archaeology & Anthropological Theory and Education. RPRP 6-7. pp. 31–34.

Bickle, P., Whittle, A. (Eds.), 2013. The first farmers of central Europe: Diversity in LBK lifeways. Oxford Books, Oxford.

Bocquet-Appel, J.-P., 2008. Explaining the Neolithic Demographic Transition. In: Bocquet-Appel, J.-P., Bar-Yosef, O. (Eds.), The Neolithic Demographic Transition and Its Consequences. Springer, pp. 35–57.

Bocquet-Appel, J.-P., 2011. When the world's population took off: the springboard of the Neolithic Demographic Transition. Science 333, 560–1.

Bocquet-Appel, J.-P., Demars, P.-Y., Noiret, L., Dobrowsky, D., 2005. Estimates of Upper Palaeolithic meta-population size in Europe from archaeological data. Journal of Archaeological Science 32, 1656–1668.

Bogácsi-Szabó, E., Kalmár, T., Csányi, B., Tömöry, G., Czibula, Á., Priskin, K., Horváth, F., Downes, C.S., Raskó, I., 2005. Mitochondrial DNA of ancient Cumanians: culturally Asian steppe nomadic immigrants with substantially more western Eurasian mitochondrial DNA lineages. Human Biology 77, 639–662.

Bogucki, P., 2000. How agriculture came to northcentral Europe. In: Price, T.D. (Ed.), Europe's First Farmers. Cambridge University Press, pp. 197–219.

Bogucki, P., Crabtree, P.J. (Eds.), 2004. Ancient Europe 8000 B.C.--A.D. 1000: Encyclopedia of the Barbarian World, Volume 1, 1. ed. Charles Scribner's Sons.

Bogucki, P., Grygiel, R., 1993. The first farmers of Central Europe: a survey article. Journal of Field Archaeology 20, 399–426.

Bollongino, R., Edwards, C.J., Alt, K.W., Burger, J., Bradley, D.G., 2006. Early history of European domestic cattle as revealed by ancient DNA. Biology Letters 2, 155–159.

Bollongino, R., Nehlich, O., Richards, M.P., Orschiedt, J., Thomas, M.G., Sell, C., Fajkosová, Z., Powell, A., Burger, J., 2013. 2000 years of parallel societies in Stone Age Central Europe. Science 342, 479–81.

Borić, D., 2009. Absolute dating of metallurgical innovations in the Vinča culture of the Balkans. In: Kienlin, T.L., Roberts, B.W. (Eds.), Metals and Societies. Studies in Honour of Barbara S. Ottaway. R. Habelt, Bonn, pp. 191–246.

Bramanti, B., 2008. Ancient DNA: genetic analysis of aDNA from sixteen skeletons of the Vedrovice. Anthropologie (Brno) 46, 153–160.

Bramanti, B., Thomas, M.G., Haak, W., Unterlaender, M., Jores, P., Tambets, K., Antanaitis-Jacobs, I., Haidle, M.N., Jankauskas, R., Kind, C.-J., Lueth, F., Terberger, T., Hiller, J., Matsumura, S., Forster, P., Burger, J., 2009. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. Science 326, 137–40.

Brandstätter, A., Sänger, T., Lutz-Bonengel, S., Parson, W., Béraud-Colomb, E., Wen, B., Kong, Q.-P., Bravi, C.M., Bandelt, H.-J., 2005. Phantom mutation hotspots in human mitochondrial DNA. Electrophoresis 26, 3414–29.

Brandt, G., Haak, W., Adler, C.J., Roth, C., Szécsényi-Nagy, A., Karimnia, S., Moller-Rieker, S., Meller, H., Ganslmeier, R., Friederich, S., Dresely, V., Nicklisch, N., Pickrell, J.K., Sirocko, F., Reich, D., Cooper, A., Alt, K.W., 2013. Ancient DNA reveals key stages in the formation of Central European mitochondrial genetic diversity. Science 342, 257–261.

Brandt, G., Szécsényi-Nagy, A., Roth, C., Alt, K.W., Haak, W., 2014. Human paleogenetics of Europe – The known knowns and the known unknowns. Journal of Human Evolution 10.1016/j.jhevol.2014.06.017.

Breasted, J.H., 1916. Ancient times, a history of the early world: an introduction to the study of ancient history and the career of early man. Boston.

Brotherton, P., Haak, W., Templeton, J., Brandt, G., Soubrier, J., Jane Adler, C., Richards, S.M., Sarkissian, C. Der, Ganslmeier, R., Friederich, S., Dresely, V., van Oven, M., Kenyon, R., Van der Hoek, M.B., Korlach, J., Luong, K., Ho, S.Y.W., Quintana-Murci, L., Behar, D.M., Meller, H., Alt, K.W., Cooper, A., 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nature Communications 4, 1764.

Budja, M., 1999. The transition to farming in Mediterranean Europe - an indigenous response. Documenta Praehistorica 26, 119–141.

Budja, M., 2005. The process of Neolithisation in South-eastern Europe: from ceramic female figurines and cereal grains to entoptics and human nuclear DNA polymorphic markers. Documenta Praehistorica 32, 53–72.

Busby, G.B.J., Brisighelli, F., Sánchez-Diz, P., Ramos-Luis, E., Martinez-Cadenas, C., Thomas, M.G., Bradley, D.G., Gusmão, L., Winney, B., Bodmer, W., Vennemann, M., Coia, V., Scarnicci, F., Tofanelli, S., Vona, G., Ploski, R., Vecchiotti, C., Zemunik, T., Rudan, I., Karachanak, S., Toncheva, D., Anagnostou, P., Ferri, G., Rapone, C., Hervig, T., Moen, T., Wilson, J.F., Capelli, C., 2012. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. Proceedings of the Royal Society B 279, 884–92.

Cann, R.L., Stoneking, M., Wilson, A.C., 1987. Mitochondrial DNA and human evolution. Nature 325, 31–6.

Caramelli, D., Lalueza-fox, C., Vernesi, C., Lari, M., Casoli, A., Mallegni, F., Chiarelli, B., Dupanloup, I., Bertranpetit, J., Barbujani, G., Bertorelle, G., 2003. Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. Proc Natl Acad Sci U S A 100, 6593–6597.

Caramelli, D., Milani, L., Vai, S., Modi, A., Pecchioli, E., Girardi, M., Pilli, E., Lari, M., Lippi, B., Ronchitelli, A., Mallegni, F., Casoli, A., Bertorelle, G., Barbujani, G., 2008. A 28,000 years old Cro-Magnon mtDNA sequence differs from all potentially contaminating modern sequences. PloS ONE 3, e2700.

Chan, D.C., 2006. Mitochondria: dynamic organelles in disease, aging, and development. Cell 125, 1241–52.

Chandler, H., Sykes, B., Zilhão, J., 2005. Using ancient DNA to examine genetic continuity at the Mesolithic-Neolithic transition in Portugal. In: Arias, P., Ontañón, R., García-Moncó, C. (Eds.), Actas Dell III Congreso Del Neolítico En La Península Ibérica, Santander, Monografías Del Instituto Internacional de Investigaciones Prehistóricas de Cantabria 1. pp. 781–786.

Chandler, H.C., 2003. Using Ancient DNA to Link Culture and Biology in Human Populations Using Ancient DNA to Link Culture and Biology in Human. Doctoral Thesis. University College.

Chapman, J., 1981. The Vinča culture of South-East Europe. Studies in Chronology, economy and society. BAR International Series 117. Oxford.

Chapman, J., 1993. Social power in the Iron Gates Mesolithic. In: Chapman, J., Dolukhanov, P. (Eds.), Cultural Transformation and Interactions in Eastern Europe. Aldershot: Avebury Press., pp. 71–122.

Chapman, J., 1994. The origins of farming in South East Europe. Prehistoire Europeenne 6, 133–156.

Chikhi, L., Destro-Bisol, G., Bertorelle, G., Pascali, V., Barbujani, G., 1998. Clines of nuclear DNA markers suggest a largely neolithic ancestry of the European gene pool. Proc Natl Acad Sci U S A 95, 9053–8.

Chikhi, L., Nichols, R.A., Barbujani, G., Beaumont, M.A., 2002. Y genetic data support the Neolithic demic diffusion model. Proc Natl Acad Sci U S A 99, 11008–13.

Childe, G., 1942. What Happened in History. Penguin Books, Harmondsworth.

Childe, G., 1957. Dawn of European Civilisation, 6. ed. New York.

Cinnioğlu, C., King, R., Kivisild, T., Kalfoğlu, E., Atasoy, S., Cavalleri, G.L., Lillie, A.S., Roseman, C.C., Lin, A. a, Prince, K., Oefner, P.J., Shen, P., Semino, O., Cavalli-Sforza, L.L., Underhill, P. a, 2004. Excavating Y-chromosome haplotype strata in Anatolia. Human Genetics 114, 127–48.

Clark, J.D.G., 1965. Radiocarbon dating and the expansion of farming culture from the Near East over Europe. Proceedings of the Prehistoric Society 31, 57–73.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W., McCabe, A.M., 2009. The Last Glacial Maximum. Science 325, 710–714.

College, S., Conolly, J., Shennan, S., 2004. Archaeobotanical evidence for the spread of farming in the Eastern Mediterranean. Current Anthropology 45, S35–S58.

Cruciani, F., La Fratta, R., Trombetta, B., Santolamazza, P., Sellitto, D., Colomb, E.B., Dugoujon, J.-M., Crivellaro, F., Benincasa, T., Pascone, R., Moral, P., Watson, E., Melegh, B., Barbujani, G., Fuselli, S., Vona, G., Zagradisnik, B., Assum, G., Brdicka, R., Kozlov, A.I., Efremov, G.D., Coppa, A., Novelletto, A., Scozzari, R., 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E-M78 and J-M12. Molecular Biology and Evolution 24, 1300–11.

Czene, A., 2008. Újkőkori falu Budakeszin. In: Gyöngyössy, M. (Ed.), Képek a Múltból - Az Elmúlt Évek Ásatásaiból Pest Megyében. Szentendre, pp. 14–15.

Czene, A., Ottományi, K., 2007. Budakeszi, Szőlőskert-Tangazdaság (MRT 7. k. 4/8 lh.). Régészeti Kutatások Magyarországon 2006, 162.

Dabney, J., Meyer, M., Pääbo, S., 2013. Ancient DNA damage. Cold Spring Harbor Perspectives in Biology 5, doi: 10.1101/cshperspect.a012567.

Dennel, R., 1992. The origin of crop agriculture in Europe. In: Cowan, C.W., Watson, P.J. (Eds.), The Origins of Agriculture. An International Perspective. Smithsonian Institution Press, Washington, pp. 71–100.

Der Sarkissian, C., Balanovsky, O., Brandt, G., Khartanovich, V., Buzhilova, A., Koshel, S., Zaporozhchenko, V., Gronenborn, D., Moiseyev, V., Kolpakov, E., Shumkin, V., Alt, K.W., Balanovska, E., Cooper, A., Haak, W., 2013. Ancient DNA Reveals Prehistoric Gene-Flow from Siberia in the Complex Human Population History of North East Europe. PLoS Genetics 9, e1003296.

Derenko, M., Malyarchuk, B., Grzybowski, T., Denisova, G., Dambueva, I., Perkova, M., Dorzhu, C., Luzina, F., Lee, H.K., Vanecek, T., Villems, R., Zakharov, I., 2007. Phylogeographic analysis of mitochondrial DNA in northern Asian populations. American Journal of Human Genetics 81, 1025–41.

Dimitrijević, S., 1974. Materijali Saveza arheoloških društava Jugoslavie [Das Problem der Gliederung der Starčevo-kultur mit Besonderer Rücksicht auf den Beitrag der Südpannonischen Fundstellen zur Lösung dieses Problems]. Poceci ranih zemljoradnickih kultura u Vojvodini i srpskom Podunavlju, Materijali Saveza arheoloških društava Jugoslavie X, 59–121.

Dimitrijević, S., 1979. Lasinjska kultura. In: Praistorija Jugoslavenskih Zemalja III. Eneolitsko Doba. Akademija nauka i umjetnosti Bosne i Hercegovine. Centar za balkanološka ispitivanja, Sarajevo, pp. 137–181.

Dombay, J., 1960. Die Siedlung und das Gräberfeld in Zengővárkony (Beiträge zur Kultur des Aeneolithikums in Ungarn). Arch. Hung. 37., Archaeologica Hungarae. Bonn.

Domboróczki, L., 2003. Radiocarbon data from Neolithic archaeological sites in Heves County (North-Eastern Hungary). Agria 39, 5–71.

Domboróczki, L., 2009. Settlement structures of the Alföld Linear Pottery Culture (ALPC) in Heves county (north-eastern Hungary): development models and historical reconstructions on micro, meso and macro levels. In: Kozlowski, J.K. (Ed.), Interaction Between Different Models of Neolithization North of the Central European Agro-Ecological Barrier. Polska Akademia Umiejetnosci, Prace Komisji Prehistorii Karpat, Krakow, pp. 75–127.

Domboróczki, L., Kaczanowska, M., Kozłowski, J., 2010. The Neolithic settlement at Tiszaszőlős-Domaháza-puszta and the question of the northern spread of the Körös Culture. Atti Della Soc Preistoria E Protostoria Della Reg Friuli-Venezia Giulia 17, 101–155.

Eichmann, W.J., Kertész, R., Marton, T., 2010. Mesolithic in the LBK Heartland of Transdanubia,Western Hungary. In: Gronenborn, D., Petrasch, J. (Eds.), Die

Neolithisierung Mitteleuropas. The Spread of Neolithic to Central Europe. International Conference, 24-26th June 2005. Mainz RGZM. RGZM, Mainz, pp. 211–233.

El-Sibai, M., Platt, D.E., Haber, M., Xue, Y., Youhanna, S.C., Wells, R.S., Izaabel, H., Sanyoura, M.F., Harmanani, H., Bonab, M.A., Behbehani, J., Hashwa, F., Tyler-Smith, C., Zalloua, P. a, 2009. Geographical structure of the Y-chromosomal genetic landscape of the Levant: a coastal-inland contrast. Annals of Human Genetics 73, 568–81.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10, 564–7.

F. Petres, É., 1959. Neolithic graves at Bicske. Az István király Múzeum Közleményei, Székesfehérvár 9.

Fernández, E., Pérez-Pérez, A., Gamba, C., Prats, E., Cuesta, P., Anfruns, J., Molist, M., Arroyo-Pardo, E., Turbón, D., 2014. Ancient DNA Analysis of 8000 B.C. Near Eastern Farmers Supports an Early Neolithic Pioneer Maritime Colonization of Mainland Europe through Cyprus and the Aegean Islands. PLoS Genetics 10, e1004401.

Forenbaher, S., Miracle, P.T., 2006. The spread of farming in the Eastern Adriatic. Documenta Praehistorica 33, 89–100.

Fu, Q., Mittnik, A., Johnson, P.L.F., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., Ronchitelli, A.M., Martini, F., Cremonesi, R.G., 2013. A revised timescale for human evolution based on ancient mitochondrial genomes. Current Biology 23, 1–7.

Galeta, P., Bruzek, J., 2009. Demographic model of the Neolithic transition in Central Europe. Documenta Praehistorica 36, 139–150.

Galeta, P., Sládek, V., Sosna, D., Bruzek, J., 2011. Modeling neolithic dispersal in central Europe: demographic implications. American Journal of Physical Anthropology 146, 104–15.

Gamba, C., Fernández, E., Tirado, M., Deguilloux, M.F., Pemonge, M.H., Utrilla, P., Edo, M., Molist, M., Rasteiro, R., Chikhi, L., Arroyo-Pardo, E., 2012. Ancient DNA from an Early

Neolithic Iberian population supports a pioneer colonization by first farmers. Molecular Ecology 21, 45–56.

Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T.F.G., Hofreiter, M., Bradley, D.G., Pinhasi, R., 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nature Communications 5, 5257.

Gamble, C., Davies, W., Pettitt, P., Hazelwood, L., Richards, M., 2005. The Archaeological and Genetic Foundations of the European Population during the Late Glacial: Implications for. Cambridge Archaeological Journal 15, 193.

Gilbert, M.T.P., Bandelt, H.-J., Hofreiter, M., Barnes, I., 2005. Assessing ancient DNA studies. Trends in Ecology & Evolution 20, 541–4.

Gilbert, M.T.P., Hansen, A.J., Willerslev, E., Rudbeck, L., Barnes, I., Lynnerup, N., Cooper, A., 2003. Characterization of genetic miscoding lesions caused by postmortem damage. American Journal of Human Genetics 72, 48–61.

Gilbert, M.T.P., Tomsho, L.P., Rendulic, S., Packard, M., Drautz, D.I., Sher, A., Tikhonov, A., Dalén, L., Kuznetsova, T., Kosintsev, P., Campos, P.F., Higham, T., Collins, M.J., Wilson, A.S., Shidlovskiy, F., Buigues, B., Ericson, P.G.P., Germonpré, M., Götherström, A., Iacumin, P., Nikolaev, V., Nowak-Kemp, M., Willerslev, E., Knight, J.R., Irzyk, G.P., Perbost, C.S., Fredrikson, K.M., Harkins, T.T., Sheridan, S., Miller, W., Schuster, S.C., 2007. Whole-Genome Shotgun Sequencing of Mitochondria from Ancient Hair Shafts. Science 317, 1927–1930.

Giles, R.E., Blanc, H., Cann, H.M., Wallace, D.C., 1980. Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A 77, 6715–9.

Gkiasta, M., Russell, T., Shennan, S., Steele, J., 2003. Neolithic transition in Europe : the radiocarbon record revisited. Antiquity 77, 45–62.

Gläser, R., 1996. Zur absoluten Datierung der Vinča-Kultur anhand von 14C Daten. In: The Vinča Culture, Its Role and Cultural Connections: International Symposium on the Vinca

Culture, Its Role and Cultural Connections, Timișoara, Romania, October 1995. Timisoara, pp. 175–212.

Gokcumen, O., Dulik, M.C., Pai, A.A., Zhadanov, S.I., Rubinstein, S., Osipova, L.P., Andreenkov, O. V, Tabikhanova, L.E., Gubina, M.A., Labuda, D., Schurr, T.G., 2008. Genetic variation in the enigmatic Altaian Kazakhs of South-Central Russia: insights into Turkic population history. American Journal of Physical Anthropology 136, 278–93.

Gronenborn, D., 1999. A variation on a basic theme : The transition to farming in southern Central Europe. Journal of World Prehistory 13, 123–210.

Gronenborn, D., 2007. Beyond the models: “Neolithisation” in Central Europe. In: Whittle, A., Cummings, V. (Eds.), Going over: The Mesolithic-Neolithic Transition in North-West Europe. Oxford University Press/British Academy, Oxford, pp. 73–98.

Guba, Z., Hadadi, É., Major, Á., Furka, T., Juhász, E., Koós, J., Nagy, K., Zeke, T., 2011. HVS-I polymorphism screening of ancient human mitochondrial DNA provides evidence for N9a discontinuity and East Asian haplogroups in the Neolithic Hungary. Journal of human genetics 56, 784–96.

H. Simon, K., 1994. Frühneolithische Kultgegenstände bei Gellénháza (Kom. Zala). In: A Kőkortól a Középkorig. Tanulmányok Trogmayer Ottó 60. Születésnapjára [Von Steinzeit Bis Zum Mittelalter. Studien Zum 60.Geburtstag von Ottó Trogmayer]. Lőrinczy, G, Szeged.

Haak, W., Balanovsky, O., Sanchez, J.J., Koshel, S., Zaporozhchenko, V., Adler, C.J., Der Sarkissian, C., Brandt, G., Schwarz, C., Nicklisch, N., Dresely, V., Fritsch, B., Balanovska, E., Villems, R., Meller, H., Alt, K.W., Cooper, A., 2010. Ancient DNA from European early Neolithic farmers reveals their Near Eastern affinities. PLoS Biology 8, e1000536.

Haak, W., Brandt, G., de Jong, H.N., Meyer, C., Ganslmeier, R., Heyd, V., Hawkesworth, C., Pike, A.W.G., Meller, H., Alt, K.W., 2008. Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. Proc Natl Acad Sci U S A 105, 18226–31.

Haak, W., Forster, P., Bramanti, B., Matsumura, S., Brandt, G., Tänzer, M., Villems, R., Renfrew, C., Gronenborn, D., Alt, K.W., Burger, J., 2005. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. Science 310, 1016–8.

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Fu, Q., Mittnik, A., Bánffy, E., Szécsényi-Nagy, A., Roth, C., Economou, C., Francken, M., Friederich, S., Pena, R.G., Hallgren, F., Khartanovich, V., Khokhlov, A., Kunst, M., Kuznetsov, P., Meller, H., Mochalov, O., Moiseyev, V., Nicklisch, N., Risch, R., Guerra, M.A.R., Wahl, J., Meyer, M., Krause, J., Brown, D., Anthony, D., Cooper, A., Alt, K.W., Reich, D., 2015. Massive migration from the steppe is a source for Indo-European languages in Europe. In preparation.

Hammer, M.F., 1995. A recent common ancestry for human Y chromosomes. Nature 378, 376–8.

Hammer, M.F., Zegura, S.L., 2002. The Human Y Chromosome Haplogroup Tree: Nomenclature and Phylogeography of Its Major Divisions. Annual Review of Anthropology 31, 303–321.

Hartz, S., Lübke, H., Terberger, T., 2007. From fish and seal to sheep and cattle: new research into the process of neolithisation in northern Germany. In: Academy, B. (Ed.), Going over: The Mesolithic-Neolithic Transition in North-West Europe. pp. 564–594.

Hebsgaard, M.B., Phillips, M.J., Willerslev, E., 2005. Geologically ancient DNA: fact or artefact? Trends in microbiology 13, 212–20.

Hertelendi, E., Kalicz, N., Raczky, P., Horvath, F., Veres, M., Svingor, Futo, I Bartosiewicz, L., 1995. Re-evaluation of the Neolithic in Eastern Hungary based oncalibrated radiocarbon dates. Radiocarbon 37, 239–244.

Hervella, M., Izagirre, N., Alonso, S., Fregel, R., Alonso, A., Cabrera, V.M., de la Rúa, C., 2012. Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the Neolithic expansion into Europe. PloS ONE 7, e34417.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C., 1984. DNA sequences from the quagga, an extinct member of the horse family. Nature 312, 282–4.

Hodder, I., 1990. The Domestication of Europe., 1. ed. Basil Blackwell, Oxford.

Hofreiter, M., Jaenicke, V., Serre, D., Haeseler, A. Von, Pääbo, S., 2001a. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient 29, 4793–4799.

Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., Pääbo, S., 2001b. Ancient DNA. Nature Reviews Genetics 2, 353–359.

Horváth, F., 2005. Grozsa. Előzetes eredmények az újkőkori tell 1978 ás 1996 közötti feltárásából. In: Hétköznapok Vénuszai. Hódmezővásárhely, pp. 51–83.

Horváth, F., 2006. Comments on the connections between the Vinča complex and the Carpathian Basin. In: Tasic, N., Grozdanov, C. (Eds.), Homage to Milutin Garašin. Serbian Academy of Sciences and Arts, Macedonian Academy of Sciences and Arts, Belgrade, pp. 309–324.

Horváth, L.A., H. Simon, K., 2003. Das Neolithikum und die Kupferzeit in Südwesttransdanubien., Inventaria. ed. Magyar Nemzeti Múzeum, Budapest.

Ingman, M., Kaessmann, H., Pääbo, S., Gyllensten, U., 2000. Mitochondrial genome variation and the origin of modern humans. Nature 408, 708–13.

Irwin, J., Egyed, B., Saunier, J., Szamosi, G., O'Callaghan, J., Padar, Z., Parsons, T.J., 2007. Hungarian mtDNA population databases from Budapest and the Baranya county Roma. International journal of legal medicine 121, 377–83.

ISOGG, 2014. International Society of Genetic Genealogy (2014). Y-DNA Haplogroup Tree 2014, Version: 9.54, Date: 30. April 2014, http://www.isogg.org/tree.

Jakucs, J., Voicsek, V., 2014. The northermost distribution of the early Vinča Culture in the Danube valley: A case study from Szederkény-Kukorica-dűlő (Baranya County, south Hungary). Antaeus Communicationes ex Instituto Archaeologico Academiae Scientarum Hungaricae In press.

Janosic, I., 1984. Arheološka istraživanja na području općine Vinkovci. Arheološka istraživanja u istočnoj Slavoniji i Baranji. Izdanja Hrvatskog arheoloskog drustva 9, 143–151.

Jelinek, J., 1973. Die neolitische und bronzezeitliche Besiedlung der heutigen Tschechoslowakei. In: Schwidetzky, I. (Ed.), Die Anfänge Des Neolithikums Vom Orient Bis Nordeuropa. Anthropologie. Teil VIIIab. Köln-Wien, pp. 186–199.

Jeran, N., Havas Augustin, D., Grahovac, B., Kapović, M., Metspalu, E., Villems, R., Rudan, P., 2009. Mitochondrial DNA heritage of Cres Islanders--example of Croatian genetic outliers. Collegium antropologicum 33, 1323–8.

Jobling, M., Hollox, E., Kivisild, T., Tyler-Smith, C., Hurles, M., 2013. Human Evolutionary Genetics: Origins, Peoples and Disease, 2. ed. Garland Science, New York, London.

Jovanovic, B., 1969. Chronological frames of the Iron Gates Group of the Early Neolithic period. Archaeologica Iugoslavica 10, 23–38.

Juhász, I., 2004. Palynological evidence of Preneolithization in South-Western Transdanubia. Antaeus Communicationes ex Instituto Archaeologico Academiae Scientarum Hungaricae 27, 213–225.

K. Zoffmann, Z., 2012. Anthropological analysis of the burials from the LBK settlement at Balatonszárszó-Kis-erdei-dűlő. Acta Archaeologica Academiae Scientarium Hungaricae 63, 301–316.

Kalicz, N., 1969. A rézkori Balatoni-csoport Veszprém megyében. Veszprém Megyei Múzeumok Közleményei 8.

Kalicz, N., 1973. Über die chronologische Stellung der Balaton-Gruppe. In: Symposium Über Die Entstehung Und Chronologie Der Badener-Kultur. Bratislava. Verlag der Slowakischen Akademie der Wissenschaften, Bratislava.

Kalicz, N., 1983. A Balaton—Lasinja kultúra történeti kérdései és fémleletei. Archaeológiai Értesítő 109.

Kalicz, N., 1988. Beiträge zur Entsterungsphase der LengyelKultur. Slovenská archeológia 36, 105–118.

Kalicz, N., 1990. Frühneolithische Siedlungsfunde aus Südwestungarn. Quellenanalyse zur Geschichte der Starčevo-Kultur, Inventaria Praehistorica Hungariae. Inventaria Praehistorica Hungariae, Budapest.

Kalicz, N., 1994a. A Dunántúli (Közép-európai) vonaldíszes kerámia legidősebb leletei és a korai Vinca kultúra [Die Ältesten Funde der Transdanubischen (Mitteleuropäischen) Linienbandkeramik und die frühe Vinča-Kultur]. In: A Kőkortól a Középkorig. Tanulmányok Trogmayer Ottó 60. Születésnapjára [Von Steinzeit Bis Zum Mittelalter. Studien Zum 60.Geburtstag von Ottó Trogmayer]. Lőrinczy, G, Szeged, pp. 67–84.

Kalicz, N., 1994b. Wenden des Spätneolithikums im Oberen Theißgebiet. Jósa András Múzeum Évkönyve 36, 263–290.

Kalicz, N., 1995. Die Balaton-Lasinja-Kultur in der Kupferzeit Südost- und Mitteleuropas, Neuere Daten zur Siedlungsgeschichte und Chronologie der Kupferzeit des Karpatenbeckens., Inventaria. ed. Magyar Nemzeti Múzeum, Budapest.

Kalicz, N., 2010. An Grenze „zweier Welten" -Transdanubien (Ungarn) im Frühneolithikum. In: Gronenborn, D., Petrasch, J. (Eds.), Die Neolithisierung Mitteleuropas. The Spread of Neolithic to Central Europe. International Conference, 24-26th June 2005. Mainz RGZM. RGZM, Mainz, pp. 235–254.

Kalicz, N., Kreiter, A., Kreiter, E., Tokai, Z.A., Tóth, M., Bajnóczi, B., 2012. A neolitikum történeti és kronológiai kérdései Becsehely-Bükkaljai-dűlő lelőhelyen. In: Barbara, K. (Ed.), ΜΩΜΟΣ IV. Őskoros Kutatók IV. Összejövetelének Konferenciakötete. Debrecen, 2005. Március 22-24. Déri Múzeum Régészeti Tár, Debrecen, pp. 87–170.

Kalicz, N., M. Virág, Z., T. Bíró, K., 1998. The northern periphery of the Early Neolithic Starcevo Culture in southwestern Hungary—A case study of an excavation at the Lake Balaton. Porocilo o raziskovanju paleolilika, neolitika in eneolitika v Sloveniji 25, 151–187.

Kalicz, N., Makkay, J., 1972. A neolitikus Sopot-Bicske kultúra. [Die neolitische Sopot-Bicske-Kultur]. Archaeológiai Értesítő 99, 3–14.

Kalicz, N., Makkay, J., 1977. Die Linienbandkeramik in der Großen Ungarischen Tiefebene. Studia Archaeologica 7. Budapest.

Kalicz, N., Molnár, S., Rózsás, M., 2007. Az élelemtermelés kezdetei Somogy megyében a Kr.e. 7–6. évezred fordulóján. Az újkőkor (neolitikum) legidősebb szakasza. Beginnings of food production in Somogy county at the turn of the 7th–6th millenia B.C. The earliest phase of the Neolithic Period. Communicationes Archaeologicae Hungariae 19–64.

Karafet, T.M., Lansing, J.S., Redd, A.J., Reznikova, S., Watkins, J.C., Surata, S.P.K., Arthawiguna, W.A., Mayer, L., Bamshad, M., Jorde, L.B., Hammer, M.F., 2005. Balinese Y-chromosome perspective on the peopling of Indonesia: genetic contributions from pre-neolithic hunter-gatherers, Austronesian farmers, and Indian traders. Human Biology 77, 93–114.

Karafet, T.M., Mendez, F.L., Meilerman, M.B., Underhill, P. a, Zegura, S.L., Hammer, M.F., 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome research 18, 830–8.

Karlsson, A.O., Wallerström, T., Götherström, A., Holmlund, G., 2006. Y-chromosome diversity in Sweden – A long-time perspective. European Journal of Human Genetics 14, 963–970.

Keerl, V., 2014. A River Runs Through It- Ancient DNA Data on the Neolithic Populations of the Great Hungarian Plain. Doctoral thesis. Johannes Gutenberg University.

Keller, A., Graefen, A., Ball, M., Matzas, M., Boisguerin, V., Maixner, F., Leidinger, P., Backes, C., Khairat, R., Forster, M., Stade, B., Franke, A., Mayer, J., Spangler, J., McLaughlin, S., Shah, M., Lee, C., Harkins, T.T., Sartori, A., Moreno-Estrada, A., Henn, B., Sikora, M., Semino, O., Chiaroni, J., Rootsi, S., Myres, N.M., Cabrera, V.M., Underhill, P. a, Bustamante, C.D., Vigl, E.E., Samadelli, M., Cipollini, G., Haas, J., Katus, H., O'Connor, B.D., Carlson, M.R.J., Meder, B., Blin, N., Meese, E., Pusch, C.M., Zink, A., 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nature Communications 3, 698.

Keyser, C., Bouakaze, C., Crubézy, E., Nikolaev, V.G., Montagnon, D., Reis, T., Ludes, B., 2009. Ancient DNA provides new insights into the history of south Siberian Kurgan people. Human Genetics 126, 395–410.

King, R.J., Ozcan, S.S., Carter, T., Kalfoğlu, E., Atasoy, S., Triantaphyllidis, C., Kouvatsi, A., Lin, a a, Chow, C.-E.T., Zhivotovsky, L. a, Michalodimitrakis, M., Underhill, P. a, 2008. Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. Annals of Human Genetics 72, 205–14.

Kiss, V., Sebők, K., 2007. Balatonszemes-Bagódomb. In: Belényesy, K., Honti, S., Kiss, V. (Eds.), Gördülő Idő. Régészeti Feltárások Az M7-Es Autópálya Somogy Megyei Szakaszán Zamárdi És Ordacsehi Között. Somogyi Múzeumok Igazgatósága, pp. 113–121.

Kivisild, T., Reidla, M., Metspalu, E., Rosa, A., Brehm, A., Pennarun, E., Parik, J., Geberhiwot, T., Usanga, E., Villems, R., 2004. Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. American Journal of Human Genetics 75, 752–70.

Kivisild, T., Rootsi, S., Metspalu, M., Mastana, S., Kaldma, K., Parik, J., Metspalu, E., Adojaan, M., Tolk, H.-V., Stepanov, V., Gölge, M., Usanga, E., Papiha, S.S., Cinnioğlu, C., King, R., Cavalli-Sforza, L., Underhill, P. a, Villems, R., 2003. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. American Journal of Human Genetics 72, 313–32.

Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G., Kronenberg, F., 2011. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Human Mutation 32, 25–32.

Köhler, K., 2004. Anthropological finds of the Lengyel culture from Csabdi-Télizöldes. Alba Regia XXXIII., 7–24.

Köhler, K., 2013. Biological reconstruction of the Late Neolithic Lengyel Culture. Dissertationes Archaeologicae Ser 3 No 1 179–204.

Köhler, K., 2014. A Budakeszi, Szőlőskert -Tangazdaság lelőhelyen feltárt, a közép-európai vonaldíszes kultúra körébe tartozó temetkezések embertani vizsgálatának eredményei. Folia Anthropologica 13, 47–61.

Köhler, K., Pálfi, G., Molnár, E., Zalai-Gaál, I., Osztás, A., Bánffy, E., Kirinó, K., Kiss, K.K., 2012. A Late Neolithic Case of Pott's Disease from Hungary. International Journal of Osteoarchaeology 7.

Komso, D., 2006. The Mesolithic in Croatia., Opuscula a. ed.

Kotsakis, K., 2001. Mesolithic to Neolithic in Greece. Continuity, discontinuity or change of course. Documenta Praehistorica XXVIII, 63–73.

Krause, J., Briggs, A.W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., Pääbo, S., 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. Current Bology 20, 231–6.

Krause, J., Dear, P.H., Pollack, J.L., Slatkin, M., Spriggs, H., Barnes, I., Hofreiter, M., Lister, A.M., Ebersberger, I., Pa, S., 2006. Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae 439, 10–13.

Kristiansen, K., Larsson, T.B., 2005. The Rise of Bronze Age Society: Travels, Transmissions and Transformations. Cambridge University Press.

Kustár, R., 2003. 182. Harta-Gátőrház. Régészeti Kutatások Magyarországon 2003, 227.

Lacan, M., 2011. La Néolithisation du bassin méditerranéen. These. L'Université de Toulouse.

Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Duranthon, F., Guilaine, J., Crubézy, E., Ludes, B., 2011a. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. Proc Natl Acad Sci U S A 108, 9788–91.

Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Tarrús, J., Bosch, A., Guilaine, J., Crubézy, E., Ludes, B., 2011b. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. Proc Natl Acad Sci U S A 108, 18255–9.

Lalueza-Fox, C., 1996. Physical anthropological aspects of the Mesolithic-Neolithic transition in the Iberian Peninsula. Current Anthropology 37, 689–95.

Lalueza-Fox, C., Sampietro, M.L., Gilbert, M.T.P., Castri, L., Facchini, F., Pettener, D., Bertranpetit, J., 2004. Unravelling migrations in the steppe: mitochondrial DNA sequences from ancient central Asians. Proceedings of the Royal Society B 271, 941–7.

Lappalainen, T., Laitinen, V., Salmela, E., Andersen, P., Huoponen, K., Savontaus, M., Lahermo, P., 2008. Migration waves to the Baltic Sea region. Annals of Human Genetics 72, 337–348.

Larmuseau, M.H.D., Vanoverbeke, J., Van Geystelen, A., Defraene, G., Vanderheyden, N., Matthys, K., Wenseleers, T., Decorte, R., 2013. Low historical rates of cuckoldry in a Western European human population traced by Y-chromosome and genealogical data. Proceedings of the Royal Society B 280, 20132400.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Gallego Romero, I., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C.A., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513, 409–413.

Lee, E.J., Krause-Kyora, B., Rinne, C., Schütt, R., Harder, M., Müller, J., von Wurmb-Schwarkl, N., Nebel, A., 2013. Ancient DNA insights from the Middle Neolithic in Germany. Archaeological and Anthropological Sciences 6, 199–204.

Lee, E.J., Makarewicz, C., Renneberg, R., Harder, M., Krause-Kyora, B., Müller, S., Ostritz, S., Fehren-Schmitz, L., Schreiber, S., Müller, J., von Wurmb-Schwark, N., Nebel, A., 2012. Emerging genetic patterns of the European Neolithic: perspectives from a late Neolithic Bell Beaker burial site in Germany. American Journal of Physical Anthropology 148, 571–9.

Lenneis, E., Stadler, P., 1995. Zur Absolutchronologie der Linearbandkeramik aufgrund von 14C-Daten. Archaologie Osterreichs 6, 4–13.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics (Oxford, England) 25, 1451–2.

Lichter, C., 2005. Western Anatolia in the Late Neolithic and Early Chalcolithic : the actual state of research. In: Lichter, C. (Ed.), How Did Farming Reach Europe? Anatolian-European Relations from the Second Half of the 7th Through the First Half of the 6th Millennium Cal BC. BYZAS 2. Istanbul, pp. 59–74.

Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature 362, 709–15.

Makkay, J., 1975. A bicskei neolitikus telep és temető. Az István király Múzeum Közleményei D sorozat Székesfehérvár.

Makkay, J., 1982. A magyarországi neolitikum kutatásának új eredmeényei. az időrend és a népi azonosítás kérdései. [Neue Ergebnisse der Forschung des ungarländischen Neolithikums. Die Fragen der Datierung und der ethnischen Identifizierung]. Budapest.

Makkay, J., Starnini, E., Tulok, M., 1996. Excavations at Bicske-Galagonyás (part III). The Notenkopf and Sopot-Bicske cultural phases. Societá per la Preistoria e protostoria della regione Friuli – Venezia Giulia Quaderno 6.

Malmström, H., Gilbert, M.T.P., Thomas, M.G., Brandström, M., Storå, J., Molnar, P., Andersen, P.K., Bendixen, C., Holmlund, G., Götherström, A., Willerslev, E., 2009. Ancient DNA

reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. Current Biology 19, 1758–62.

Malmström, H., Svensson, E.M., Gilbert, M.T.P., Willerslev, E., Götherström, A., Holmlund, G., 2007. More on Contamination : The Use of Asymmetric Molecular Behavior to Identify Authentic Ancient Human DNA. Molecular Biology and Evolution 24, 998–1004.

Malyarchuk, B., Grzybowski, T., Derenko, M., Perkova, M., Vanecek, T., Lazur, J., Gomolcak, P., Tsybovsky, I., 2008. Mitochondrial DNA phylogeny in Eastern and Western Slavs. Molecular Biology and Evolution 25, 1651–8.

Malyarchuk, B.A., Rogozin, I.B., Berikov, V.B., Derenko, M. V, 2002. Analysis of phylogenetically reconstructed mutational spectra in human mitochondrial DNA control region. Human Genetics 111, 46–53.

Mannino, M.A.-, Catalano, G., Talamo, S., Mannino, G., Salvo, R. Di, Schimmenti, V., Lalueza-Fox, C., Messina, A., Petruso, D., Caramelli, D., Richards, M.P., Sineo, L., 2012. Origin and Diet of the Prehistoric Hunter-Gatherers on the Mediterranean Island of Favignana (Egadi Islands, Sicily). PloS ONE 7(11): e49802.

Marlowe, F.W., 2004. Marital residence among foragers. Current Anthropology 45, 277–284.

Marton, T., Oross, K., 2010. Siedlungsforschung in linienbandkeramischen Fundorten in Zentral- und Südtransdanubien- Wiege, Peripherie oder beides? In: Kreienbrink, F., Cladders, M., Stäuble, H., Tischendorf, T., Wolfram, S. (Eds.), Siedlungsstruktur Und Kulturwandel in Der Bandkeramik. Beiträge Der Internationalen Tagung “Neue Fragen Zur Bandkeramik Oder Alles Beim Alten?!” Leipzig, 23. Bis 24. September 2010. Arbeits- Und Forschungsberichte Zur Sächsischen Bodendenkmalpflege, Beihe. Druckhaus Dresden GmbH, Dresden, pp. 220–239.

Meier-Arendt, W., 2007. Zum Begriff der Kultur im Neolitikum Europas. SCIVA tomul 58, 401–407.

Menozzi, P., Piazza, A., Cavalli-Sforza, L., 1978. Synthetic maps of human gene frequencies in Europeans. Science 201, 786–92.

Meyer, M., Fu, Q., Aximu-Petri, A., Glocke, I., Nickel, B., Arsuaga, J.-L., Martínez, I., Gracia, A., de Castro, J.M.B., Carbonell, E., Pääbo, S., 2014. A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature 505, 403–6.

Meyer-Arendt, W., 2003. Zum Kultur- und Ethnosbegriff in der prähistorischen Archäologie. In: Jerem, E., Raczky, P. (Eds.), Morgenrot Der Kulturen: Frühe Etappen Der Menschheitsgeschichte in Mittel- Und Südosteuropa. Festschrift Für Nándor Kalicz Zum 75. Geburtstag. Archaeolingua, Budapest, pp. 23–28.

Mikič, Ž., 1989. Anthropological remains from the Neolithic sites in Serbia. In: Srejović, D. (Ed.), The Neolithic of Serbia. Archaeological Research 1948-1988. Belgrade, pp. 20–23.

Milisauskas, S., Harding, A.F., Kruk, J., Jochim, M., Wells, P.S., 2011. European Prehistory: A Survey. Springer Science & Business Media.

Minichreiter, K., 1992. Starčevačka kultura u sjevernoj Hrvatskoj. Arheološki zavod Filozofskog fakultea Sveučilišta u Zagrebu. Sveučilišta u Zagrebu.

Minichreiter, K., 2001. The architecture of Early and Middle Neolithic settlements of the Starcevo culture in Northern Croatia. Documenta Praehistorica XXVIII, 199–214.

Minichreiter, K., 2002. Ukopi stanovnika u naseljima starevake kulture u Hrvatskoj (Burials in the Settlements of Star$evo Culture in Croatia). Histria Antiqua 8, 63–72.

Minichreiter, K., Bronic, I.K., 2006. New radiocarbon dates for the Early Starčevo Culture in Croatia. Pril Inst arh Zagreb 23, 5–16.

Minichreiter, K., Markovic, Z., 2011. Architecture of Lasinja culture settlements in the light of new investigations in northern Croatia. Documenta Praehistorica XXXVIII, 333–343.

Mlekuz, D., Budja, M., Payton, R., Bonsall, C., 2008. “Mind the Gap”: Caves, Radiocarbon Sequences, and the Mesolithic–Neolithic Transition in Europe—Lessons from the Mala Triglavca Rockshelter Site. Geoarchaeology: An International Journal 23, 398–416.

Mullis, K.B., Faloona, F.A., 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods in Enzymology 155, 335–50.

Muscheler, R., Kromer, B., Björck, S., Svensson, A., Friedrich, M., Kaiser, K.F., Southon, J., 2008. Tree rings and ice cores reveal 14C calibration uncertainties during the Younger Dryas. Nature Geoscience 1, 263–267.

Nagy, E., 2007. Az M6 – M60 gyorsforgalmi út Baranya megyei szakasza. Magyar múzeumok 2007/1 21–24.

Nasidze, I., Quinque, D., Dupanloup, I., Rychkov, S., Naumova, O., Zhukova, O., Stoneking, M., 2004. Genetic evidence concerning the origins of South and North Ossetians. Annals of Human Geneticsuman genetics 68, 588–99.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia Univ. Press, New York.

Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A 76, 5269–5273.

Nikitin, A.G., Newton, J.R., Potekhina, I.D., 2012. Mitochondrial haplogroup C in ancient mitochondrial DNA from Ukraine extends the presence of East Eurasian genetic lineages in Neolithic Central and Eastern Europe. Journal of human genetics 57, 610–2.

Nikolova, L., 1998. Neolithic Sequence: The Upper Stryama Valley in Western Thrace (with Appendix: Radiocarbon Dating of the Balkan Neolithic). Documenta Praehistorica XXV, 99–131.

Obelić, B., Krznarić Škrivanko, M., Marijan, B., Boronić, I.K., 2004. Radiocarbon dating of Sopot culture sites (Late Neolithic) in eastern Croatia. Radiocarbon 46, 245–258.

Olalde, I., Allentoft, M.E., Sánchez-Quinto, F., Santpere, G., Chiang, C.W.K., DeGiorgio, M., Prado-Martinez, J., Rodríguez, J.A., Rasmussen, S., Quilez, J., Ramírez, O., Marigorta, U.M., Fernández-Callejo, M., Prada, M.E., Encinas, J.M.V., Nielsen, R., Netea, M.G., Novembre, J., Sturm, R. a, Sabeti, P., Marquès-Bonet, T., Navarro, A., Willerslev, E., Lalueza-Fox, C., 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature 507, 225–8.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T.,

Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M. V, Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A.S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J. V, Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J., Willerslev, E., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature 499, 74–8.

Oross, K., 2013. Balatonszárszó-Kis-Erdei-dűlő lelőhely középső neolit településszerkezete és közép-európai párhuzamai. Doktori disszertáció. ELTE University.

Oross, K., Bánffy, E., 2009. Three successive waves of Neolithisation: LBK development in Transdanubia. Documenta Praehistorica 36, 175–189.

Oross, K., Marton, T., 2012. Neolithic burials of the Linearbandkeramik settlement at Balatonszárszó and their European context. Acta Archaeologica Academiae Scientarium Hungaricae 63, 257–299.

Osztás, A., Zalai-Gaál, I., Bánffy, E., 2012. Alsónyék-Bátaszék: a new chapter in the research of Lengyel culture. Documenta Praehistorica 39, 377–396.

Özdoğan, M., 2008. An Alternative Approach in Tracing Changes in Demographic Composition. In: Bocquet-Appel, J.-P., Bar-Yosef, O. (Eds.), The Neolithic Demographic Transition and Its Consequences. Springer, pp. 139–178.

Özdoğan, M., 2011. Archaeological Evidence on the Westward Expansion of Farming Communities from Eastern Anatolia to the Aegean and the Balkans. Current Anthropology 52, S415–S430.

Pääbo, S., 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A 86, 1939–43.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., 2004. Genetic analyses from ancient DNA. Annuall Review of Genetics 38, 645–79.

Palanichamy, M.G., Zhang, C.-L., Mitra, B., Malyarchuk, B., Derenko, M., Chaudhuri, T.K., Zhang, Y.-P., 2010. Mitochondrial haplogroup N1a phylogeography, with implication to the origin of European farmers. BMC evolutionary biology 10, 304.

Pavúk, J., 1972. Neolithisches Gräberfeld in Nitra. Slovenská archeológia 20, 5–105.

Penrose, L.S., 1954. Distance, size and shape. Annals of Eugenics 18, 137–343.

Pericić, M., Lauc, L.B., Klarić, I.M., Rootsi, S., Janićijevic, B., Rudan, I., Terzić, R., Colak, I., Kvesić, A., Popović, D., Sijacki, A., Behluli, I., Dordevic, D., Efremovska, L., Bajec, D.D., Stefanović, B.D., Villems, R., Rudan, P., 2005. High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. Molecular Biology and Evolution 22, 1964–75.

Perlès, C., 2001. An alternate (and old-fashioned) view of Neolithisation in Greece. Documenta Praehistorica 30, 99–113.

Perlès, C., 2005. From the Near East to Greece: let's reverse the focus – cultural elements that did not transfer. In: Lichter, C. (Ed.), BYZAS 2. How Did Farming Reach Europe? Anatolian-European Relations from the Second Half of the 7th through the First Half of the 6th Millennium Cal BC. Veröffentlichungen Des Proceedings of the International Workshop Istanbul, 20-22 May 2004. Veröffentlichungen des Deutschen Archäologischen Instituts Istanbul. Ege Yayinlari, Istanbul, pp. 275–290.

Petrasch, J., 2001. Seid fruchtbar und mehret euch und füllet die Erde und machet sie euch untertan. Archäologisches Korrespondenzblatt 31, 13–25.

Pickrell, J.K., Reich, D., 2014. Toward a new history and geography of human genes informed by ancient DNA. Trends in Genetics 30, 377–389.

Pinhasi, R., Fort, J., Ammerman, A.J., 2005. Tracing the origin and spread of agriculture in Europe. PLoS Biology 3, e410.

Pinhasi, R., von Cramon-Taubadel, N., 2009. Craniometric data supports demic diffusion model for the spread of agriculture into Europe. PloS ONE 4, e6747.

Pinhasi, R., von Cramon-Taubadel, N., 2012. A craniometric perspective on the transition to agriculture in Europe. Human Biology 84, 45–66.

Poinar, H., Kuch, M., Mcdonald, G., Martin, P., Pääbo, S., 2003. Nuclear Gene Sequences from a Late Pleistocene Sloth Coprolite. Current Biology 13, 1150–1152.

Poznik, G.D., Henn, B.M., Yee, M.-C., Sliwerska, E., Euskirchen, G.M., Lin, A.A., Snyder, M., Quintana-Murci, L., Kidd, J.M., Underhill, P.A., Bustamante, C.D., 2013. Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. Science 341, 562–5.

Price, T.D. (Ed.), 2000. Europe's first farmers. Cambridge University Press, Cambridge.

Raczky, P., 1974. A Lengyeli-kultúra legkésőbbi szakaszának leletei a Dunántúlon. Archeológiai Értesítő 101.

Radic, D., 2005. Vela spila : preliminarna analiza starijeneolitičkih i mezolitičkih naslaga iz sonde istražene 2004. godine. Opuscula Archaeologica Papers of the Department of Archaeology 29, 323–348.

Rasteiro, R., Chikhi, L., 2013. Female and male perspectives on the Neolithic transition in Europe: clues from ancient and modern genetic data. PloS ONE 8, e60844.

Regenye, J., 1996a. Die Sopot-Kultur in Transdanubien und ihre südlichen Beziehungen. In: The Vinča Culture, Its Role and Cultural Connections: International Symposium on the Vinca Culture, Its Role and Cultural Connections, Timișoara, Romania, October 1995. Mirton, Timisoara, pp. 163–175.

Regenye, J., 1996b. a Sopot kultúra lelőhelyei a Balaton-felvidéken. Communicationes Archaeologicae Hungariae 23–41.

Regenye, J., 2002. Chronological situation of the Sopot culture in Hungary. A Sopot kultúra kronológiai helyzete Magyarországon. Veszprém Megyei Múzeumok Közleményei 22, 31–42.

Regenye, J., 2006. Temetkezések Veszprém, Jutasi út lelőhelyen. Veszprém Megyei Múzeumok Közleményei 24, 7–36.

Regenye, J., 2010. What about the other side: Starčevo and LBK settlements north of Lake Balaton. In: Kozlowski, J.K., Raczky, P. (Eds.), Neolithization of the Carpathian Basin: Northernmost Distribution of the Starčevo/Körös Culture. Krakow, Budapest, pp. 53–64.

Regenye, J., 2011. Kő és Agyag. Település és életmód a neolitikum-rézkor fordulóján a Dunántúlon. Veszprém Megyei Múzeumok Igazgatósága, Veszprém.

Renfrew, C., 1987. Archaeology and Language. The Puzzle of Indo-European Origins. Cambridge University Press, Cambridge.

Richards, M., Côrte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A., Papiha, S., Hedges, R., Bandelt, H.J., Sykes, B., 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. American Journal of Human Genetics 59, 185–203.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Gölge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Calì, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di Rienzo, A., Novelletto, A., Oppenheim, A., Nørby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozari, R., Torroni, A., Bandelt, H.J., 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. American Journal of Human Genetics 67, 1251–76.

Richards, M., Macaulay, V., Sykes, B., Pettitt, P., Hedges, R.E.M., Forster, P., Bandelt, H.-J., 1997. Reply to Cavalli-Sforza and Minch. American Journal of Human Genetics 61, 251–54.

Robb, J., Miracle, P., 2007. Beyond “migration” versus “acculturation”: new models for the spread of agriculture. In: Whittle, A., Cummings, V. (Eds.), Going over: The Mesolithic-Neolithic Transition in North-West Europe. Oxford, pp. 99–115.

Roostalu, U., Kutuev, I., Loogväli, E.-L., Metspalu, E., Tambets, K., Reidla, M., Khusnutdinova, E.K., Usanga, E., Kivisild, T., Villems, R., 2007. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Molecular Biology and Evolution 24, 436–48.

Rootsi, S., Magri, C., Kivisild, T., Benuzzi, G., Help, H., Bermisheva, M., Kutuev, I., Barać, L., Pericić, M., Balanovsky, O., Pshenichnov, A., Dion, D., Grobei, M., Zhivotovsky, L. a, Battaglia, V., Achilli, A., Al-Zahery, N., Parik, J., King, R., Cinnioğlu, C., Khusnutdinova, E., Rudan, P., Balanovska, E., Scheffrahn, W., Simonescu, M., Brehm, A., Goncalves, R., Rosa, A., Moisan, J.-P., Chaventre, A., Ferak, V., Füredi, S., Oefner, P.J., Shen, P., Beckman, L., Mikerezi, I., Terzić, R., Primorac, D., Cambon-Thomsen, A., Krumina, A., Torroni, A., Underhill, P. a, Santachiara-Benerecetti, a S., Villems, R., Semino, O., 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. American Journal of Human Genetics 75, 128–37.

Rosser, Z.H., Zerjal, T., Hurles, M.E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., Beckman, G., Beckman, L., Bertranpetit, J., Bosch, E., Bradley, D.G., Brede, G., Cooper, G., Côrte-Real, H.B., de Knijff, P., Decorte, R., Dubrova, Y.E., Evgrafov, O., Gilissen, A., Glisic, S., Gölge, M., Hill, E.W., Jeziorowska, A., Kalaydjieva, L., Kayser, M., Kivisild, T., Kravchenko, S.A., Krumina, A., Kucinskas, V., Lavinha, J., Livshits, L.A., Malaspina, P., Maria, S., McElreavey, K., Meitinger, T.A., Mikelsaar, A. V, Mitchell, R.J., Nafa, K., Nicholson, J., Nørby, S., Pandya, A., Parik, J., Patsalis, P.C., Pereira, L., Peterlin, B., Pielberg, G., Prata, M.J., Previderé, C., Roewer, L., Rootsi, S., Rubinsztein, D.C., Saillard, J., Santos, F.R., Stefanescu, G., Sykes, B.C., Tolun, A., Villems, R., Tyler-Smith, C., Jobling, M.A., 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. American Journal of Human Genetics 67, 1526–43.

Ruttkay, E., 1995. Spätneolithikum. In: Jungsteinzeit Im Osten Österreich. Niederösterreichisches Pressehaus, St. Pölten-Wien, p. 224.

Sampietro, M.L., Lao, O., Caramelli, D., Lari, M., Pou, R., Martí, M., Bertranpetit, J., Lalueza-Fox, C., 2007. Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. Proceedings of the Royal Society B 274, 2161–7.

Sánchez-Quinto, F., Schroeder, H., Ramirez, O., Avila-Arcos, M.C., Pybus, M., Olalde, I., Velazquez, A.M. V, Marcos, M.E.P., Encinas, J.M.V., Bertranpetit, J., Orlando, L., Gilbert, M.T.P., Lalueza-Fox, C., 2012. Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. Current Biology 22, 1494–9.

Scally, A., Durbin, R., 2012. Revising the human mutation rate: implications for understanding human evolution. Nature Reviews Genetics 13, 745–53.

Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jäger, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J.L., Kjellström, A., Wu, H., Stewart, G.R., Taylor, G.M., Bauer, P., Lee, O.Y.-C., Wu, H.H.T., Minnikin, D.E., Besra, G.S., Tucker, K., Roffey, S., Sow, S.O., Cole, S.T., Nieselt, K., Krause, J., 2013. Genome-wide comparison of medieval and modern Mycobacterium leprae. Science 341, 179–83.

Schwidetzky, I., 1969. Grazilisation und Degrazilisation: Merkmalsstistische Untersuchungen zur Anthropologie des Neolithikum. HOMO 20, 160–174.

Scozzari, R., Massaia, A., D'Atanasio, E., Myres, N.M., Perego, U.A., Trombetta, B., Cruciani, F., 2012. Molecular dissection of the basal clades in the human Y chromosome phylogenetic tree. PloS ONE 7, e49170.

Sebők, K., 2008. Bölcske, Gyűrűsvölgy. Régészeti Kutatások Magyarországon 2008, 154.

Sebők, K., 2013. Two ceramic-covered burials from the Middle Neolithic of the Carpathian Basin. In: Anders, A., Kulcsár, G. (Eds.), Moments in Time. Budapest, pp. 249–257.

Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., Benedictis, G. De, Francalacci, P., Kouvatsi, A., Limborska, S., Marcikiæ, M., Mika, A., Mika, B., Primorac, D., Santachiara-Benerecetti, A.S., Cavalli-Sforza, L.L., Underhill, P.A., 2000. The genetic legacy of paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 290, 1155–1159.

Sengupta, S., Zhivotovsky, L.A., King, R., Mehdi, S.Q., Edmonds, C.A., Chow, C.-E.T., Lin, A.A., Mitra, M., Sil, S.K., Ramesh, A., Usha Rani, M. V, Thakur, C.M., Cavalli-Sforza, L.L., Majumder, P.P., Underhill, P.A., 2006. Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. American Journal of Human Genetics 78, 202–21.

Serjovic, 1988. The Neolithic of Serbia: a review of research. In: Serjovic, D. (Ed.), The Neolithic of Serbia: Archaeological Research 1948–1988. Center for Archaelogical Research, University of Belgrade., Belgrade, pp. 5–19.

Shaw, I., Jameson, R. (Eds.), 2008. A Dictionary of Archaeology. John Wiley & Sons.

Shennan, S., 2009. Evolutionary demography and the population history of the European early Neolithic. Human Biology 81, 339–55.

Skoglund, P., Malmström, H., Omrak, A., Raghavan, M., Valdiosera, C., Günther, T., Hall, P., Tambets, K., Parik, J., Sjögren, K.-G., Apel, J., Willerslev, E., Storå, J., Götherström, A., Jakobsson, M., 2014. Genomic diversity and admixture differs for Stone-age Scandinavian foragers and farmers. Science 344, 747–750.

Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T.P., Götherström, A., Jakobsson, M., 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science 336, 466–9.

Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139, 457–62.

Slaus, M., 2005. Rezultati antropološke analize ljudskog osteološkog materijala s nalazišta Vela spila kod Vele Luke, Stručni izvještaj EP – 97 – 03/05 Unpublished report for Hrvatska Akademija znanosti i umjet-nosti Odsjek za arhaeologiju. Zagreb.

Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V., Bandelt, H.-J., Torroni, A., Richards, M.B., 2010. The archaeogenetics of Europe. Current Biology 20, R174–83.

Soares, P., Alshamali, F., Pereira, J.B., Fernandes, V., Silva, N.M., Afonso, C., Costa, M.D., Musilová, E., Macaulay, V., Richards, M.B., Cerny, V., Pereira, L., 2012. The Expansion of mtDNA Haplogroup L3 within and out of Africa. Molecular Biology and Evolution 29, 915–27.

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V., Richards, M.B., 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. American Journal of Human Genetics 84, 740–59.

Sokal, R.R., Oden, N.L., Wilson, C., 1991. Genetic evidence for the spread of agriculture in Europe by demic diffusion. Nature 351, 143–5.

Stadler, P., Kotova, N., 2010. Early Neolithic settlement from Brunn Wolfholz from Lower Austria and the problem of the origin of (Western) LBK. In: Kozlowski, J.K., Raczky, P. (Eds.), Neolithization of the Carpathian Basin: Northernmost Distribution of the Starčevo/Körös Culture. Kraków-Budapest, pp. 325–348.

Stadler, P., Ruttkay, E., Doneus, M., Friesinger, H., Lauermann, E., Kutschera, W., Mateiciucová, I., Neubauer, W., Neugebauer-Maresch, C., Trnka, G., Weninger, F., Wild E., M., 2006. Absolutchronologie der Mährisch- Ostösterreichischen Gruppe (MOG) der bemalten Keramik aufgrund von neuen 14C-Datierungen. Absolutchronologie der Mährisch-Ostösterreichischen Gruppe (MOG) der bemalten Keramik aufgrund von neuen 14C-Datierungen,. In: Krenn-Leeb, A. (Ed.), Festschrift E. Ruttkay - Ein Lächeln Für Die Jungsteinzeit. Ausgewählte Beiträge Zum Neolithikum Ostösterreichs. pp. 53–81.

Sümegi, P., Kertész, R., Hertelendi, E., 2002. Environmental Change and Human Adaptation in the Carpathian Basin at the lateglacial/postglacial transition. Archaeometry ’98 Proceedings of the 31th Symposium on Archaeometry, Budapest, April 26–May 3 1998 171–177.

Szécsényi-Nagy, A., Brandt, G., Keerl, V., Jakucs, J., Haak, W., Fecher, M., Moeller-Rieker, S., Köhler, K., Mende, B.G., Oross, K., Marton, T., Osztás, A., Kiss, V., Pálfy, G., Molnár, E., Raczky, P., Anders, A., Sebők, K., Czene, A., Kustár, R., Paluch, T., Somogyi, K., Šlaus, M., Novak, M., Tóth, G., Ősz, B., Voicsek, V., Bánffy, E., Alt, K.W., 2014a. Tracing the genetic

orgigin of Europe's first farmers reveals insights into their social structure. BioRxiv http://dxdoiorg/101101/008664 (In press).

Szécsényi-Nagy, A., Keerl, V., Jakucs, J., Brandt, G., Bánffy, E., Alt, K.W., 2014b. Ancient DNA Evidence for a Homogeneous Maternal Gene Pool in sixth Millennium cal BC Hungary and the Central European LBK. In: Whittle, A., Bickle, P. (Eds.), Early Farmers. The View from Archaeology and Science Proceedings of the British Academy 198. Oxford University Press/British Academy, Oxford, pp. 71–93.

Tajima, F., 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105, 437–60.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–95.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10, 512–26.

Thomas, J., 1996. The cultural context of the first use of domesticates in continental Central and Northwest-wast Europe. In: Harris, D.R. (Ed.), The Origins and Spread of Agriculture and Pastoralism in Eurasia. UCL Press, pp. 310–322.

Thomson, R., Pritchard, J.K., Shen, P., Oefner, P.J., Feldman, M.W., 2000. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. Proc Natl Acad Sci U S A 97, 7360–5.

Thorpe, I.J., 1996. The origins of agriculture in Europe, Geoarchaeology. Routledge, London.

Thrash, J.C., Boyd, A., Huggett, M.J., Grote, J., Carini, P., Yoder, R.J., Robbertse, B., Spatafora, J.W., Rappé, M.S., Giovannoni, S.J., 2011. Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. Scientific Reports 1, 13.

Tilford, C.A., Kuroda-Kawaguchi, T., Skaletsky, H., Rozen, S., Brown, L.G., Rosenberg, M., McPherson, J.D., Wylie, K., Sekhon, M., Kucaba, T.A., Waterston, R.H., Page, D.C., 2001. A physical map of the human Y chromosome. Nature 409, 943–5.

Tillmann, A., 1993. Kontinuität oder Diskontinuität? Zur Frage einer bandkeramischen Landnahme im südlichen Mitteleuropa. Archäologische Informationen 16, 157–87.

Tömöry, G., Csanyi, B., Boga, E., Priskin, K., Downes, C.S., 2007. Comparison of Maternal Lineage and Biogeographic Analyses of Ancient and Modern Hungarian Populations. American Journal of Physical Anthropology 368, 354–368.

Torroni, a, Bandelt, H.J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., Parik, J., Tolk, H. V, Tambets, K., Forster, P., Karger, B., Francalacci, P., Rudan, P., Janicijevic, B., Rickards, O., Savontaus, M.L., Huoponen, K., Laitinen, V., Koivumäki, S., Sykes, B., Hickey, E., Novelletto, A., Moral, P., Sellitto, D., Coppa, A., Al-Zaheri, N., Santachiara-Benerecetti, a S., Semino, O., Scozzari, R., 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. American Journal of Human Genetics 69, 844–52.

Torroni, A., Schurr, T.G., Cabell, M.F., Brown, M.D., Neel, J. V, Larsen, M., Smith, D.G., Vullo, C.M., Wallace, D.C., 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. American Journal of Human Genetics 53, 563–90.

Tringham, R., 2000. Southeastern Europe in the transition to agriculture in Europe: bridge, buffer, or mosaic. In: Price, D. (Ed.), Europe's First Farmers. Cambridge University Press, Cambridge, pp. 19–56.

Underhill, P. a, Poznik, G.D., Rootsi, S., Järve, M., Lin, A. a, Wang, J., Passarelli, B., Kanbar, J., Myres, N.M., King, R.J., Di Cristofaro, J., Sahakyan, H., Behar, D.M., Kushniarevich, A., Sarac, J., Saric, T., Rudan, P., Pathak, A.K., Chaubey, G., Grugni, V., Semino, O., Yepiskoposyan, L., Bahmanimehr, A., Farjadian, S., Balanovsky, O., Khusnutdinova, E.K., Herrera, R.J., Chiaroni, J., Bustamante, C.D., Quake, S.R., Kivisild, T., Villems, R., 2014. The phylogenetic and geographic structure of Y-chromosome haplogroup R1a. European journal of human genetics : EJHG 1–8.

Vajda-Kiss, O., 2009. Lánycsók, M6 116. lelőhely. Régészeti Kutatások Magyarországon 2008, 220–221.

Van Andel, H., Runnels, N., 1995. The earliest farmers in Europe. Antiquity 69, 481–500.

Van Oven, M., Kayser, M., 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Human Mutation 30, E386–94.

Vigne, J.-D., 2008. Zooarchaeological Aspects of the Neolithic Diet Transition in the Near East and Europe, and Their Putative Relationships with the Neolithic Demographic Transition. In: Bocquet-Appel, J.-P., Bar-Yosef, O. (Eds.), The Neolithic Demographic Transition and Its Consequences. Springer, pp. 179–206.

Voicsek, V., 2010. Lánycsók - Gata Csotola (Baranya megye B 67 lelőhely). Évkönyv és jelentés a KÖSZ 2008 évi feltárásairól 23–24.

Von Cramon-Taubadel, N., Pinhasi, R., 2011. Craniometric data support a mosaic model of demic and cultural Neolithic diffusion to outlying regions of Europe. Proceedings of the Royal Society B 278, 2874–80.

Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58, 236–244.

Wells, R.S., Yuldasheva, N., Ruzibakiev, R., Underhill, P. a, Evseeva, I., Blue-Smith, J., Jin, L., Su, B., Pitchappan, R., Shanmugalakshmi, S., Balakrishnan, K., Read, M., Pearson, N.M., Zerjal, T., Webster, M.T., Zholoshvili, I., Jamarjashvili, E., Gambarov, S., Nikbin, B., Dostiev, A., Aknazarov, O., Zalloua, P., Tsoy, I., Kitaev, M., Mirrakhimov, M., Chariev, A., Bodmer, W.F., 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc Natl Acad Sci U S A 98, 10244–9.

Whittle, A., 1996. Europe in the Neolithic. Cambridge University Press, Cambridge.

Whittle, A., Bartosiewicz, L., Borić, D., Pettitt, P., Michael, R., 2002. In the Beginning: New Radiocarbon dates for the Early Neolithic in northern Serbia and southeast Hungary. Antaeus Communicationes ex Instituto Archaeologico Academiae Scientarum Hungaricae 25, 63–117.

Whittle, A., Cummings, V. (Eds.), 2007. Going Over. The Mesolithic-Neolithic transition in North-West Europe, Proceeding. ed. OUP/British Academy, Oxford.

Wilde, S., Timpson, A., Kirsanow, K., Kaiser, E., Kayser, M., Unterländer, M., Hollfelder, N., Potekhina, I.D., Schier, W., Thomas, M.G., Burger, J., 2014. Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. Proc Natl Acad Sci U S A.

Willerslev, E., Cooper, A., 2005. Ancient DNA. Proceedings of the Royal Society B 272, 3–16.

Woodward, S.R., Weyand, N.J., Bunnell, M., 1994. DNA sequence from Cretaceous period bone fragments. Science 266, 1229–32.

Y Chromosome, C., 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Research 12, 339–48.

Yunusbayev, B., Metspalu, M., Järve, M., Kutuev, I., Rootsi, S., Metspalu, E., Behar, D.M., Varendi, K., Sahakyan, H., Khusainova, R., Yepiskoposyan, L., Khusnutdinova, E.K., Underhill, P. a, Kivisild, T., Villems, R., 2012. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. Molecular Biology and Evolution 29, 359–65.

Zagwijn, W.H., 1992. The beginning of the Ice Age in Europe and its major subdivisions. Quaternary Science Reviews 11, 583–591.

Zalai-Gaál, I., 1988. Közép-Európai neolitikus temetők szociálarcheológiai elemzése [Sozialarchäologische Untersuchungendes mitteleuropäischen Schanzwerke im südlichen Transdanubien]. Zalai Múzeum 2, 31–46.

Zalai-Gaál, I., 2002. Die neolithische Gräbergruppe-B1 von Mórágy-Tüzködomb. I. Die archäologischen Funde und Befunde. Szekszárd-Saarbrücken.

Zalai-Gaál, I., 2010. Die soziale Differenzierung im Spätneolithikum Südtransdanubiens., Varia Arch. ed. Archaeolingua/MTA Régészeti Intézet, Budapest.

Zalloua, P. a, Xue, Y., Khalife, J., Makhoul, N., Debiane, L., Platt, D.E., Royyuru, A.K., Herrera, R.J., Hernanz, D.F.S., Blue-Smith, J., Wells, R.S., Comas, D., Bertranpetit, J., Tyler-Smith, C., 2008. Y-chromosomal diversity in Lebanon is structured by recent historical events. American Journal of Human Genetics 82, 873–82.

Zerjal, T., Pandya, A., Thangaraj, K., Ling, E.Y.S., Kearley, J., Bertoneri, S., Paracchini, S., Singh, L., Tyler-Smith, C., 2007. Y-chromosomal insights into the genetic impact of the caste system in India. Human Genetics 121, 137–44.

Zhao, Z., Khan, F., Borkar, M., Herrera, R., Agrawal, S., 2009. Presence of three different paternal lineages among North Indians: a study of 560 Y chromosomes. Annals of Human Biology 36, 46–59.

Zilhão, J., 2001. Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. Proc Natl Acad Sci U S A 98, 14180–14185.

Zischler, H., Höss, M., Handt, O., von Haeseler, A., van der Kuyl, A.C., Goudsmit, J., 1995. Detecting dinosaur DNA. Science 268, 1192–3; author reply 1194.

Zoffmann K, Z., 2005. Embertani leletek a rézkori Balaton-Lasinja kultúra Keszthely-Fenékpuszta lelőhelyen feltárt gödreiből. Zalai Múzeum 14, 53–60.

Zoffmann K., Z., 2004. Õslakosok és bevándorlók a neolitikus és rézkori Kárpát-medencében az embertani adatok alapján. Somogyi Múzeumok Közleményei 16., 127–138.

Zoffmann K., Z., 2005. Embertani adatok a dél-alföldi neolitikum biológiai és történelmi rekonstrukciójához.Anthropological data to the biological and historical reconstruction of the Neolithic of the southern part of the Great Hungarian Plain. In: Bende, L., Lőrinczy, G. (Eds.), Hétköznapok Vénuszai. Hódmezővásárhely, pp. 145–157.

Zoffmann, Z.K., 1978. Excavation at Bicske II. Anthropological finds from the Neolithic cemetery. Alba Regia 16, 61–69.

Zoffmann, Z.K., 2004. A Lengyeli kultúra Mórágy B1 temetkezési csoportjának embertani ismertetése. Wosinszky Mór Múzeum Évkönyve 26, 137–152.

Zvelebil, M., 1989. On the transition to farming in Europe, or what was spreading with the Neolithic: a reply to Ammerman. Antiquity 63, 379–383.

Zvelebil, M., 1998. Genetic and cultural diversity of Europe: a comment on Cavalli-Sforza. Journal of Anthropological Research 54, 411–17.

Zvelebil, M., 2000. The social context of the agricultural transition in Europe. In: Renfrew, C., Boyle, K. (Eds.), Archaeogenetics: DNA and the Population Prehistory of Europe. McDonald Inst., Cambridge.

Zvelebil, M., 2001. The agricultural transition and the origins of Neolithic society in Europe. Documenta Praehistorica 28, 1–26.

Zvelebil, M. (Ed.), 2009. Hunters in Transition: Mesolithic Societies of Temperate Eurasia and Their Transition to Farming. Cambridge University Press.

Zvelebil, M., Lillie, M.C., 2000. Transition to agriculture in eastern Europe. In: Price, D. (Ed.), Europe's First Farmers. Cambridge University Press, Cambridge, pp. 57–92.

11 Glossary of abbreviations used in text, figures, and tables

aDNA	ancient DNA
ALBK	Alföld *Linearbandkeramik* /Linear Pottery culture/population on the Alföld
AMOVA	Analysis of molecular variance
ASHA	Ancestral shared haplotype analysis
BAC	Baalberge culture/population/dataset
BAK	Bronze Age population/dataset from Kazakhstan
BAS	Bronze Age population/dataset from Siberia
BBC	Bell Beaker culture/population/dataset
BEC	Bernburg culture/population/dataset
BL	Balaton-Lasinja culture/population/dataset
BP	Before Present date
cal BC	Calibrated Before Christ date
CAR	Cardial culture/population/dataset
CAT	Catacomb culture/population/dataset
CWC	Corded Ware culture/population/dataset
ENL	Eneolithic population/dataset from the north Pontic-Caspian Steppe
FBC	Funnel Beaker or *Trichterbecher* culture/population/dataset
GDM	Genetic distance map
HGCN	Hunter-gatherer metapopulation/dataset in Central and North Europe
HGE	Hunter-gatherer metapopulation/dataset in Eastern Europe
HGSW	Hunter-gatherer metapopulation/dataset in southwest Europe
HVS	Hyper Variable Segment of the mitochondrial genome
LBK	*Linearbandkeramik* or Linear Pottery culture/population in Central Europe
LBKT	*Linearbandkeramik* or Linear Pottery culture/population in Transdanubia
LGM	Last Glacial Maximum
LGY	Lengyel culture/population/dataset
MDS	Multidimensional Scaling
mtDNA	mitochondrial DNA
NRY	non-recombining part of the human Y-chromosome
NPO	Neolithic population/dataset from Portugal

NQB	Neolithic population/dataset from Basque country and Navarra
np	nucleotide position
PCA	Principal component analysis
PCR	Polymerase chain reaction
PPN	Pre-Pottery Neolithic
PWC	Pitted Ware culture/population/dataset
rCRS	revised Cambridge Reference Sequence
RSC	Rössen culture/population/dataset
RSRS	Reconstructed Sapiens Reference Sequence
SBE	Single Base Extension
SCG	Schöningen group/population/dataset
SNP	Single nucleotide polymorphism
SOP	Sopot culture/population/dataset
STA	Starčevo culture/population/dataset
STR	Short Tandem Repeat
TMRCA	Time of the Most Recent Common Ancestor
TPC	Test of population continuity
TRE	Treilles culture/population/dataset
UC	Únětice or *Aunjetitz* culture/population/dataset
VIN	Vinča culture/population/dataset
YAM	Yamnaya culture/population/dataset

12 Acknowledgements

13 List of Figures in the text

14 List of Tables in the text

15 Supplement

15.1 Supplementary information

15.1.1 Machines used in the laboratory

In Pre-PCR Lab:

Machine	Producer
Dremel Multitool	Dremel Europe B.V.,Breda, Netherlands
Electric Saw, KaVO EWL	KaVO Elektrotechnisches Werk, Neukirchen, Germany
Spot blasting unit P-G 400 (customised)	Harnisch+ Rieth GmbH und Co.KG, Winterbach, Germany
Spot Blasting Unit P-G 400- K (customised)	Harnisch+ Rieth GmbH und Co.KG, Winterbach, Germany
Magnetic Stirrer, Variomag Mobil Direct	Thermo Fisher Scientific, Langenselbold, Germany
Immersion UV lamp (customised)	UV-Consulting Peschl e.K., Mainz, Germany
Mixer Mill, MM200	Retsch GmbH, Haan
Reverse Osmosis Water Purification System,	IEM-Industrial Equipment and Machinery GmbH, Mainz, Germany
pH-Meter FiveEasy LE409	Mettler Toledo GmbH, Gießen, Germany
Vortexer REAX 1R	Heidolph, Wiesbaden
Micro-centrifuge 120	(Andreas) Hettich GmbH & Co.KG, Tuttlingen, Germany
Centrifuge Universal 320	(Andreas) Hettich GmbH & Co.KG, Tuttlingen, Germany
Pipettes 5ml, 1000µl, 200µl, 20µl, 10µl	ABIMed GmbH, Langenfeld, Germany
Vortexer M52 Minishaker	IKA-Werk, Janke & Kunkel GmbH & Co.KG, .
Hybridisation Oven RPM6	Appligene S.A., Strasbourg, France (now Appligene Oncor S.A.)
Rotator Stuart SB2	Bibby Scientific Ltd., Staffordshire, UK
Thermomixer Thermostart Plus	Eppendorf GmbH, Hamburg, Germany
Canon IXUS S 105 Camera	Canon
Tripod, Cullmann 2800	Cullmann

Precision scale, types 440 / EMB	Kern & Sohn GmbH, Balingen-Frommern, Germany
Racks	Carl Roth Ltd., Karlsruhe, Germany
Glassware	Carl Roth Ltd., Karlsruhe, Germany, Schott Ltd. Mainz, Germany
Metering pump	Carl Roth Ltd., Karlsruhe, Germany

In Post-PCR Lab:

Machine	**Producer**
ABI PRISM™ 3130 Genetic Analyzer	Applied Biosystems, Weiterstadt, Germany
Digital Agarose Gel Documentation System	INTAS Science Imaging Instruments GmbH, Göttingen, Germany
Transilluminator	INTAS Science Imaging Instruments GmbH, Göttingen, Germany
Electrophoresis Chamber	Carl Roth GmbH, Karlsruhe, Germany
Consort E143 Mini Electrophoresis Power Supply	Sigma-Aldrich Chemie GmbH, Taufkirchen bei München, Germany
Eppendorf® Biophotometer	eppendorf AG, Hamburg, Germany
Eppendorf® Centrifuge 5415C	eppendorf AG, Hamburg, Germany
Eppendorf® Mastercycler gradient	eppendorf AG, Hamburg, Germany
Eppendorf® Mastercycler nexus gradient	eppendorf AG, Hamburg, Germany
Eppendorf Mastercycler	eppendorf AG, Hamburg, Germany
Eppendorf® Thermomixer comfort 5355	eppendorf AG, Hamburg, Germany
EQUIBIO Easyject Prisma Elektroporator	peQLab Biotechnologie GmbH, Erlangen, Germany
Combimag RCT Magnetic Stirrer	IKA-Werke GmbH & CO. KG, Staufen, Germany
Water Filtration Unit "Umwelt 3", inc. 40l tank	Novodirect GmbH, Kehl/Rhein, Germany
Pipettes 5ml, 1000µl, 200µl, 20µl, 10µl	ABIMed GmbH, Langenfeld, Germany
MILIPORE Milli-Qplus Filteranlage	Merk Millipore Ltd., Darmstadt, Germany
Incubator,Type B 5042	Heraeus, Hanau, Germany
VWR Lab Dancer Mini Vortexer	IKA-Werke GmbH & CO. KG, Staufen, Germany
Sterilizer M3-20-EC	Midmark Corp., Versailles, OH, USA
Precision scale	Kern & Sohn GmbH, Balingen-Frommern, Germany
Microwave Oven	

15.1.2 Laboratory consumables

Article	Producer
0,5 ml Safe Lock Tubes	Eppendorf AG, Hamburg, Germany
0,5 ml microtubes	Saarstedt AG, Nümbrecht, Germany
0,5 ml DNA LoBind tubes	Eppendorf AG, Hamburg, Germany
1,5 ml microtubes	Saarstedt AG, Nümbrecht, Germany
0.2ml microtube strips	Biozym Scientific Ltd., Hessisch Oldendorf, Germany
0.2 microtube strips	Saarstedt AG, Nümbrecht, Germany
2 ml Eppendorf Tubes	Saarstedt AG, Nümbrecht, Germany
50 kDa Amicon Ultra	Merk Millipore Ltd., Darmstadt, Germany
15ml, 50ml Falcons,	Saarstedt AG, Nümbrecht
ART® 1000E, 100E, 20P, 10 REACH pipette tips	Molecular Bio Products™ , San Diego
5ml pipette tips	eppendorf AG, Hamburg, Germany
Electroporation cuvette	peQLab Biotechnologie GmbH
SHIELD polythene oversleeves, DA01	HPC Healthline Ltd., Morden, Surrey, UK
TyvekR Overshoes, model POBO	DuPont de Nemour GmbH, Neu-Isenburg, Germany
TyvekR Classic Coveralls, model CHF5	DuPont de Nemour GmbH, Neu-Isenburg, Germany
Surgical face masks	Hansa Medical/ Industrial, Kirchhoff Group, Hamburg, Germany
Head covers	Hansa Medical/ Industrial, Kirchhoff Group, Hamburg, Germany
Parafilm M(R) laboratory film	Pechiney Plastic Packaging, Inc., Chicago, IL, USA/ Bemis Flexible Packaging, Inc.,Neenah, WI, USA
Non-sterile Latex Examination Gloves Hansa-Medical 24 Shield-superb	Hansa Medical/ Industrial, Kirchhoff Group, Hamburg, Germany
Nitril Gloves	Hansa Medical/ Industrial, Kirchhoff Group, Hamburg, Germany
MultiScreen PCR µ96 Filter Plate	Merk Millipore Ltd., Darmstadt, Germany
MultiScreen384 SEQ Filter Plates	Merk Millipore Ltd., Darmstadt, Germany
Fiber-free precision wipes Kimtech	Kimberly-Clark Inc., Roswell, GA, USA
96-well sequencing plates	Applied Biosystems, Weiterstadt, Germany
96-well plate septa	Applied Biosystems, Weiterstadt, Germany

Weighing paper	Carl Roth Ltd., Karlsruhe, Germany
Petri dishes	Carl Roth Ltd., Karlsruhe, Germany

15.1.3 R Scripts

15.1.3.1 R script example of the PCA, plotting PC1-2

```
rm(list=ls(all=TRUE))
data <- read.csv2("C:/ PCA/PCA pop/ Slatkin LGY.csv", header=TRUE, row.names=1)

PCA <- prcomp(data, scale.=T)
scores <- PCA$x
scores1 <- scores[,1] * -1
scores2 <- scores[,2] *-1
scores
write.csv(scores, file ="C:/Users/Anna/Desktop/ PCA/PCA pop/Slatkin LGY_scores.csv",
append = FALSE, quote = TRUE, sep = "", eol ="\n", na = "NA", dec = ".", row.names = TRUE,
col.names = TRUE, qmethod = c("escape","double"))
summary(scores)
summary(PCA)
plot(PCA, type = "lines")

color.out = rep(c("black","black","black","black"), times=c(30,0,0,0))
color.in = rep(c("dodgerblue4", "cornflowerblue", "paleturquoise1", "deepskyblue4",
"steelblue1", "steelblue4", "Yellowgreen", "peru"), times=c(5,4,3,14,4,12,31,6))
symbol.pop=rep(c(16,15,0,17,18), times=c(12,8,21,14,42))
symbol.in=rep(c("black","black","white","black","white"), times=c(3,8,21,14,42))

SUMMARY <- summary(PCA)
        PC1 <- SUMMARY$importance[2,1]
        PC1 <- PC1 * 100
        PC1 <- round(PC1,1)
```

```
xlabel <- paste("PC 1 - ",PC1,"% variance")
        PC2 <- SUMMARY$importance[2,2]
        PC2 <- PC2 * 100
        PC2 <- round(PC2,1)

ylabel <- paste("PC 2 - ",PC2,"% variance")
        plot(scores1,scores2,
        pch=21,
        col=color.out,
        bg=color.in,
        cex=1.5,
        cex.axis=0.9,
        font.lab=2,
        font.axis=2,
        xlab=xlabel,
        ylab=ylabel,)

        text(scores1,scores2, row.names(data), font=1, cex=0.45, pos=1, col="black")
vector.exp <- 6
VAR1<-(PCA$rotation[,1])*vector.exp *-1
VAR2<-(PCA$rotation[,2])*vector.exp *-1
VAR12<-cbind(VAR1,VAR2)
points(VAR12,col="grey50",type="n")
text(VAR12,rownames(VAR12), cex=0.6, font=2, pos=1, col="grey50")
        for (i in 1:length(VAR1)) {
        segments(0,0,VAR1[i],VAR2[i],col="grey50",lty=3,lwd=0.5)
        }
```

15.1.3.2 R script of Ward clustering

```
scores1 <-read.csv("C: /Ward Clustering/regions transdanubia _scores.csv",header=TRUE,
row.names=1)
require(graphics)
```

```
library(pvclust)
fit <- pvclust(scores1, method.hclust="ward",method.dist="euclidean")
plot(fit)
pvrect(fit, alpha=.95)
```

15.1.3.3 R script example of MDS

```
rm(list=ls(all=TRUE))
library(vegan)
matrix <- read.csv("C:/Users/Anna/Desktop/MDS/transd.csv", header=TRUE, row.names=1)
fit <- metaMDS(matrix, distance="euclidean", k=2, trymax=50, autotransform=FALSE)
fit$points
fit
color.out = rep(c("black","black","black","black"), times=c(30,0,0,0))
color.in = rep(c("grey50", "red","peru", "orange", "gold", "goldenrod2", "yellow", "red",
"brown"), times=c(4,4,4,5,3,0,6,0,6))
x <- fit$points[,1]
y <- fit$points[,2]
plot (x, y,
        xlab="Coordinate 1",
        ylab="Coordinate 2",
        main="",
        type = "p",
        pch=21,
        col=color.out,
        bg=color.in,
        cex=1.5,
        cex.axis=0.9,
        font.lab=2,
        font.axis=2)
text(x, y, labels = row.names(matrix), font=1, cex=0.6, pos=1, col="black")
```

```
# Shepard plot
label <- paste("Stress - ",fit$stress)
stressplot(fit,
        main=label,
        pch=21,
        p.col="black",
        bg="grey70",
        l.col="black",
        lwd=2)
```

15.1.3.4 Script levelplot (an example)

```
rm(list=ls(all=TRUE))
library(lattice)
matrix <- read.csv("C: /29POP_Fst.csv", header=TRUE, row.names=1)
m <- as.matrix(matrix)
rotmatrix <- t(m)[ , ncol(m):1]

col.l <- colorRampPalette(c("white", "grey", "grey2", "black" ))
levelplot(rotmatrix, col.regions = col.l, cuts = 64, xlab = "", ylab = "",
scales=list(x=list(rot=90)))
```

15.1.3.5 Script Benjamin-Hochberg correction

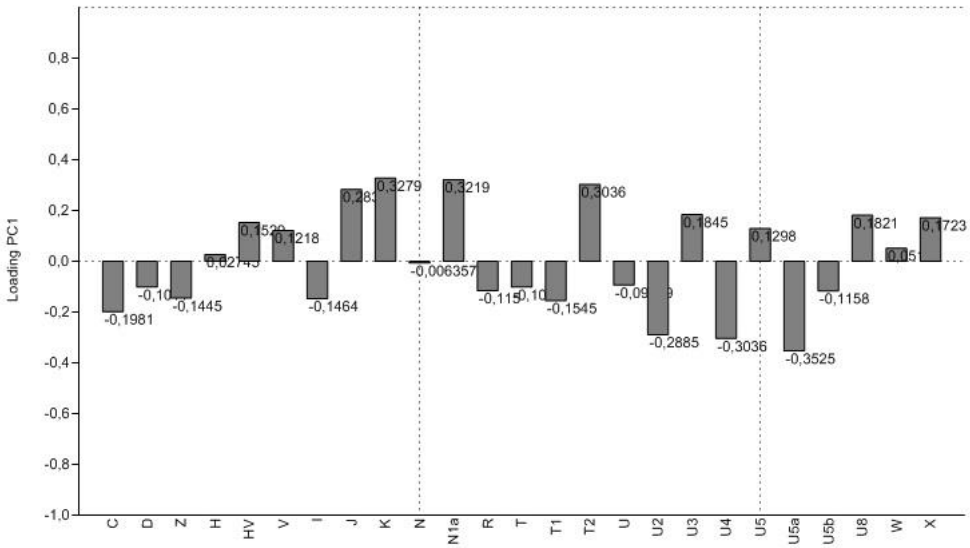
```
P <-c (matrix)
p.BH <- p.adjust(p, "BH")
p.BH
```

15.2 Supplementary figures

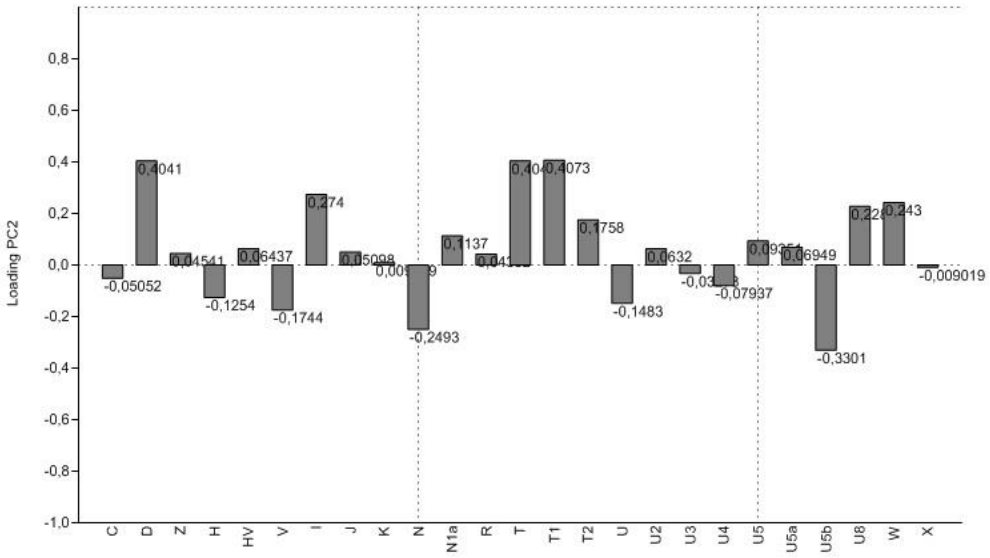
15.2.1 List of Supplementary Figures

Supplementary Figure 1a-b. Variable correlations with the PC1 (A) and PC2 (B) of the PCA with 29 prehistoric populations and metapopulations (corresponding to Figure 14).

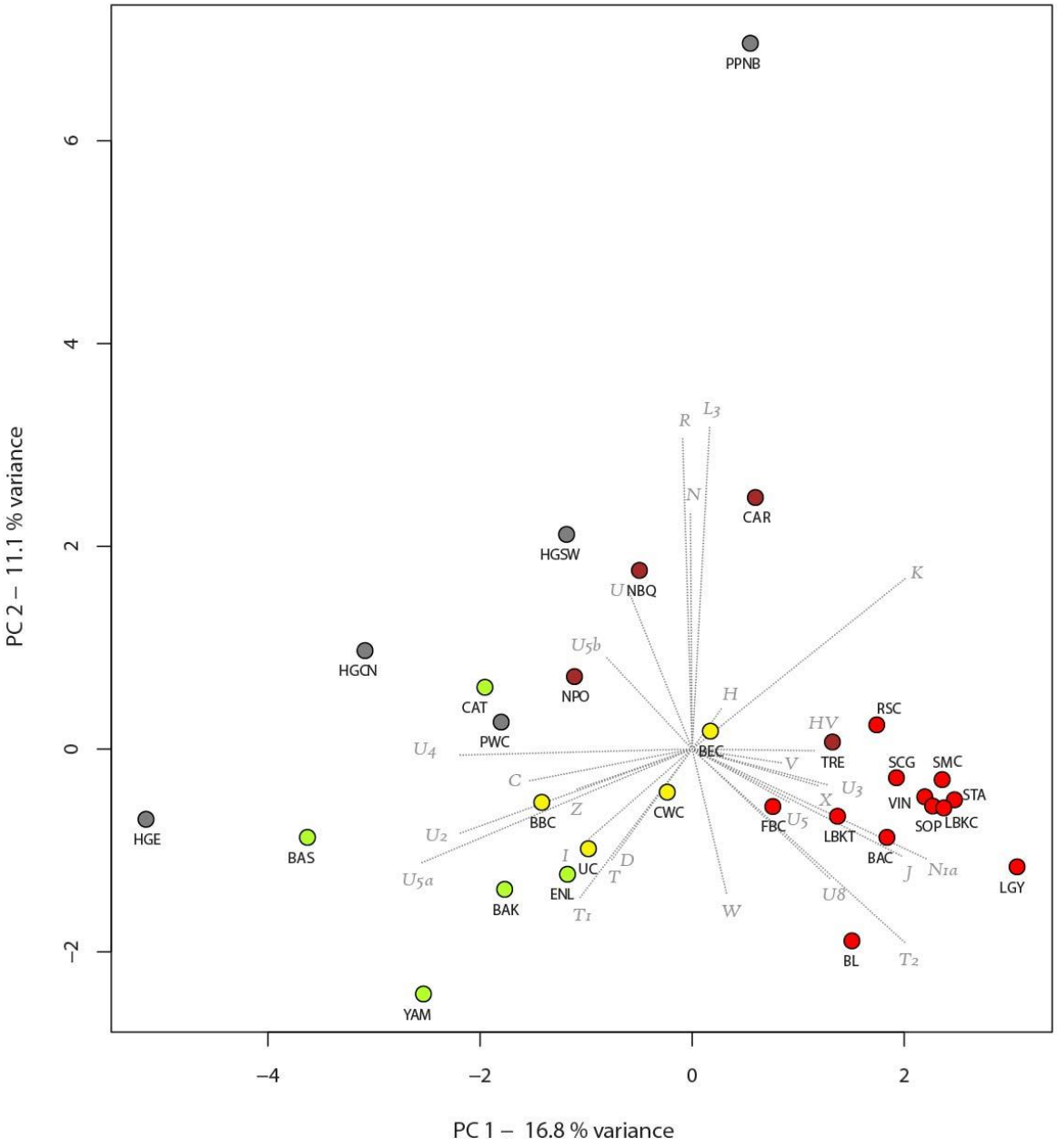
A: PC1:



B: PC2:

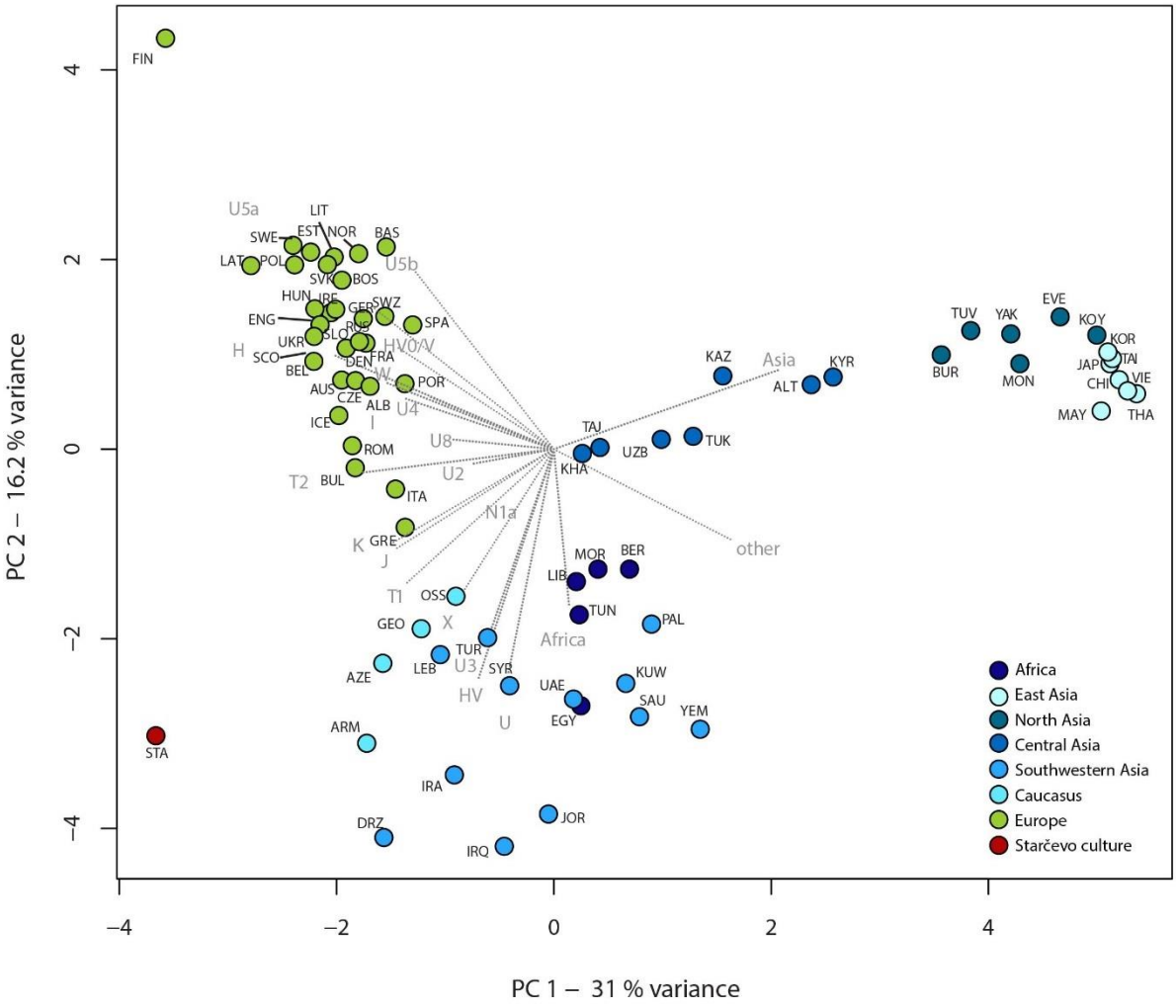


Supplementary Figure 2. PCA with the Pre-Pottery Neolithic B from Levant, containing 30 prehistoric populations.

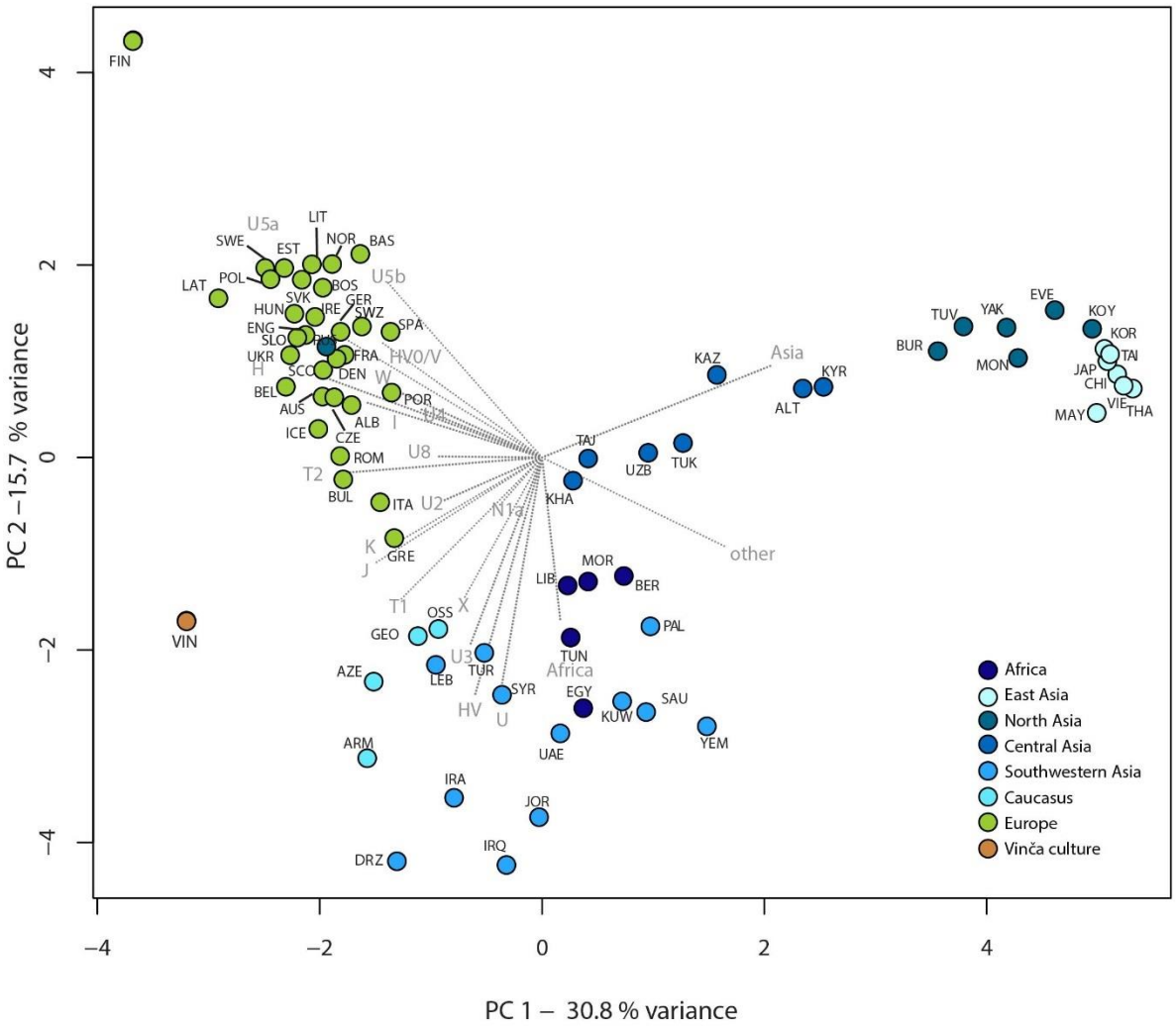


Supplementary Figure 3a-e. Two dimensional PCA with 73 modern populations and the Starčevo (a), LBKT (b), Vinča (c), Sopot (d), Lengyel (e) datasets.

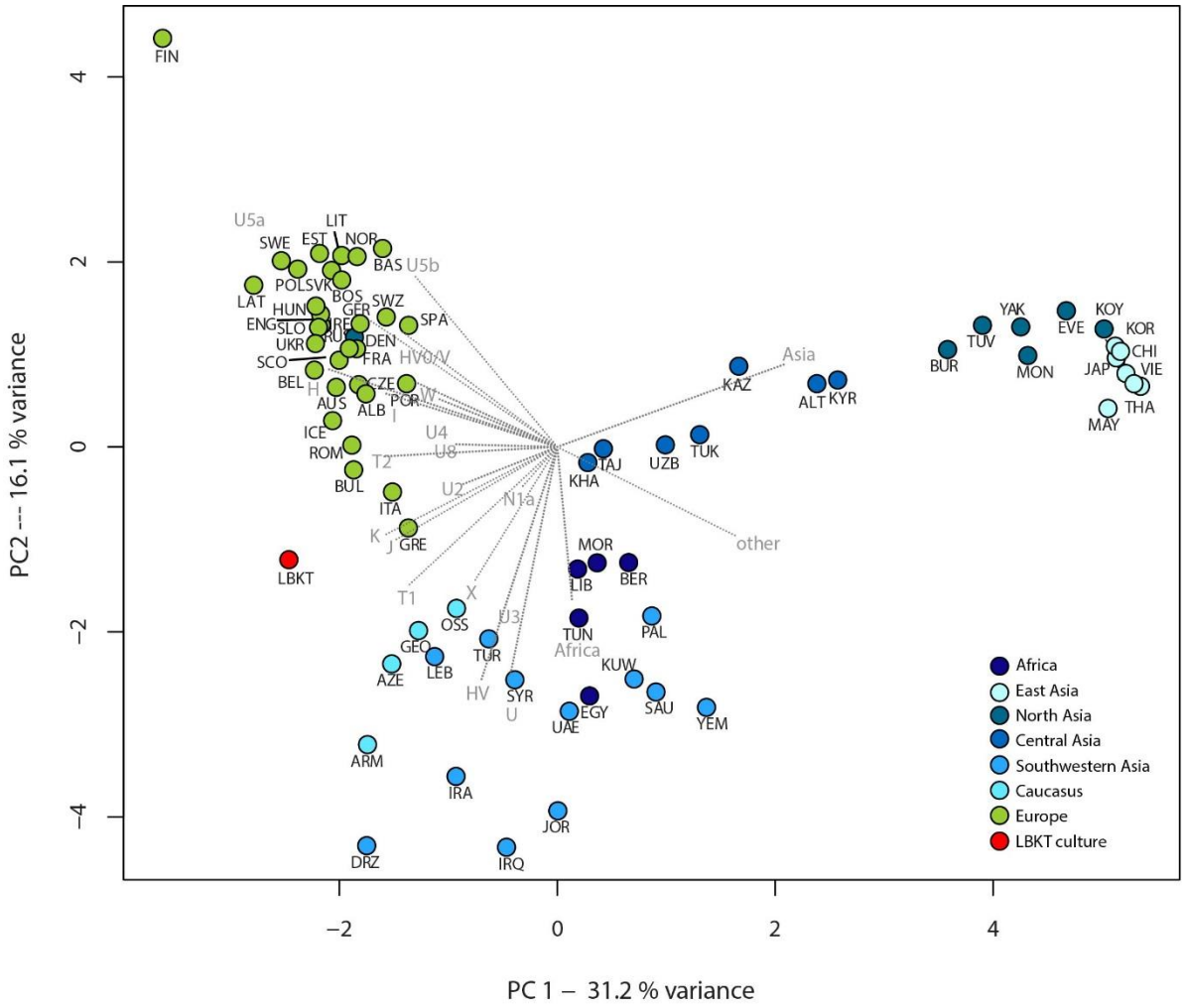
Population codes: Albanians, Macedonians (ALB), Altai (ALT), Arabs in UAE, Oman, Qatar (ARA), Armenians (ARM), Austrians (AUS), Azeris (AZE), Basques (BAS), Belarusians (BEL), Berber (BER), Bosnians, Croatians, Serbians (BOS), British (ENG), Bulgarians (BUL), Buryats (BUR), Chinese, Tibetan (CHI), Czechs (CZE), Danes (DEN), Druze (DRZ), Egyptians (EGY), Estonians (EST), Evenks (EVE), Finns (FIN), French (FRA), Georgians (GEO), Germans (GER), Greeks (GRE), Hungarians (HUN), Icelanders (ICE), Iranians (IRA), Iraqi (IRQ), Irish (IRE), Italians (ITA), Japanese (JAP), Jordanians (JOR), Kazakhs (KAZ), Koreans (KOR), Koryaks (KOY), Kuwaiti (KUW), Kyrgyz (KYR), Latvians (LAT), Lebanese (LEB), Libyans (LIB), Lithuanians (LIT), Malays (MAY), Khants, Mansi (KHA), Mongolians (MON), Moroccans (MOR), Norwegians (NOR), Ossetians (OSS), Palestinians (PAL), Poles (POL), Portuguese (POR), Romanians (ROM), Russians (RUS), Saudi Arabians (SAU), Scots (SCO), Slovaks (SVK), Slovenians (SLO), Spaniards (SPA), Swedes (SWE), Swiss (SWZ), Syrians (SYR), Taiwanese (TAI), Tajiks (TAJ), Thai (THA), Tunisians (TUN), Turkmens (TUK), Turks (TUR), Tuvinians (TUV), Ukrainians (UKR), Uzbeks (UZB), Vietnamese (VIE), Yakuts, Yukaghir (YAK), Yemenis (YEM). The prehistoric cultures are abbreviated as: Balaton-Lasinja culture (BL), Linearbandkeramik culture in Transdanubia (LBKT), Lengyel culture (LGY), Sopot culture (SOP), Starčevo culture (STA), Vinča culture (VIN).



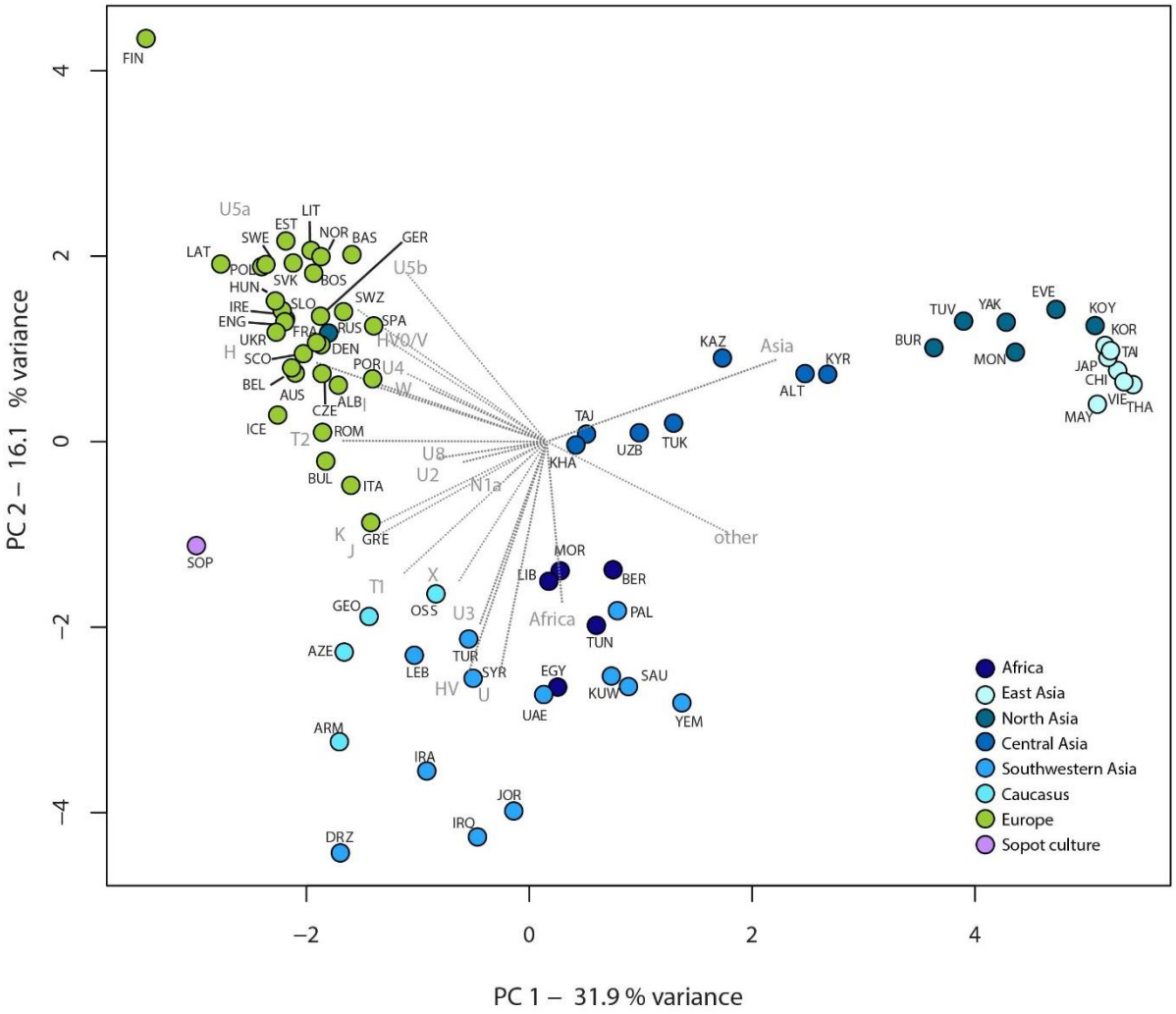
B:


PC 2 – 15.7 % variance
PC 1 – 30.8 % variance
FIN
VIN
LIT
SWE
EST
NOR
BAS
LAT
POL
BOS
SVK
GER
HUN
IRE
SWZ
ENG
SPA
SLO
FRA
UKR
SCO
DEN
BEL
POR
AUS
ALB
ICE
CZE
ROM
BUL
ITA
GRE
U5a
U5b
HV0/V
W
I
U8
T2
U2
K
J
N1a
other
KAZ
Asia
KYR
ALT
TAJ
UZB
TUK
KHA
MOR
LIB
BER
T1
OSS
X
GEO
PAL
TUN
U3
TUR
AZE
LEB
SYR
Africa
EGY
SAU
KUW
HV
U
YEM
UAE
ARM
IRA
JOR
IRQ
DRZ
TUV
YAK
EVE
KOY
KOR
BUR
MON
TAI
JAP
CHI
VIE
THA
MAY
Africa
East Asia
North Asia
Central Asia
Southwestern Asia
Caucasus
Europe
Vinča culture

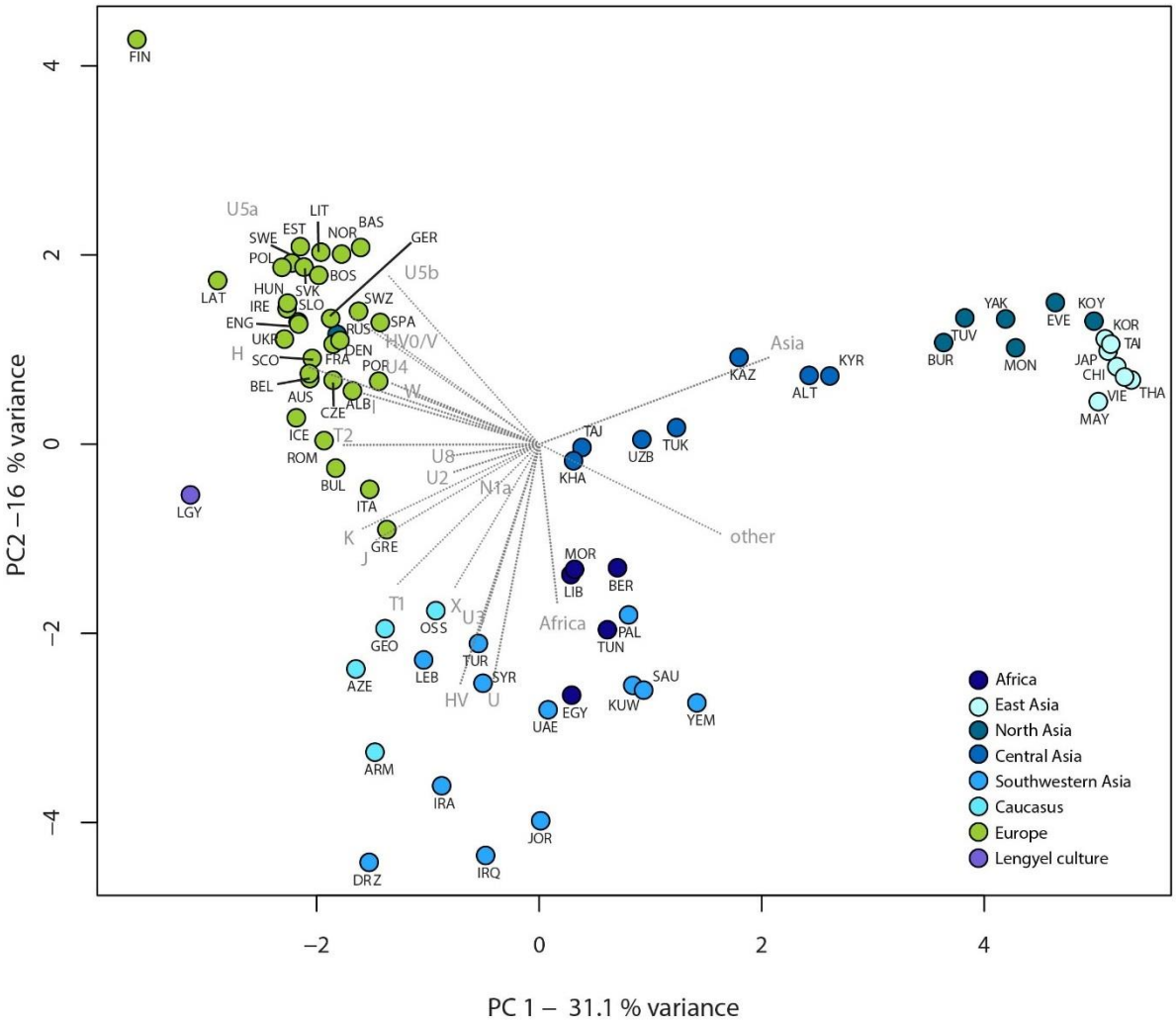
C:


PC2 --- 16.1 % variance
PC 1 – 31.2 % variance
FIN
LIT
EST
NOR
SWE
BAS
POLSVK
LAT
BOS
SWZ
HUN
GER
ENG
SLO
SPA
UKR
DEN
SCO
FRA
BEL
CZE
AUS
ALB
POR
ICE
ROM
BUL
ITA
GRE
LBKT
U5a
U5b
HV0/V
W
I
U4
U8
T2
U2
N1a
K
J
T1
X
U3
OSS
GEO
TUR
LEB
AZE
HV
SYR
U
Africa
ARM
IRA
JOR
DRZ
IRQ
Asia
KAZ
ALT
KYR
TAJ
KHA
UZB
TUK
other
MOR
LIB
BER
TUN
PAL
KUW
SAU
EGY
UAE
YEM
YAK
TUV
BUR
KOY
EVE
KOR
CHI
JAP
MON
VIE
THA
MAY
Africa
East Asia
North Asia
Central Asia
Southwestern Asia
Caucasus
Europe
LBKT culture

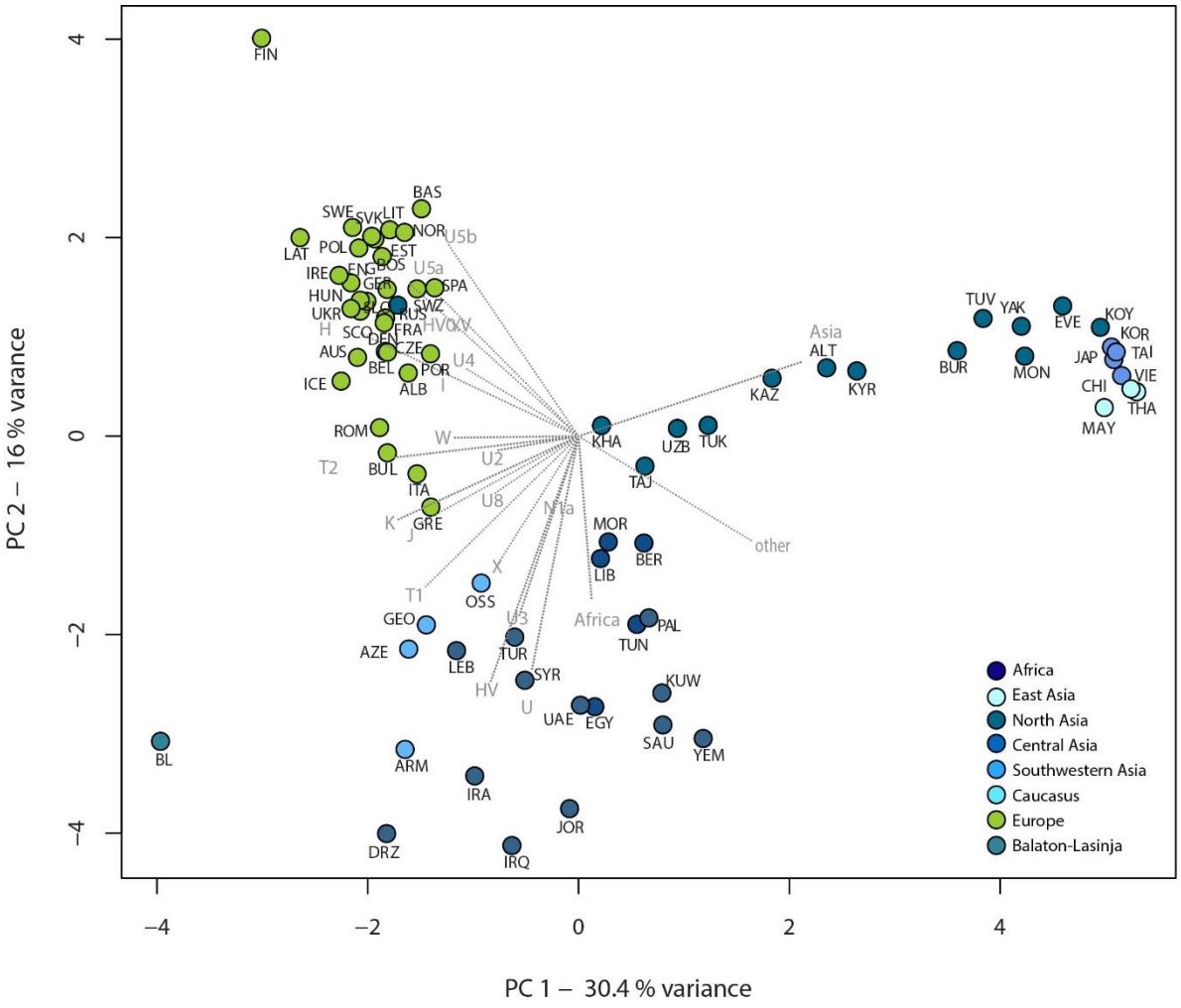
D:

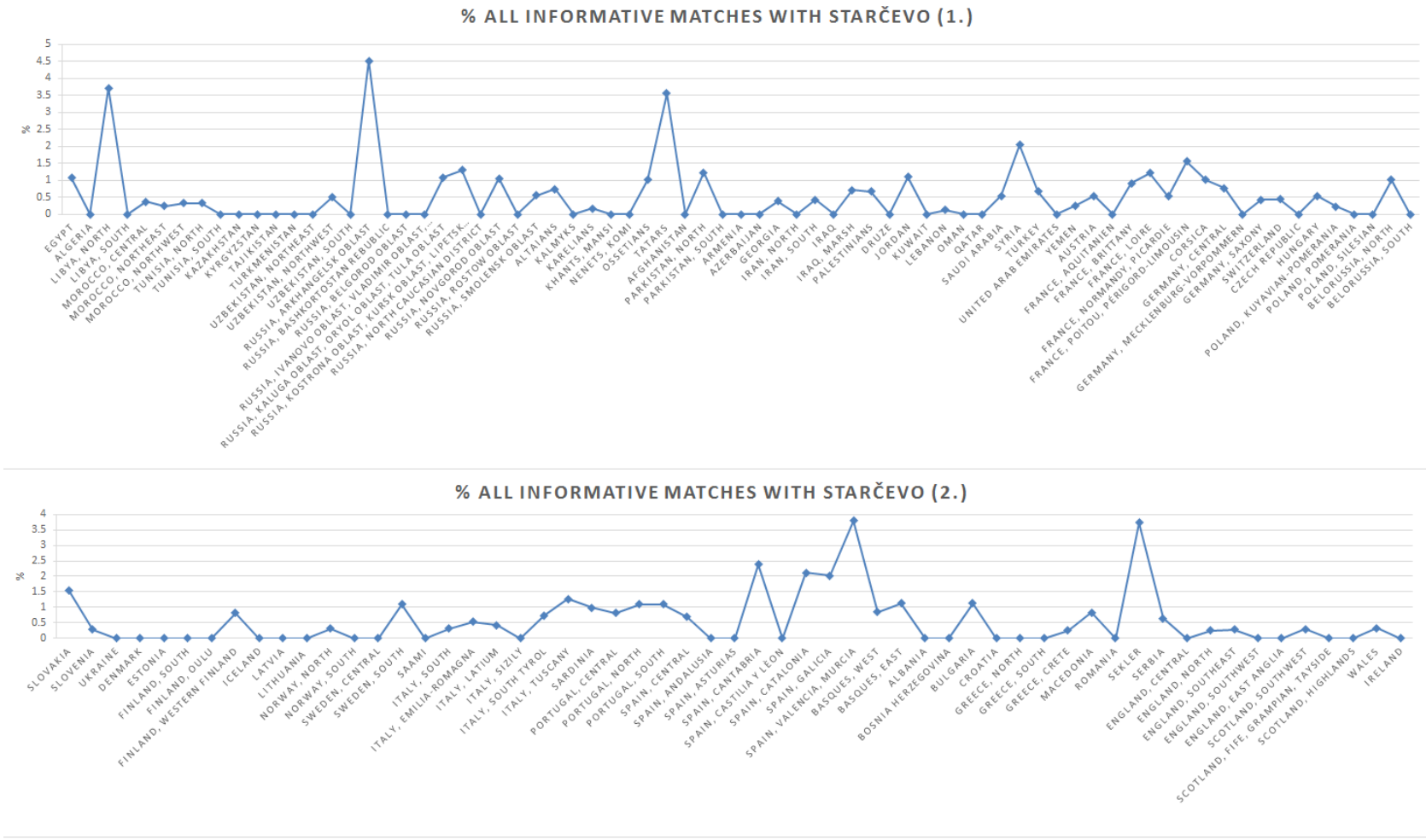

PC 2 – 16.1 % variance
PC 1 – 31.9 % variance
FIN
LAT
POL
SWE
EST
LIT
NOR
BAS
GER
SVK
BOS
HUN
SWZ
IRE
SLO
ENG
SPA
RUS
UKR
FRA
DEN
SCO
POR
BEL
AUS
ALB
CZE
ICE
ROM
BUL
ITA
GRE
SOP
U5a
U5b
H
HV0/V
U4
W
T2
U8
U2
N1a
K
J
other
X
T1
GEO
OSS
U3
AZE
TUR
LEB
SYR
HV
U
ARM
IRA
JOR
IRQ
DRZ
MOR
LIB
BER
Africa
PAL
TUN
EGY
SAU
KUW
UAE
YEM
KAZ
Asia
KYR
ALT
TAJ
UZB
TUK
KHA
TUV
YAK
EVE
KOY
BUR
MON
KOR
TAI
JAP
CHI
VIE
THA
MAY
Africa
East Asia
North Asia
Central Asia
Southwestern Asia
Caucasus
Europe
Sopot culture

E:


PC2 −16 % variance
PC 1 − 31.1 % variance
FIN
LAT
LGY
U5a
LIT
BAS
SWE
EST
NOR
GER
POL
BOS
U5b
HUN
SVK
SWZ
IRE
SLO
ENG
RUS
SPA
UKR
DEN
HV0/V
H
SCO
FRA
POR
J4
BEL
AUS
W
CZE
ALB
ICE
T2
ROM
U8
BUL
U2
N1a
ITA
K
GRE
J
T1
X
OSS
U3
GEO
TUR
LEB
AZE
SYR
HV
U
ARM
IRA
JOR
DRZ
IRQ
Asia
KAZ
KYR
ALT
TAJ
UZB
TUK
KHA
other
MOR
LIB
BER
Africa
PAL
TUN
EGY
UAE
KUW
SAU
YEM
YAK
EVE
KOY
TUV
KOR
TAI
BUR
MON
JAP
CHI
VIE
THA
MAY
Africa
East Asia
North Asia
Central Asia
Southwestern Asia
Caucasus
Europe
Lengyel culture

F:


PC 2 – 16 % varance
PC 1 – 30.4 % variance
FIN
BAS
SWE
SVK
LIT
NOR
LAT
POL
EST
U5b
IRE
ENG
BOS
U5a
GER
SPA
HUN
SWZ
UKR
RUS
H
SCO
FRA
HV0
DEN
CZE
AUS
U4
BEL
POR
ICE
ALB
I
ROM
W
KHA
UZB
TUK
T2
BUL
U2
TAJ
ITA
U8
N1a
K
GRE
J
MOR
other
BER
LIB
X
T1
OSS
GEO
U3
Africa
PAL
TUN
AZE
LEB
TUR
SYR
HV
KUW
U
UAE
EGY
SAU
YEM
BL
ARM
IRA
JOR
DRZ
IRQ
Asia
ALT
KAZ
KYR
TUV
YAK
EVE
KOY
KOR
BUR
MON
JAP
TAI
VIE
CHI
THA
MAY
Africa
East Asia
North Asia
Central Asia
Southwestern Asia
Caucasus
Europe
Balaton-Lasinja



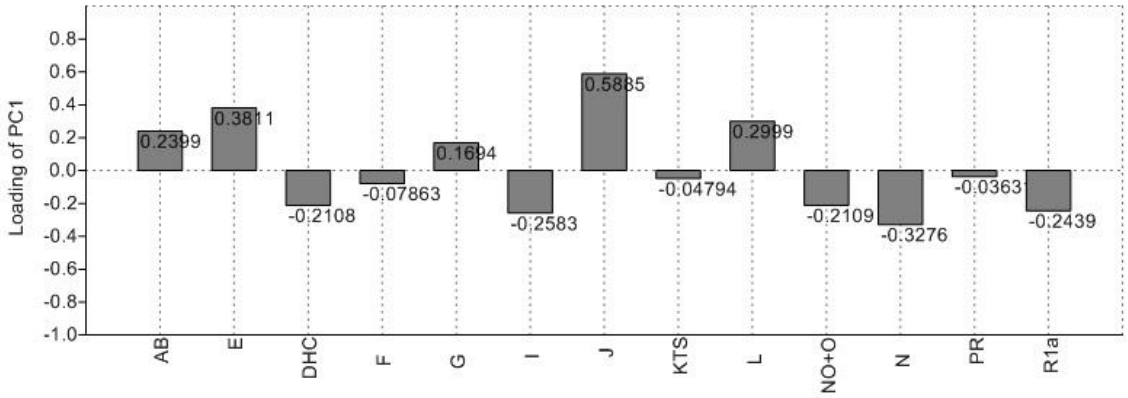
Supplementary Figure 4. Shared haplotypes among the Starčevo dataset and 130 modern Eurasian populations.

Percentage of informative matches were indicated on this two diagrams.

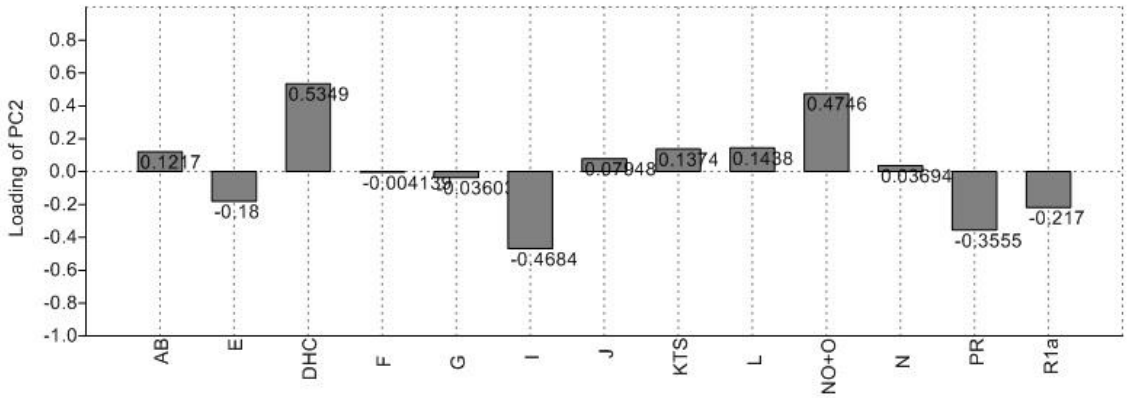
Supplementary Figure 5a-c. Variable correlations with the PC1-3 of the PCA with Y-chromosomal data of the STA-LBKT dataset and 80 modern populations.

The values are given as coefficient values. The data were plotted in PAST program, based on the original haplogroup frequency table, used for the PCA presented in Figure 36.

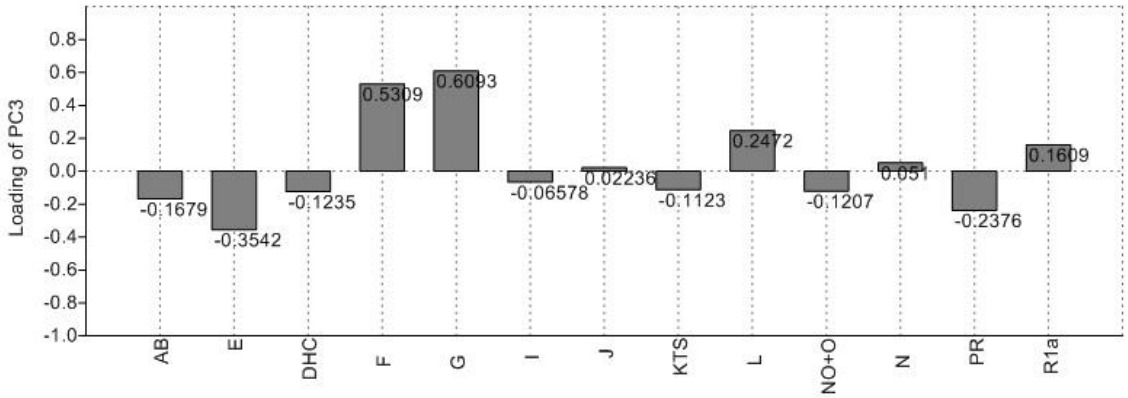
A: PC1:



B: PC2:



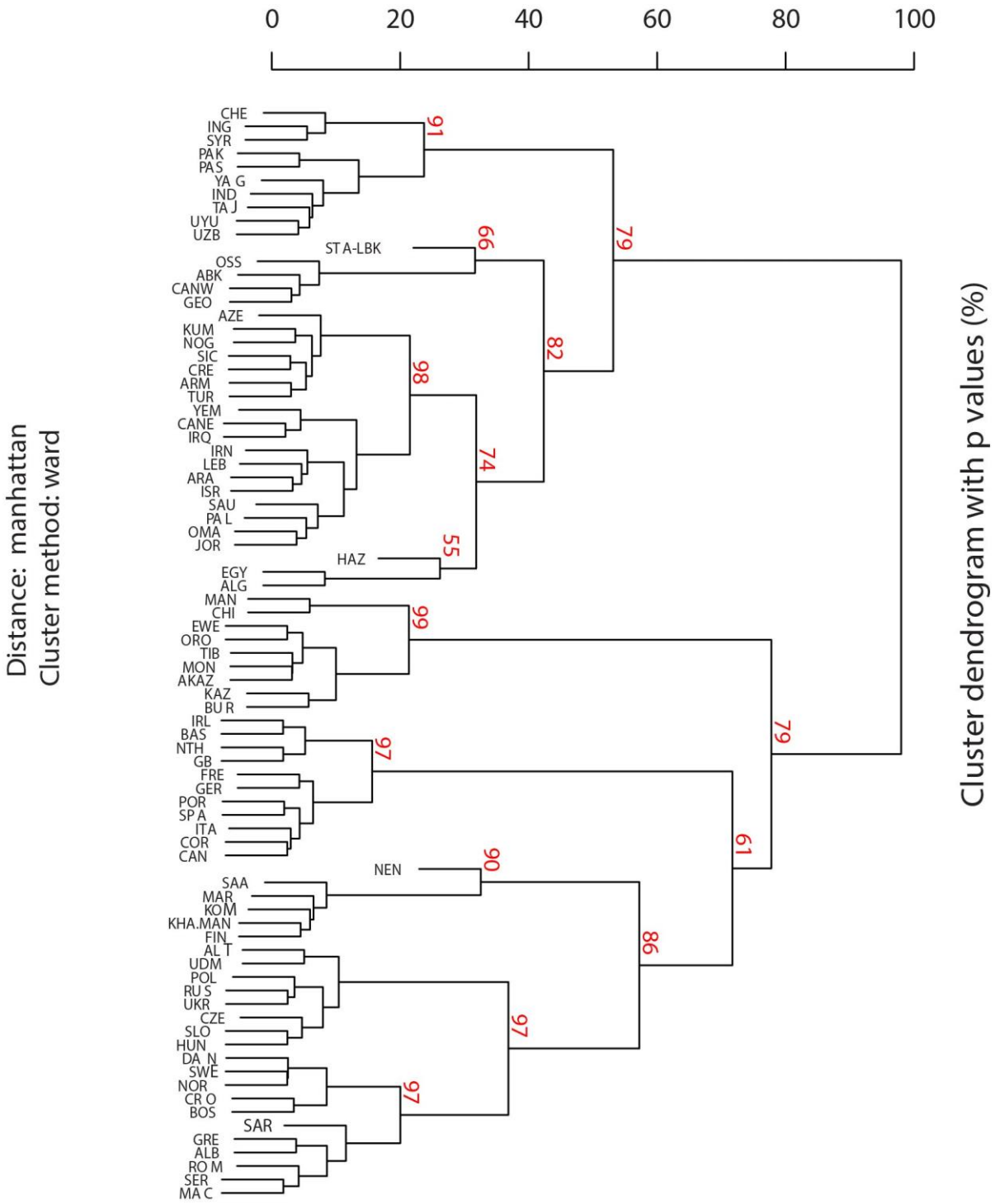
C: PC3:

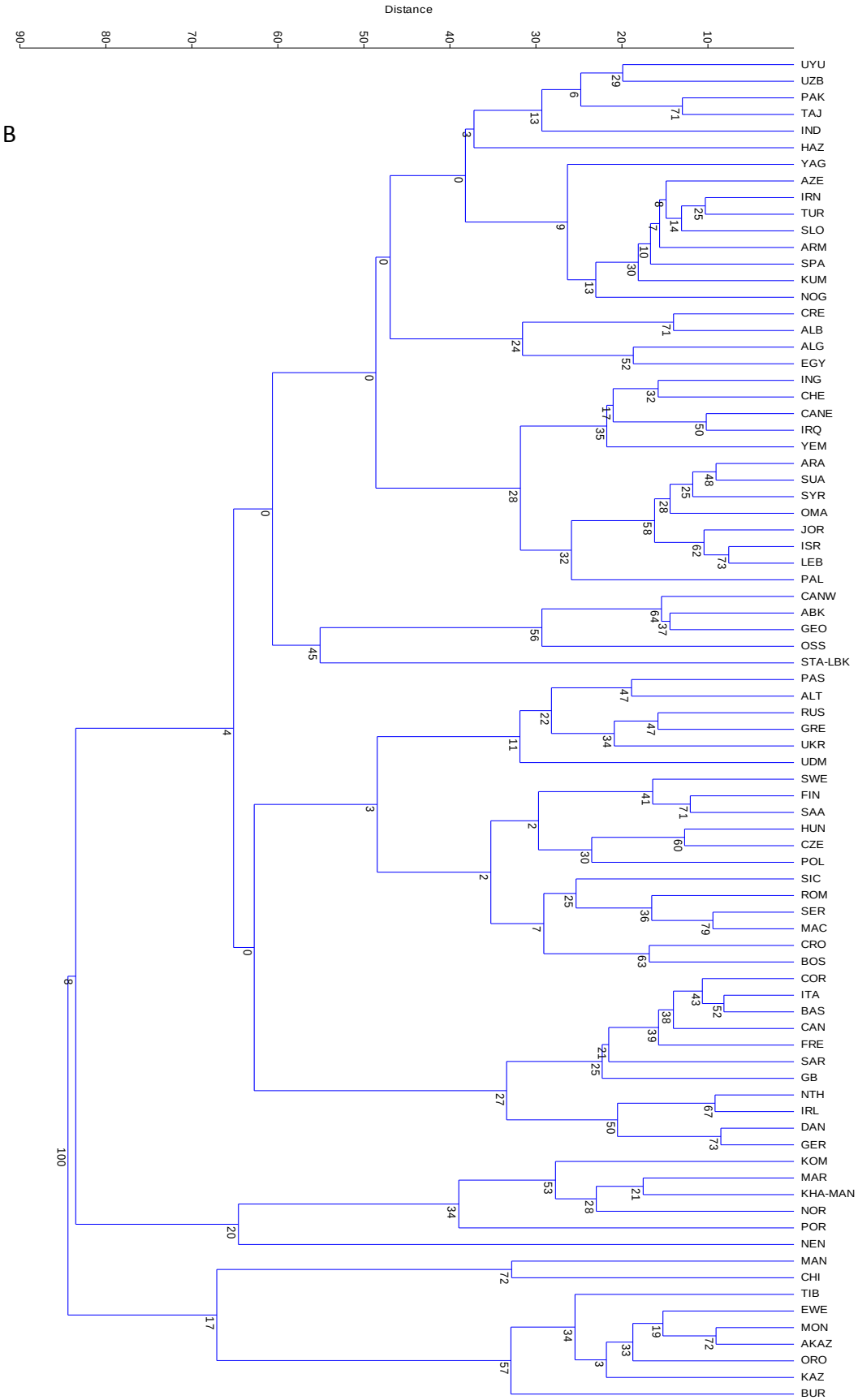


Supplementary Figure 6a-b. Clustering of ward type (A) and paired group (B) with the STA-LBK Y-chromosomal dataset and 80 modern populations.

For population codes see Figure 36. On the Figure B, the percentage of random replicates of bootstrapping, where the cluster is still supported (containing the same set of taxa) is given at the root of the clusters. The copphenetic coefficient was 0.8395. Paired group algorithm was used with correlation similarity measuring method.

A

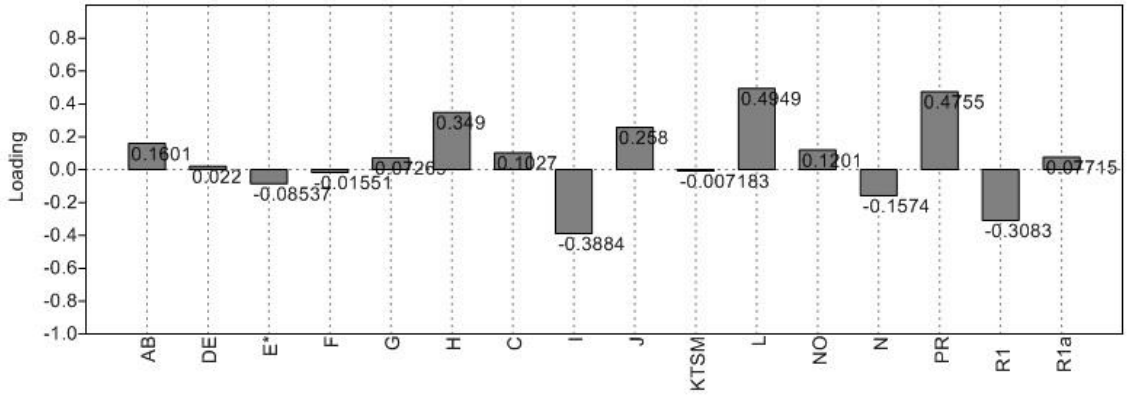



B
Distance
90
80
70
60
50
40
30
20
10
UYU
UZB
PAK
TAJ
IND
HAZ
YAG
AZE
IRN
TUR
SLO
ARM
SPA
KUM
NOG
CRE
ALB
ALG
EGY
ING
CHE
CANE
IRQ
YEM
ARA
SUA
SYR
OMA
JOR
ISR
LEB
PAL
CANW
ABK
GEO
OSS
STA-LBK
PAS
ALT
RUS
GRE
UKR
UDM
SWE
FIN
SAA
HUN
CZE
POL
SIC
ROM
SER
MAC
CRO
BOS
COR
ITA
BAS
CAN
FRE
SAR
GB
NTH
IRL
DAN
GER
KOM
MAR
KHA-MAN
NOR
POR
NEN
MAN
CHI
TIB
EWE
MON
AKAZ
ORO
KAZ
BUR

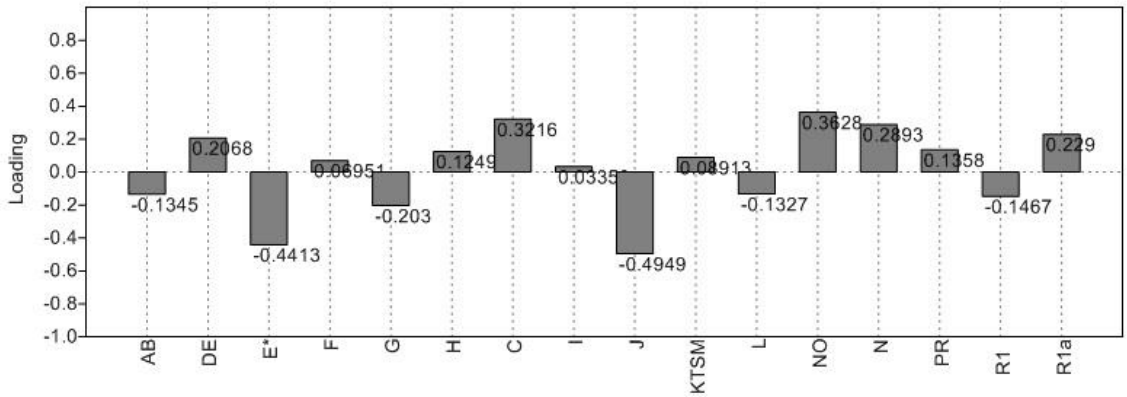
Supplementary Figure 7a-c. Variable correlations with the PC1-3 of the PCA with the SOP-LGY Y-chromosomal dataset and 79 modern populations.

The values are given as coefficient values. The data were plotted in PAST program, based on the original haplogroup frequency table, used for the PCA. The correlations correspond to Figure 36.

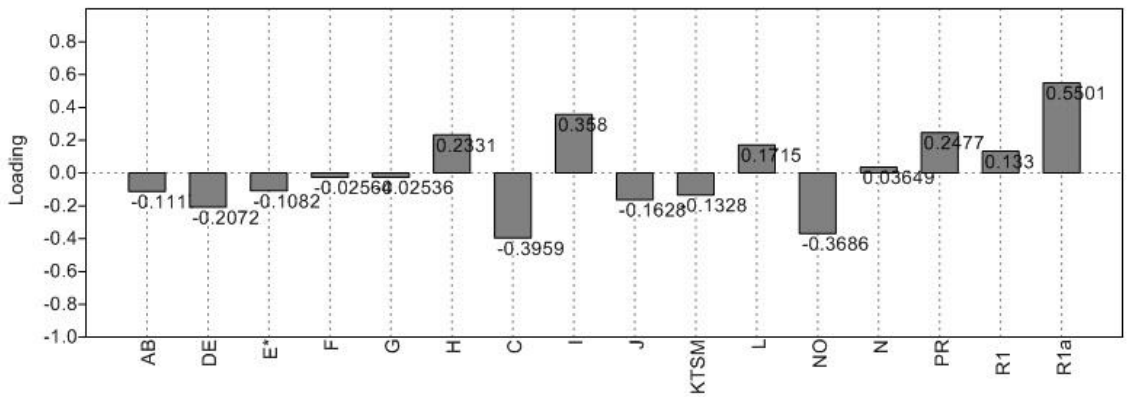
A: PC1



B: PC2

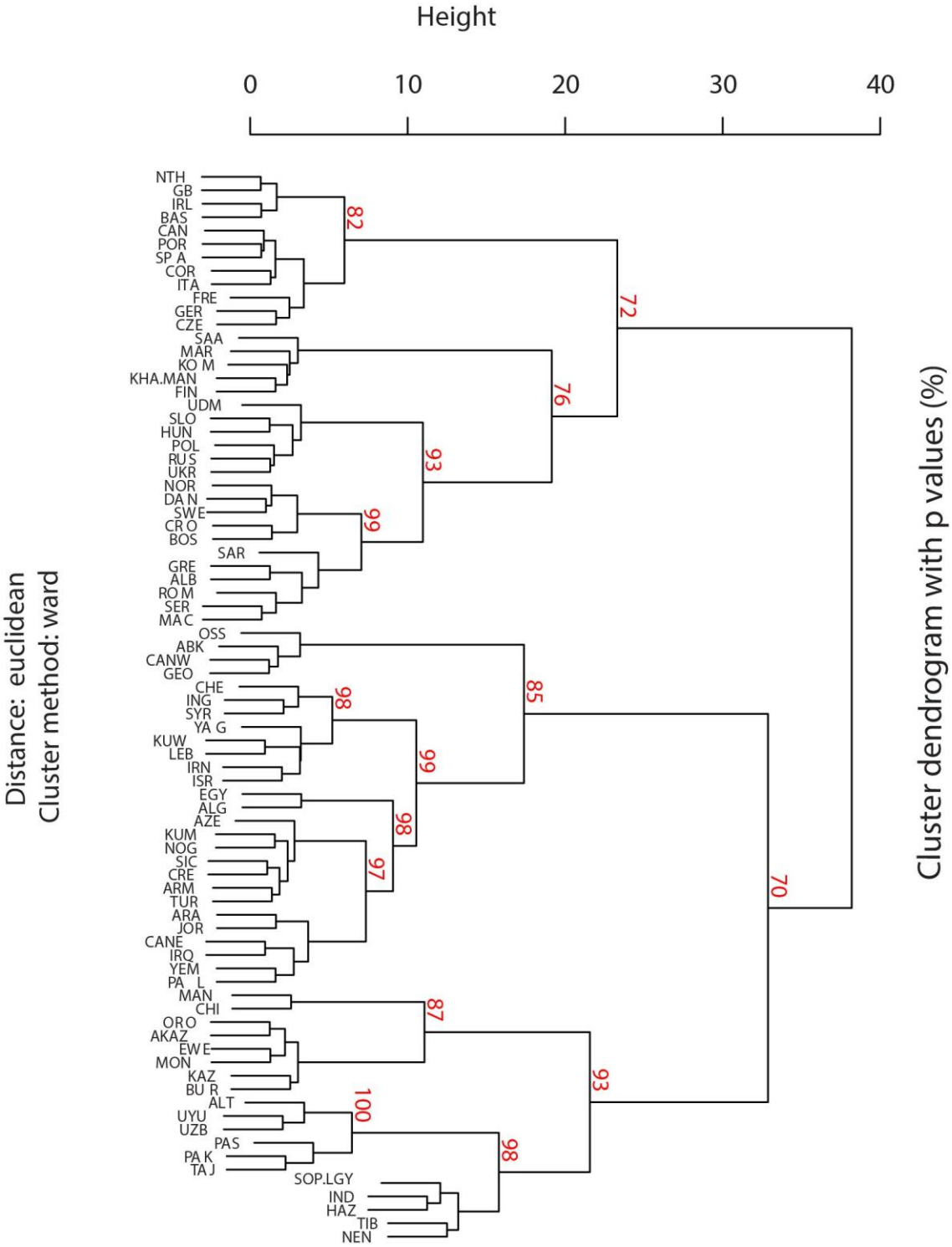


C: PC3

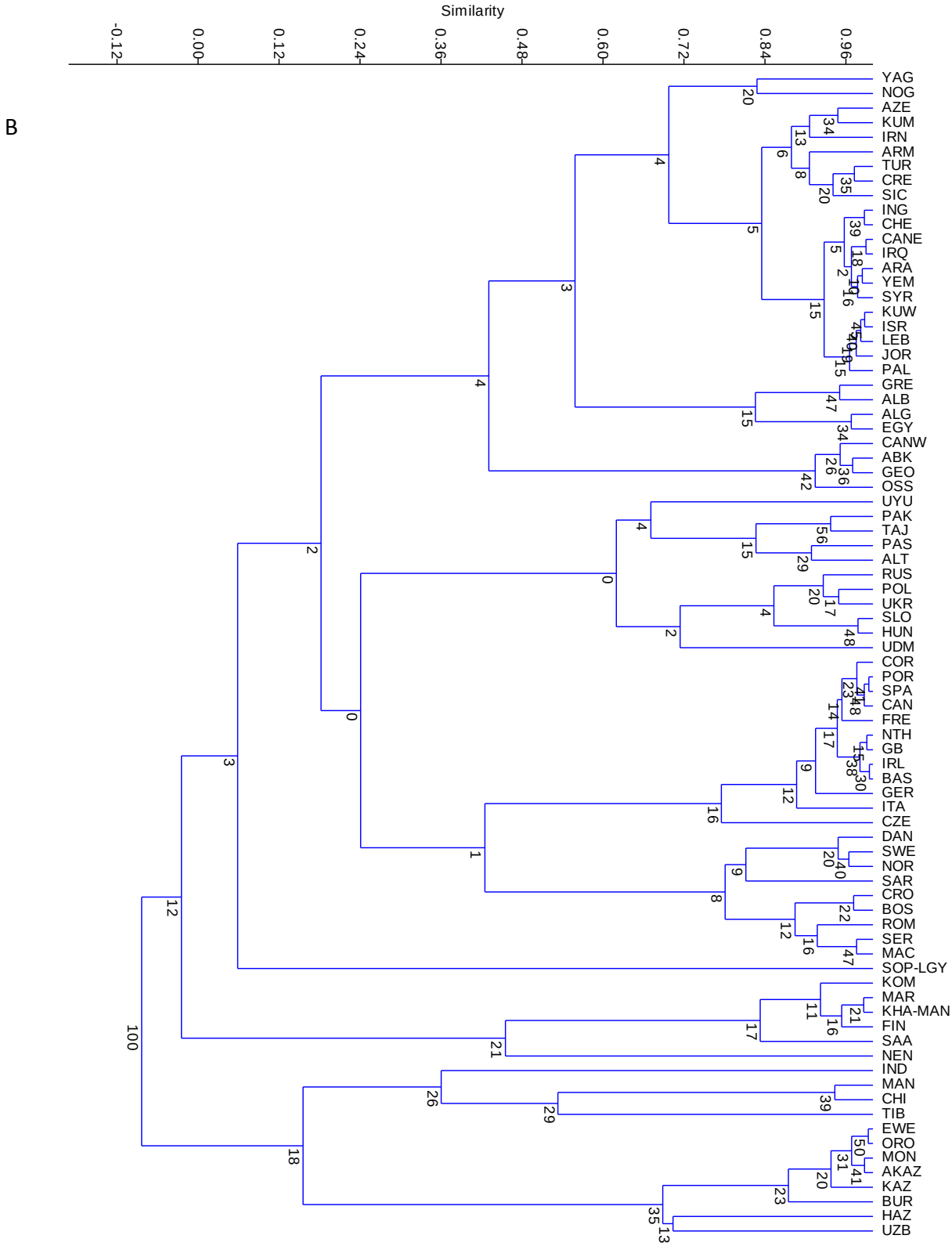


Supplementary Figure 8a-b. Clustering of ward type (A) and paired group (B) with the SOP-LGY Y-chromosomal dataset and 80 modern populations.

A



B



The diagram A was calculated with Euclidean distance measurement, and clustered by Ward type method. At the B dendogram, the “cophenetic coefficient” is 0.847, the algorithm was “paired group”, and the similarity measurement was done by the correlation method.

15.3 Supplementary Tables

15.3.1 List of the Supplementary Tables

Supplementary Table 1. Individual radiocarbon data of the studied skeletons.

The radiocarbon dates were calibrated using OxCal 4.2.3 software, IntCal 13 calibration curve. (Bronk Ramsey 2013; Reimer et al. 2013). All available (published and unpublished) ^{14}C dates were listed in the table.

Site	Labor Nr.	Obj.	SNR	Culture	Basis of dating	Labor Nr. ^{14}C	^{14}C age (BP)	Cal 1 sigma	Cal 2 sigma	References of radiocarbon dates
Vela Spila/Island Korčula	**STANKO**			Mesolithic	stratigraphy	VERA 2340	Layer 7/4: 7200±30 BP	6075-6020	6205-6000	Komso 2006
Alsónyék-Bátaszék, Mérnöki telep	**BAM01**	0688.		Starčevo	^{14}C date	MAMS-11926	6649±29	5630-5550	5640-5520	Szécsényi-Nagy et al. 2014a
	BAM02	0721.		Starčevo	^{14}C date	MAMS-11927	6852±31	5770-5670	5810-5660	Szécsényi-Nagy et al. 2014a
	BAM04	0745.		Starčevo	^{14}C date	MAMS-11928	6677±27	5640-5560	5650-5540	Szécsényi-Nagy et al. 2014a
	BAM05	0746.		Starčevo	^{14}C date	MAMS-11929	6571±34	5550-5480	5620-5470	Szécsényi-Nagy et al. 2014a
	BAM06	0775.		Starčevo	^{14}C date	MAMS-11930	6672±35	5640-5560	5650-5520	Szécsényi-Nagy et al. 2014a
	BAM08	0797.		Starčevo	^{14}C date	MAMS-11931	6657±30	5630-5560	5640-5530	Szécsényi-Nagy et al. 2014a
	BAM11	1372.		Starčevo	^{14}C date	MAMS-11932	6661±25	5630-5560	5640-5540	Szécsényi-Nagy et al. 2014a
	BAM13	1435.		Starčevo	^{14}C date	MAMS-11933	6704±34	5660-5570	5710-5550	Szécsényi-Nagy et al. 2014a
	BAM14	1436./female skeleton		Starčevo	^{14}C date	MAMS-11934	6800±35	5720-5660	5740-5630	Szécsényi-Nagy et al. 2014a
	BAM17	1483.		Starčevo	^{14}C date	MAMS-11935	6857±31	5780-5700	5840-5660	Szécsényi-Nagy et al. 2014a
	BAM21	1525.		Starčevo	^{14}C date	MAMS-11936	6698±34	5650-5560	5680-5540	Szécsényi-Nagy et al. 2014a
	BAM22	1527.		Starčevo	^{14}C date	MAMS-11937	6709±34	5670-5570	5710-5550	Szécsényi-Nagy et al. 2014a
	BAM23	1528.		Starčevo	^{14}C date	MAMS-11938	6617±38	5620-5520	5630-5490	Szécsényi-Nagy et al. 2014a
	BAM25	1532.		Starčevo	^{14}C date	MAMS-11939	6695±40	5650-5560	5710-5530	Szécsényi-Nagy et al. 2014a
	BAM26	1533.		Starčevo	^{14}C date	MAMS-11940	6853±38	5780-5670	5840-5660	Szécsényi-Nagy et al. 2014a
	BAM27	1535.		Lengyel	^{14}C date	MAMS-11941	5790±33	4710-4600	4720-4540	Szécsényi-Nagy et al. 2014a
Lánycsók-Csata-alja	**M6-116.2**	190.		Bronze Age	^{14}C date	MAMS-14127	4145±23	2870-2660	2880-2630	Szécsényi-Nagy et al. 2014a
	M6-116.3	312.		Bronze Age	^{14}C date	MAMS-14128	4077±23	2840-2570	2850-2490	Szécsényi-Nagy et al. 2014a
	M6-116.8	281.		Bronze Age	^{14}C date	MAMS-14129	4120±22	2860-2620	2870-2580	Szécsényi-Nagy et al. 2014a
	M6-116.9	360.		Starčevo	^{14}C date	MAMS-14130	6712±25	5660-5610	5680-5560	Szécsényi-Nagy et al. 2014a
	M6-116.10	369.		Bronze Age	^{14}C date	MAMS-14131	4113±23	2860-2620	2870-2570	Szécsényi-Nagy et al. 2014a
	M6-116.12	221.		Balaton-Lasinja	^{14}C date	MAMS-14132	5300±23	4230-4050	4240-4040	Szécsényi-Nagy et al. 2014a
Balatonszárszó-Kis-erdei-dűlő	**BSZ 3**	278		LBK	^{14}C date	MAMS-14139	6105±25	5060-4980	5210-4940	Szécsényi-Nagy et al. 2014a
Budakeszi 4/8	**BUD 2**	199.		LBK	^{14}C date	MAMS-14140	6171±25	5210-5060	5220-5040	Szécsényi-Nagy et al. 2014a

		(181.)							
		201.							
	BUD 3	(181.)		LBK	^{14}C date	MAMS-14141	6170±25	5210-5060	5220-5040	Szécsényi-Nagy et al. 2014a
	BUD 13	503.		LBK	^{14}C date	MAMS-14142	6066±28	5020-4930	5060-4850	Szécsényi-Nagy et al. 2014a
	BUD 14	629. (509.)		LBK	^{14}C date	MAMS-14143	6152±26	5210-5040	5220-5010	Szécsényi-Nagy et al. 2014a
M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	KON 1	55.	56	LBK	arch. context					
	KON 2	223.	233	Balaton-Lasinja	^{14}C date	Beta - 310033	5380±30	4330-4170	4340-4070	Szécsényi-Nagy et al. 2014a
	KON 5	612.	647, 002	LBK	^{14}C date	Beta - 310035	6040±40	5000-4850	5050-4830	Szécsényi-Nagy et al. 2014a
	KON 6	748.	785	Bronze Age	^{14}C date	Beta - 310036	3530±30	1920-1770	1950-1760	Szécsényi-Nagy et al. 2014a
Tolna-Mözs	TOLM3	2392.	2559.	LBK	^{14}C date	MAMS-14144	6143±24	5210-5020	5210-5000	Szécsényi-Nagy et al. 2014a
	TOLM4	1649.	1748.	LBK	^{14}C date	MAMS-14145	6233±23	5300-5200	5310-5070	Szécsényi-Nagy et al. 2014a
Nitra	NITR2	2/64		LBK	arch. context	OxA-24095	6298±33		5330-5210	Bickle&Whittle 2013
	NITR14	32/65		LBK	arch. context	OxA24578	6138±34		5220-4980	Bickle&Whittle 2013
VerSend-Gilencsa	VEGI1	415.		Vinča	^{14}C date	MAMS-14830	6321±28	5330-5220	5360-5220	this study
	VEGI12	1163.		Vinča	^{14}C date	MAMS-14831	6202±31	5220-5070	5290-5050	this study
	VEGI14	1394.		Vinča	^{14}C date	MAMS-14832	6226±30	5300-5080	5310-5060	this study
	VEGI22	2030.		Vinča	^{14}C date	MAMS-14833	6186±29	5220-5070	5230-5040	this study
Szederkény-Kukorica-dűlő	SEKU1	119.		Vinča	arch. context	MAMS-14808	6079±33	5040-4940	5210-4850	this study
	SEKU6	2398.		Vinča	arch. context	MAMS-14809	6267±33	5300-5220	5330-5080	this study
	SEKU7	3413.		Vinča	arch. context	MAMS-14810	6224±29	5300-5070	5300-5060	this study
	SEKU8	2436.		Vinča	arch. context	MAMS-14811	6362±33	5380-5300	5470-5230	this study
	SEKU11	2842.		Vinča	arch. context	MAMS-14812	6220±29	5290-5070	5300-5060	this study
Szemely-Hegyes	SZEH2	627		Sopot	^{14}C date	Beta - 310034	6020±40	4960-4840	5020-4790	this study
	SZEH3	883		Sopot	^{14}C date	Beta - 310037	5990±40	4940-4800	5000-4780	this study
	SZEH 4	1001		LBK	^{14}C date	Beta - 310038	6110±30	5200-4980	5210-4940	this study
	SZEH5	1003		Sopot	arch. context	Beta - 310039	5920±40	4840-4720	4910-4700	this study
	SZEH7	1085		Sopot	^{14}C date	Beta - 310040	5930±40	4850-4720	4930-4710	this study
	SZEH 9	1139		LBK	^{14}C date	Beta - 310041	6140±30	5210-5000	5210-5000	this study
Alsónyék-elkerülő 2. lh.	ALE1	210.		Sopot	^{14}C date	MAMS-14813	6008±32	4940-4840	5000-4800	this study
	ALE4	220A		Sopot	^{14}C date	MAMS-14814	6032±32	4990-4850	5020-4830	this study
	ALE11	396.		Sopot	^{14}C date	MAMS-14815	5989±32	4940-4830	4970-4790	this study
	ALE13	449.		Bronze Age	^{14}C date	MAMS-14816	3682±28	2140-2020	2200-1970	this study
	ALE14	463.		Sopot	^{14}C date	MAMS-14817	6049±29	5000-4910	5030-4840	this study

	ALE17	471.		Sopot	^{14}C date	MAMS-14818	5937±32	4850-4740	4910-4720	this study
Veszprém-Jutasi út	**VEJ2**	98.	2.	Lengyel	arch. context, 14C date	MAMS-14826	5610±33	4490-4370	4510-4350	this study
	VEJ4	71.	4.	Lengyel	arch. context, 14C date	MAMS-14827	5861±26	4780-4700	4800-4680	this study
	VEJ9	280.	9.	Balaton-Lasinja	arch. context, 14C date	MAMS-14828	5418±29	4330-4260	4340-4230	this study
	VEJ10	555.	**13.**	Balaton-Lasinja	arch. context, 14C date	MAMS-14829	5213±31	4050-3970	4220-3960	this study

Supplementary Table 2. Names and types of the studied samples.

Tooth notation followed the FDI World Dental Federation if determinable otherwise, the abbreviations M (molar), P (premolar), C (caninus), and I (incisivus) was used. Numerous possible teeth are separated by slashes. Bone samples are named by anatomical notation followed by r. (right) or l. (left), if determinable.

	Site	**Country**	**Nr. im Labor**	**Feature**	**Grave nr.**	**Sample A**	**Sample B**	**Sample C**	**Sample D**
Mesolithic	Vela Spilja	Croatia	**STANKO**	Stratum 12 /2004		M37	M28		
Starčevo culture (STA)	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 1**	0688.		M17	M16		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 2**	0721.		M37	M18		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 3**	0727.		femur r.	femur l.		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 4**	0745.		M36	M37		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 5**	0746.		M36	I42		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 6**	0775.		M27	M28		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 7**	0792.		M46	M85		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 8**	0797.		M36	M37		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 9**	1061.		femur l.	femur r.		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 10**	1362.		M46/47	M47/48		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 11**	1372.		M47	M48		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 12**	1398.		pars petrosa	femur r.		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 13**	1435.		M16	pars petrosa		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 14**	1436. female		femur l.	femur r.		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 15**	1449.		M47	M46		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 16**	1461.		femur r.	femur l.		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 17**	1483.		M36	pars petrosa		
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 18**	1495.		M36	M37		

	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 19**	1513.		M16	M46	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 20**	1516.		M46	M47	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 21**	1525.		femur r.	femur l.	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 22**	1527.		femur r.	femur l.	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 23**	1528.		M44	M34	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 24**	1531.		M46	M47	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 25**	1532.		M46	M47	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM 26**	1533.		M46	M47	
	Lánycsók Gata Csotola	Hungary	**LGCS1**	1661.		M17	M27	
	Lánycsók Gata Csotola	Hungary	**LGCS2**	1662.		femur l.	femur r.	
	Lánycsók Gata Csotola	Hungary	**LGCS3**	1784.		m55	m65	
	Lánycsók Gata Csotola	Hungary	**LGCS4**	1810. (=1658)		M46	M47	
	Lánycsók, Csata-alja	Hungary	**M6_116.1**	163/160		M26	M27	
	Lánycsók, Csata-alja	Hungary	**M6_116.4**	353/329		M38	PM25	
	Lánycsók, Csata-alja	Hungary	**M6_116.9**	360/384		M27	M36	
	Vinkovci Nama	Croatia	**VINK 1**		7.	M26	M27	
	Vinkovci Nama	Croatia	**VINK 2**		8.	M16	M17	
	Vinkovci Nama	Croatia	**VINK 3**		11.	M18	M38	
	Vinkovci Nama	Croatia	**VINK 4**		12a	femur r.	femur l.	
	Vinkovci Nama	Croatia	**VINK 5**		13.	M17	C23	
	Vinkovci Nama	Croatia	**VINK 6**		15.	M16	M17	
	Vinkovci Jugobanka	Croatia	**VINJ 1**		3.	M16	M36	
	Vinkovci Jugobanka	Croatia	**VINJ 2**		4a	M26	PM24	
	Vinkovci Jugobanka	Croatia	**VINJ 3**		7a	Ulna r.	right Humerus	
	Vinkovci Jugobanka	Croatia	**VINJ 4**		7b	C13	I12	
	Vukovar Gimnazija	Croatia	**VUKG 1**		1.	M26	M38	
	Vukovar Gimnazija	Croatia	**VUKG 2**		2.	femur l.	Ulna l.	
	Vukovar Gimnazija	Croatia	**VUKG 3**		3.	M46	M26	
	Vukovar Gimnazija	Croatia	**VUKG 4**		4.	M36	M17	
Linearbandkeramik in Transdanubia (LBKT)	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 1**	S-36		femur r.	tibia l.	
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 2**	159		femur l.	tibia r.	
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 3**	278		femur r.	tibia r.	
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 4**	289		femur	tibia	I11/21 M36/46
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 5**	510		Humerus l.	Ulna r.	M37/47 M17/27
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 6**	531		femur l.	tibia l.	M37 M17
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 7**	554		femur l.	tibia l.	
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 8**	770		femur r.	tibia l.	M18
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 9**	771		femur r.	tibia l.	
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 10**	773		femur l.	tibia	M37/47
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 11**	774		femur l.	tibia r.	M37 M27
	Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 12**	775		femur l.	tibia r.	

Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 13**	777		femur r.	tibia l.			
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 14**	778		femur r.	tibia l.			
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 15**	780		femur r.	humerus r.	M47		
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 19**	288		M54/64	I62	M74		
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 20**	779		M16/26	humerus	I51	I61	
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 21**	782		P35/45	M17/27	M18/28		
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 22**	791		M17	M27	M17		
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 23**	792		r. femur	l. femur			
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 24**	793		M37/47	M38/48	M16/26		
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 25**	796		M2 maxilla	M2 maxilla			
Balatonszárszó-Kis-erdei-dűlő	Hungary	**BSZ 26**	797		M36	P15	C23	M26	
Balatonszemes-Bagódomb	Hungary	**BAB 1**	57.	69.	M26	M85	M75		
Balatonszemes-Bagódomb	Hungary	**BAB 2**	74.	89.	femur	tibia	M3		
Balatonszemes-Bagódomb	Hungary	**BAB 3**	314.	305.	M16/17	M38			
Balatonszemes-Bagódomb	Hungary	**BAB 4**	324.	504.	M16	M36/46			
Balatonszemes-Bagódomb	Hungary	**BAB 5**	410.	604.	M38	PM24			
Balatonszemes-Bagódomb	Hungary	**BAB 6**	420.	623.	M16	M46	I21		
Bölcske-Gyűrűsvölgy	Hungary	**BÖVÖ 1**	41.	55.	M16	PM			
Bölcske-Gyűrűsvölgy	Hungary	**BÖVÖ 2**	62.	82.	M26	M38			
Bölcske-Gyűrűsvölgy	Hungary	**BÖVÖ 3**	63.	83.	M36	M3 Mandibula			
Bölcske-Gyűrűsvölgy	Hungary	**BÖVÖ 4**	116.	161.	M46	M47			
Bölcske-Gyűrűsvölgy	Hungary	**BÖVÖ 5**	117.	162.	M36	M46			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 1**	87.		M36	M28			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 2**	199. (181.)		M36	M38			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 3**	201. (181.)		M26	M28			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 4**	290.		M26	M28			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 5**	297.		r. femur	l. femur			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 6**	378.		r. femur	l. humerus			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 7**	388.		M26	M27			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 9**	389.2		M36	M37			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 10**	348.		M36	m75			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 11**	557.		r. femur	l. femur			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 12**	631. (611.)		M36	M37			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 13**	503.		M36	M48			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 14**	629. (509.)		M16	M48			
Budakeszi, Szőlőskert-Tangazdaság	Hungary	**BUD 15**	632. (611.)		M16	M26			
Harta-Gátőrház	Hungary	**HARG 1**	115.		femur	humerus	PM		
Harta-Gátőrház	Hungary	**HARG 2**	132.		M37	M26	ulna		

	Harta-Gátőrház	Hungary	**HARG 3**	137.		r. femur.	humerus	M36/46	M28
	Harta-Gátőrház	Hungary	**HARG 4**	161.		M46/36	M37/47		
	Harta-Gátőrház	Hungary	**HARG 5**	185.		M48	M36/46		
	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 1**	55.		l. femur	r. humerus.		
	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 3**	286.		M18	M16		
	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 4**	612.1		M85	M75		
	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 5**	612.2		M85	M84		
	Szemely-Hegyes	Hungary	**SZEH 4**	1001.	607-608.	M36	M27		
	Szemely-Hegyes	Hungary	**SZEH 9**	1139.	1025.	M16	M36		
	Tolna-Mözs	Hungary	**TOLM 3**	2392.		r. femur	l. femur		
	Tolna-Mözs	Hungary	**TOLM 4**	1649.		M36	M37		
	Nitra	Slovakia	**NITR2**	2/64		M18	PM45		
	Nitra	Slovakia	**NITR7**	14/64		PM45	M46		
	Nitra	Slovakia	**NITR11**	20/65		M27	M16		
	Nitra	Slovakia	**NITR14**	32/65		C23	I22		
	Nitra	Slovakia	**NITR26**	68/65		M16	M26		
Vinča culture	Versend-Gilencsa	Hungary	**VEGI1**	415.		M17	C43	C33	
	Versend-Gilencsa	Hungary	**VEGI2**	1032.		M47	M46		
	Versend-Gilencsa	Hungary	**VEGI3**	1039.		M46	PM45		
	Versend-Gilencsa	Hungary	**VEGI4**	1049.		PM45	C33		
	Versend-Gilencsa	Hungary	**VEGI5**	1078.		M36	M38		
	Versend-Gilencsa	Hungary	**VEGI6**	1098.		M38	l. pars petrosa		
	Versend-Gilencsa	Hungary	**VEGI7**	1110.		M26	m65		
	Versend-Gilencsa	Hungary	**VEGI8**	1115.		PM34	C33		
	Versend-Gilencsa	Hungary	**VEGI9**	1121.		M38	M36		
	Versend-Gilencsa	Hungary	**VEGI10**	1124.		M36	M48		
	Versend-Gilencsa	Hungary	**VEGI11**	1162.		PM24	M26		
	Versend-Gilencsa	Hungary	**VEGI12**	1163.		m75	m74		
	Versend-Gilencsa	Hungary	**VEGI13**	1290.		M36	M37		
	Versend-Gilencsa	Hungary	**VEGI14**	1394.		femur	l. pars petrosa		
	Versend-Gilencsa	Hungary	**VEGI15**	1505.		r. femur	l. femur		
	Versend-Gilencsa	Hungary	**VEGI16**	1506.		r. femur	tibia		
	Versend-Gilencsa	Hungary	**VEGI17**	1561.		M48	M46		
	Versend-Gilencsa	Hungary	**VEGI18**	1703.		m75	M36		
	Versend-Gilencsa	Hungary	**VEGI19**	1720.		M37	M36		
	Versend-Gilencsa	Hungary	**VEGI20**	1972.		M37	M36		
	Versend-Gilencsa	Hungary	**VEGI21**	1995.		M37	M36		
	Versend-Gilencsa	Hungary	**VEGI22**	2030.		M48	M46		
	Versend-Gilencsa	Hungary	**VEGI23**	2032.		M36	M46		
	Versend-Gilencsa	Hungary	**VEGI24**	1568.		r. femur	r. tibia		
	Versend-Gilencsa	Hungary	**VEGI25**	1721.		M46	M47		
	Szederkény-Kukorica-dűlő	Hungary	**SEKU1**	119.		M46	M47		
	Szederkény-Kukorica-dűlő	Hungary	**SEKU2**	159.		l. tibia	l. pars petrosa		

	Szederkény-Kukorica-dűlő	Hungary	**SEKU3**	270.	M36	M37	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU4**	344.	PM25	femur	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU5**	367.	M16	M17	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU6**	2398.	M26	M27	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU7**	3413.	M36	M37	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU8**	2436.	M16	M26	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU9**	2484.	humerus	ulna	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU10**	2491.	M16	M17	
	Szederkény-Kukorica-dűlő	Hungary	**SEKU11**	2842.	M26	M27	
	Szemely-Hegyes	Hungary	**SZEH1**	549	femur	tibia	
	Szemely-Hegyes	Hungary	**SZEH2**	627	M17	M16	
	Szemely-Hegyes	Hungary	**SZEH3**	883	M48	M36	
	Szemely-Hegyes	Hungary	**SZEH5**	1003	M26	M27	
	Szemely-Hegyes	Hungary	**SZEH6**	1070	M16	m55	
	Szemely-Hegyes	Hungary	**SZEH7**	1085	r. pars petrosa	l. pars petrosa	
	Szemely-Hegyes	Hungary	**SZEH8**	1108	M48	M47	
	Nemesvámos-Kapsa utca	Hungary	**NEK1**	17.	M26	M28	
	Nemesvámos-Baláca	Hungary	**NEB1**	1/L	M36	M37	
	Nemesvámos-Baláca	Hungary	**NEB2**	1/i	r. femur	r. humerus	
	Radovanci	Croatia	**RADOV1**		M36	M48 lose	
	Bicske-Galagonyás	Hungary	**BICS1**	1.	M37	M38	
	Bicske-Galagonyás	Hungary	**BICS2**	2.	PM45	PM44	
	Bicske-Galagonyás	Hungary	**BICS3**	3.	M46	M18	
	Bicske-Galagonyás	Hungary	**BICS4**	4.	M47	PM45	
	Bicske-Galagonyás	Hungary	**BICS5**	1959/1	M16	PM15	
	Fajsz-Garadomb	Hungary	**FAGA1**	165.	M36	M38	
	Fajsz-Garadomb	Hungary	**FAGA2**	156.	M46	M48	
	Fajsz-Garadomb	Hungary	**FAGA3**	65.	M26	M28	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE1**	210.	r. femur	l. humerus	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE2**	214B	M27	M28	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE3**	220B	M36	M37	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE4**	220A	M36	M37	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE5**	240.	M46	M47	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE6**	272.	M75	M85	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE7**	282.	M46	M37	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE8**	283.	M18	M17	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE9**	372.	M46?	PM45	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE10**	373.	r. femur	l. tibia	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE11**	396.	M16	M55	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE12**	432.	M36	M37	M38
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE14**	463.	M85	M65	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE15**	464.	M36	M38	M37
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE16**	470.	M47	M38	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE17**	471.	M36	M37	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE18**	475.	M26	M46	
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE19**	476.	M37	M36	

Sopot culture

Lengyel culture	Veszprém-Felszabadulási út	Hungary	**VEM1.**	77.1.1		r. tibia	r. femur	
	Veszprém-Felszabadulási út	Hungary	**VEM2.**	77.1.2		l. femur	r. femur	
	Veszprém-Felszabadulási út	Hungary	**VEM3.**	77.3.1		r. femur	r. humerus	
	Veszprém-Jutasi út	Hungary	**VEJ1**	102.	3.	M38	M37	
	Veszprém-Jutasi út	Hungary	**VEJ2**	98.	2.	M38.	l. femur	
	Veszprém-Jutasi út	Hungary	**VEJ3**	100.	1.	M75	M74	
	Veszprém-Jutasi út	Hungary	**VEJ4**	71.	4.	M36	M37	
	Veszprém-Jutasi út	Hungary	**VEJ5**	219.	5.	M38 lose	M37	
	Veszprém-Jutasi út	Hungary	**VEJ6**	228.	6.	M37	M38	
	Veszprém-Jutasi út	Hungary	**VEJ7**	229.	7.	M38	M37	
	Veszprém-Jutasi út	Hungary	**VEJ8**	158.	8.	r. tibia	l. tibia	
	Felsőörs-Bárókert	Hungary	**FEB1**	38.	1.	M26	M27	
	Felsőörs-Bárókert	Hungary	**FEB2**	38.	2.	M85	M75	
	Felsőörs-Bárókert	Hungary	**FEB3**	100.		PM44	M46	
	Felsőörs-Bárókert	Hungary	**FEB4**	133.		M36	M75	
	Felsőörs-Bárókert	Hungary	**FEB5**	42.		r. femur	l. femur	
	Bátaszék-Lajvér	Hungary	**BAL 1**	26.		M36	PM35	
	Bátaszék-Lajvér	Hungary	**BAL 2**	27.		M47	I42	M46
	Bátaszék-Lajvér	Hungary	**BAL 3**	35.		M46	M47	
	Bátaszék-Lajvér	Hungary	**BAL 4**	36.		M46	M47	
	Bátaszék-Lajvér	Hungary	**BAL 5**	38.		M46	M47	M48
	Bátaszék-Lajvér	Hungary	**BAL 6**	39.		M46	M47	
	Bátaszék-Lajvér	Hungary	**BAL 7**	40.		M36	M38	
	Bátaszék-Lajvér	Hungary	**BAL 8**	49.		M85	M84	
	Bátaszék-Lajvér	Hungary	**BAL 9**	65.		r. femur	M36	
	Bátaszék-Lajvér	Hungary	**BAL10**	66.		M36	M35	
	Bátaszék-Lajvér	Hungary	**BAL 11**	68.		M36	M37	
	Bátaszék-Lajvér	Hungary	**BAL 12**	69.		M46	M47	
	Bátaszék-Lajvér	Hungary	**BAL 13**	71.		r. femur	l. tibia	C
	Bátaszék-Lajvér	Hungary	**BAL 14**	72.		M46	M48	
	Bátaszék-Lajvér	Hungary	**BAL 15**	76.		M46	M47	
	Bátaszék-Lajvér	Hungary	**BAL 16**	89.		M46	M48	
	Bátaszék-Lajvér	Hungary	**BAL 17**	18.		l. humerus	l. femur	
	Bátaszék-Lajvér	Hungary	**BAL 18**	34.		l. femur	M47/48?	
	Bátaszék-Lajvér	Hungary	**BAL19**	41.		M46?	l.femur	M36?
	Bátaszék-Lajvér	Hungary	**BAL 21**	50.		M16	M18	
	Bátaszék-Lajvér	Hungary	**BAL22**	51.		M26	M27	
	Bátaszék-Lajvér	Hungary	**BAL24**	70.		M26	M28	
	Bátaszék-Lajvér	Hungary	**BAL25**	93.		r. femur	r. pars petrosa	
	Bátaszék-Lajvér	Hungary	**BAL26**	94.		l. femur	r. femur	
	Bátaszék-Lajvér	Hungary	**BAL27**	97.		r. femur	l. femur	
	Alsónyék-Bátaszék, Mérnöki telep	Hungary	**BAM27**	1535.		Pars petrosa 2x	r. femur	
	Borjád-Kenderföldek	Hungary	**BORK1**	3.Obj./ 2010.11. 05		M46	M48	
	Csabdi-Télizöldes	Hungary	**CSAT1**	1.		M16	l. femur	
	Csabdi-Télizöldes	Hungary	**CSAT2**	2.		M17	r. femur	
	Csabdi-Télizöldes	Hungary	**CSAT3**	4.		M26	M27	

	Csabdi-Télizöldes	Hungary	**CSAT5**	7.		M36	M38
	Csabdi-Télizöldes	Hungary	**CSAT6**	9.		M46	M47
	Csabdi-Télizöldes	Hungary	**CSAT7**	10.		M36	M37
	Csabdi-Télizöldes	Hungary	**CSAT8**	11.		l. femur	l. tibia
	Csabdi-Télizöldes	Hungary	**CSAT9**	12. Grube D		r. femur	tibia
	Csabdi-Télizöldes	Hungary	**CSAT11**	13.2		M16?	M17?
	Csabdi-Télizöldes	Hungary	**CSAT12**	14. 1		M46	m85
	Csabdi-Télizöldes	Hungary	**CSAT14**	15. group 1.		r. femur	l. femur
	Csabdi-Télizöldes	Hungary	**CSAT15**	15. group 2.		M46	M47
	Csabdi-Télizöldes	Hungary	**CSAT16**	15. group 3.		M36	M37
	Csabdi-Télizöldes	Hungary	**CSAT17**	16.		M26	M28
	Csabdi-Télizöldes	Hungary	**CSAT18**	17.		M16	M48
	Csabdi-Télizöldes	Hungary	**CSAT19**	18.		M46	M48
	Csabdi-Télizöldes	Hungary	**CSAT20**	19.		M46	M47
	Csabdi-Télizöldes	Hungary	**CSAT21**	20.		M36	m75
	Csabdi-Télizöldes	Hungary	**CSAT22**	21.		M36	M46
	Csabdi-Télizöldes	Hungary	**CSAT23**	23.		r. femur	l. femur
	Csabdi-Télizöldes	Hungary	**CSAT24**	24.		M26	M28
	Csabdi-Télizöldes	Hungary	**CSAT25**	25.		l. Pars petrosa	r. pars petrosa
	Csabdi-Télizöldes	Hungary	**CSAT26**	26.		M46	M17
	Csabdi-Télizöldes	Hungary	**CSAT27**	27.		M26	M27
	Csabdi-Télizöldes	Hungary	**CSAT29**	30.		PM24	PM25
	Csabdi-Télizöldes	Hungary	**CSAT30**	31.		M27	M28
	Csabdi-Télizöldes	Hungary	**CSAT31**	32.		M26	m65
	Mórágy-Tűzkődomb, B1	Hungary	**MORT1**	15.		femur?	C13
	Mórágy-Tűzkődomb, B1	Hungary	**MORT2**	16.		M27	M47
	Mórágy-Tűzkődomb, B1	Hungary	**MORT3**	36.		l. femur	r. femur
	Mórágy-Tűzkődomb, B1	Hungary	**MORT4**	43.		M26	m75
	Mórágy-Tűzkődomb, B1	Hungary	**MORT5**	44.		r. femur	l. femur
	Mórágy-Tűzkődomb, B1	Hungary	**MORT6**	46.		m75	m74
	Mórágy-Tűzkődomb, B1	Hungary	**MORT7**	47.		M46	M47
	Mórágy-Tűzkődomb, B1	Hungary	**MORT8**	48.		M36	M37
	Mórágy-Tűzkődomb, B1	Hungary	**MORT9**	51.		m65	m64
	Mórágy-Tűzkődomb, B1	Hungary	**MORT11**	55.		M46	l21
	Mórágy-Tűzkődomb, B1	Hungary	**MORT12**	56.		M46	M47
	Mórágy-Tűzkődomb, B1	Hungary	**MORT13**	57.		M37	M38
	Mórágy-Tűzkődomb, B1	Hungary	**MORT15**	59.		M18	M36
	Mórágy-Tűzkődomb, B1	Hungary	**MORT16**	60.		m55	m54
	Mórágy-Tűzkődomb, B1	Hungary	**MORT18**	63.		m75	M36
	Mórágy-Tűzkődomb, B1	Hungary	**MORT19**	65.		M36	M37
	Mórágy-Tűzkődomb, B1	Hungary	**MORT21**	67.		M36	M37
	Mórágy-Tűzkődomb, B1	Hungary	**MORT25**	81.		M36	M46
	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 7**	826.			
Balaton-Lasinja culture	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 2**	223.			
	Veszprém-Jutasi út	Hungary	**VEJ9**	280.	9.	M46	M85
	Veszprém-Jutasi út	Hungary	**VEJ10**	555.	**13.**	M36	M38

	Veszprém-Jutasi út	Hungary	**VEJ11**	556.	**14.**	M46-lose	M48-lose
Balaton-Lasinja culture	Veszprém-Jutasi út	Hungary	**VEJ12**	562.	**15.**	M16	M18
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP1**	45.	1/1/2000	M65	M64
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP2**	45.	1/2/2000	r. pars petrosa	l. pars petrosa
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP3**	45.	1/3/2000	M46	PM24
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP4**	45.	1/4/2000	M16	M28
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP5**	45.	1/5/2000	M65	M55
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP6**	45.	1/6/2000	M46	PM44
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP7**	45.	1/7/2000	m65	m64
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP13**	45. Streufund	1/8/2000	M36	M38
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP8**	46.		l. femur	
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP9**	46.		l. humerus	
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP10**	46.		r. humerus	
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP11**	46.		l. humerus	
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP12**	46.		r. humerus	
	Keszthely-Fenékpuszta Pusztaszentegyházi dűlő	Hungary	**KEFP14**	46. Streufund		l. femur	
	Tolna-Mözs	Hungary	**TOLM1**	162.		femur	Incisivus Maxilla
	Tolna-Mözs	Hungary	**TOLM2**	306.		M46	M48
	Lánycsók, Csata-alja	Hungary	**M6-116.12**	221.		M37	M36
Bronze Age	Lánycsók, Csata-alja	Hungary	**M6-116.2**	190.		li. femur	re. femur
	Lánycsók, Csata-alja	Hungary	**M6-116.3**	312.		M27	M16
	Lánycsók, Csata-alja	Hungary	**M6-116.8**	281.		M37	M36
	Lánycsók, Csata-alja	Hungary	**M6-116.10**	369.		M16	M18
	M85 Enese elkerülő 02. Kóny, Proletár-dűlő II	Hungary	**KON 6**	748.		M46.	M28.
	Alsónyék-elkerülő 2. lh.	Hungary	**ALE13**	449.		M26	M27

Supplementary Table 3. Mitochondrial DNA (HVS-I and HVS-II) results.

Sequence polymorphisms were identified relative to the revised Cambridge Reference Sequence (rCRS). Haplogroup determination is based on phylotree built 14, accessed 05 Apr 2012 (www.phylotree.com). Generally, all control region sequences are at least reproduced by two independent extracts, otherwise indicated with question mark. Individuals with insufficient DNA preservation and unreliable HVS-I sequences are signed by n.d. (stands for not determined).

Nr. im Labor	HVS-I compared to CRS	range HVS-I	HVS-II	uninformative sites HVS-II	range HVS-II	Haplogroup definition
STANKO	16270t 16311c	16056-16402				U5b2a5
BAM 1	16124c 16126c 16294t 16296t 16304c	16045-16402				T2b
BAM 2	16172c 16224c 16311c	16010-16402				K
BAM 3	16223t 16292t	16010-16405				W
BAM 4	16224c 16311c	16018-16409	73g 263g	315.1c	34-371	K
BAM 5	16126c 16153a 16294t 16296t	16045-16402				T2e
BAM 6	16298c	16057-16402				V
BAM 7	16093c 16224c 16225.1c 16311c	16055-16402	73g 263g	315.1c	59-397	K
BAM 8	16093c 16126c 16189c 16294t 16296t 16304c	16010-16409				T2b
BAM 9	16093c 16189c 16224c 16311c	16018-16406				K
BAM 10	-	16012-16409	263g	315.1c	34-397	H
BAM 11	-	16011-16402	263g	315.1c	35-397	H
BAM 12	16343g	16012-16409				U3
BAM 13	16183C 16189c 16223t 16278t	15997-16402	73g 195c 207a 263g	309.1c 315.1c	62-382	X2
BAM 14	16069t 16126c	16011-16409	73g 185a 228a 263g 295t	315.1c	61-371	J1c
BAM 15	16189c 16223t 16278t 16362c	16055-16402	73g 153g 195c 225a 263g	309.2c 315.1c	61-394	X2
BAM 16	16224c 16311c	16039-16405	73g 263g ?	315.1c	35-397	K
BAM 17	16126c 16163g 16186t 16189c 16294t	16045-16402				T1a
BAM 18	16069t 16126c	16045-16402	73g 263g 295t	315.1c	61-372	J
BAM 19	16093c 16224c 16311c	15997-16402	73g 114t 263g	315.1c	62-371	K1a1
BAM 20	16126c 16294t 16296t 16304c	16011-16402	73g 263g	315.1c	34-397	T2b
BAM 21	16126c 16294t 16304c	16027-16409				T2b
BAM 22	16086c 16147A 16172c 16223t 16248t 16320t 16355t	16028-16409				N1a1a1a
BAM 23	16179t 16189c 16223t 16278t 16362c	16025-16405	73g 153g 195c 225a 263g	315.1c	34-397	X2
BAM 24	16093c 16224c 16311c	15997-16402	73g 150t 263g	315.1c	46-397	K1
BAM 25	16147A 16172c 16193t 16223t 16248t 16355t	16057-16409				N1a1a1
BAM 26	16126c 16294t 16296t 16304c	16057-16409	73g 263g	309.1c 315.1c	34-397	T2b
LGCS1	16093c 16223t 16292t	16003-16409				W
LGCS2	16147A 16172c 16189c 16223t 16248t 16274a 16355t	16056-16402				N1a1a1
LGCS3	16304c	16013-16409				H5
LGCS4	16093c 16189c 16224c 16311c	16003-16409				K
M6_116.1	16356c	16003-16409				U4

M6_116.4	16166g 16224c 16311c	16012-16409				K
M6_116.9	16126c 16292t 16294t	16056-16402				T2c1
VINK 1	16093c 16224c 16311c	16021-16402				K
VINK 2	16069t 16126c 16224c 16261t	16025-16409				J1c
VINK 3	16298c	16056-16409				HV0
VINK 4	n.d.	-				n.d.
VINK 5	16093c 16189c 16224c 16311c	16056-16402				K
VINK 6	n.d.	-				n.d.
VINJ 1	16126c 16294t 16296t 16304c	16018-16409				T2b
VINJ 2	16093c 16224c 16261t 16311c	16018-16402				K1a4a1b2
VINJ 3	16298c	16056-16409				V
VINJ 4	16162g 16298c	16007-16409				V
VUKG 1	16069t 16126c 16302g	16018-16409	185a 228a 263g 295t	315.1c	36-397	J1c
VUKG 2	n.d.					n.d.
VUKG 3	16069t 16126c 16302g	16012-16409	73g 185a 228a 263g 295t	315.1c	36-381	J1c
VUKG 4	16126c 16294t 16296t 16304c	16020-16409				T2b
BSZ 1	n.d.	-				n.d.
BSZ 2	n.d.	-				n.d.
BSZ 3	n.d.	-				n.d.
BSZ 4	n.d.	-				n.d.
BSZ 5	16147A 16154c 16172c 16223t 16248t 16320t 16355t	16056-16409				N1a1a1a3
BSZ 6	n.d.	-				n.d.
BSZ 7	n.d.	-				n.d.
BSZ 8	n.d.	-				n.d.
BSZ 9	16069t 16126c	15997-16409				J
BSZ 10	n.d.	-				n.d.
BSZ 11	n.d.	-				n.d.
BSZ 12	n.d.	-				n.d.
BSZ 13	n.d.	-				n.d.
BSZ 14	n.d.	-				n.d.
BSZ 15	-	15997-16409				H
BSZ 19	16311c	15997-16409				HV
BSZ 20	n.d.	-				n.d.
BSZ 21	16069t 16126c	15997-16409				J
BSZ 22	n.d.	-				n.d.
BSZ 23	n.d.	-				n.d.
BSZ 24	n.d.	-				n.d.
BSZ 25	n.d.	-				n.d.
BSZ 26	n.d.	-				n.d.
BAB 1	n.d.	-				n.d.
BAB 2	n.d.	-				n.d.
BAB 3	-	16013-16402	195c 263g	315.1c	35-389	H
BAB 4	16320t	15997-16402				H26b
BAB 5	16261t	16004-16409				H
BAB 6	-	16012-16409	263g	315.1c	48-396	H
BÖVÖ 1	16093c	16003-16403				H
BÖVÖ 2	16069t 16126c	16011-16409				J
BÖVÖ 3	-	16011-16409				H
BÖVÖ 4	16224c 16311c	16004-16402				K

BÖVÖ 5	16093c 16224c 16311c	16026-16402				K
BUD 1	16192t 16249c 16256t 16270t 16399g	16028-16409				U5a1
BUD 2	16298c	16026-16409				V
BUD 3	16126c 16153a 16294t 16296t	16026-16409				T2e
BUD 4	16126c 16163g 16186t 16189c 16294t	16025-16409				T1a
BUD 5	16126c 16294t 16296t	16026-16409				T2
BUD 6	n.d.	-				n.d.
BUD 7	-	16045-16402	263g	315.1c	35-397	H
BUD 9	16051g 16092c 16179t 16274a	16045-16402				U2
BUD 10	-	16058-16402	263g	315.1c	35-397	H
BUD 11	n.d.	-				n.d.
BUD 12	16126c 16294t 16296t 16304c	16057-16402				T2b
BUD 13	16304c	15999-16409	263g	309.2c 315.1c	35-397	H5
BUD 14	16093c 16224c 16311c	15997-16409				K
BUD 15	16192t 16304c	16045-16402	263g	309.1c 315.1c	36-397	H5
HARG 1	16126c 16294t 16296t 16304c	16045-16402	73g 263g ?	309.1c 315.1c		T2b
HARG 2	16086c 16147A 16172c 16223t 16248t 16320t 16355t	16001-16409				N1a1a1a
HARG 3	16126c 16292t 16294t 16296t	15998-16409				T2c1
HARG 4	16224c 16311c 16398a	16011-16409				K
HARG 5	16126c 16294t 16296t 16304c	16025-16409	73g 263g	315.1c	42-371	T2b
KON 1	-	16045-16409				H
KON 3	16126c 16294t 16296t 16304c	16056-16402				T2b
KON 4	16126c 16147t 16294t 16296t 16297c 16304c	16045-16402	73g 263g	309.2c 315.1c	48-393	T2b23
KON 5	16126c 16147t 16294t 16296t 16297c 16304c	16046-16402	73g 263g	309.1c 315.1c	46-394	T2b23
SZEH 4	16147A 16154c 16172c 16223t 16248t 16300g 16320t 16355t	16018-16409				N1a1a1a3
SZEH 9	16147A 16154c 16172c 16223t 16248t 16320t 16355t	16025-16409				N1a1a1a3
TOLM 3	16126c 16153a 16294t 16296t	16031-16409				T2e
TOLM 4	16224c 16311c	16056-16402				K
NITR2	16093c 16224c 16311c	16048-16409				K
NITR7	16093c 16224c 16311c	16046-16401				K
NITR11	n.d.	-				n.d.
NITR14	16051g 16092c 16179t 16274a	16046-16409				U2
NITR26	n.d.	-				n.d.
VEGI1	16126c 16147t 16294t 16296t 16297c 16304c	16047-16409				T2b23
VEGI2	16069t 16126c	16025-16409	73g 152c 185a 188g 228a 263g 295t	309.1c 315.1c	60-394	J1c2
VEGI3	16126c 16294t 16296t 16304c	15997-16409				T2b
VEGI4	16093c 16189c 16224c 16311c	16046-16409				K
VEGI5	16319a 16343g	16019-16409				U3
VEGI6	16093c 16224c 16311c	15999-16409	73g 263g	309.1c 315.1c	61-394	K1
VEGI7	16093c 16224c 16311c	15997-16409	73g 263g	309.1c 315.1c	35-396	K1
VEGI8	16086c 16147a 16172c 16223t 16248t 16320t 16355t	16026-16409				N1a1a1a
VEGI9	16292t 16343g	16057-16409				U3
VEGI10	16192t 16256t 16270t	16058-16401				U5a
VEGI11	16126c 16147t 16294t 16296t 16297c 16304c	16057-16409	73g 263g	309.1c 315.1c	61-397	T2b23

VEGI12	16189c 16270t	16046-16401				U5b1b1
VEGI13	16343g 16390a	16057-16401				U3a
VEGI14	n.d.	-				n.d.
VEGI15	16298c	15997-16409				HV0
VEGI16	16093c 16256t 16270t 16320t 16399g	16057-16409				U5a1c
VEGI17	16051g 16179t 16274a	16046-16401				U2
VEGI18	16126c 16147t 16294t 16296t 16297c 16304c	15997-16409	73g 263g	309.1c 315.1c	63-396	T2b23
VEGI19	16092c 16129a 16147a 16154c 16172c 16223t 16248t 16311c 16320t 16355t	15997-16401				N1a1a1a3
VEGI20	16069t 16126c	16057-16409	73g 185a 228a 263g 295t	315.1c	46-397	J1c
VEGI21	16224c 16311c	15997-16401	73g 152c 263g	315.1c	61-397	K
VEGI22	16224c 16311c	16028-16401	73g 114t 263g	315.1c	60-397	K1a
VEGI23	16126c 16153a 16292t 16294t	16046-16409				T2e
VEGI24	n.d.	-				n.d.
VEGI25	16092c 16129a 16147a 16154c 16172c 16223t 16248t 16320t 16355t	16019-16409				N1a1a1a3
SEKU1	16093c 16224c 16311c	16046-16409				K
SEKU2	n.d.	-				n.d.
SEKU3	16126c 16294t 16304c	15997-16409				T2b
SEKU4	n.d.	-				n.d.
SEKU5	16147A 16154c 16172c 16223t 16248t 16320t 16355t	15999-16409				N1a1a1a3
SEKU6	-	15999-16409				H
SEKU7	16224c 16311c 16319a	15999-16409				K1b1a
SEKU8	16093c 16224c 16311c 16355t	16046-16409				K
SEKU9	n.d.	-				n.d.
SEKU10	16224c 16311c	15999-16409				K
SEKU11	16069t 16126c	15999-16409				J
SZEH1	16093c	16020-16401				H
SZEH2	16147A 16172c 16193t 16223t 16248t 16355t	16047-16409				N1a1a1
SZEH3	16092c 16129a 16147a 16154c 16172c 16223t 16248t 16320t 16355t	16033-16409				N1a1a1a3
SZEH5	16093c 16224c 16311c 16319a	16037-16409				K
SZEH6	16304c	16046-16401				H5
SZEH7	16093c 16224c 16311c	16026-16409				K1a
SZEH8	16069t 16126c	16019-16409				J
NEK1	16126c 16294t 16296t	16004-16406				T2
NEB1	16224c 16311c	16004-16406				K
NEB2	16183c 16189c 16223t 16278t	16045-16401	73g 153g 195c 225a 226c 263g?	309.1c 315.1c	not completed	X2b
RADOV1	16069t 16126c	16026-16409				J
BICS1	16298c	16047-16409				HV0
BICS2	16069t 16126c	16019-16401	73g 185a 228a 263g? 295t	309.1c 315.1c	not completed	J1c
BICS3	16069t 16126c 16145a 16261t	15999-16409				J1b
BICS4	16299g	16046-16401				H39
BICS5	16069t 16126c	16046-16402	73g 185a 263g 295t	309.1c 315.1c	35-396	J1c
FAGA1	16298c	16000-16409				HV0

FAGA2	-	16000-16409	152c 263g	315.1c	35-397	H
FAGA3	-	15999-16409	152c 263g	315.1c	35-397	H
ALE1	16328t 16343g	16026-16405				U3
ALE2	16304c	16019-16401	263g	315.1c	36-397	H5
ALE3	16304c	15997-16401	93g 263g	309.2c 315.1c	34-397	H5a3a1
ALE4	16126c 16292t 16294t 16296t	16012-16409				T2c1
ALE5	16170T 16172c 16224c 16311c 16344t	16046-16409				K
ALE6	16311c	16046-16401				HV
ALE7	-	16057-16401				H
ALE8	16126c 16294t 16296t 16304c	16058-16401				T2b
ALE9	16304c	16058-16409	93g 263g	309.2c 315.1c	61-397	H5a3a1
ALE10	16224c 16311c	16033-16401				K
ALE11	16304c	16057-16401	263g	315.1c	60-370	H5
ALE12	16298c 16325c	16025-16409				HV0
ALE14	16183c 16189c 16234t 16294t 16324c	16048-16409				U8b1b
ALE15	16147A 16172c 16223t 16248t 16355t	16046-16401	73g 199c 204c 263g	315.1c	34-382	N1a1a1
ALE16	16147A 16172c 16223t 16248t 16355t	16020-16404	73g 199c 204c 263g	315.1c	34-372	N1a1a1
ALE17	16126c 16147t 16294t 16296t 16297c 16304c	16012-16409				T2b23
ALE18	16187t 16192t 16256t 16270t	16046-16401				U5a
ALE19	16126c 16192t 16292t 16294t 16296t	16033-16409				T2c
VEM1.	16126c 16294t 16296t 16304c	16033-16409	73g 263g	315.1c	47-397	T2b
VEM2.	-	16023-16404				H
VEM3.	16126c 16294t 16296t 16304c	16002-16401	73g (263g) ?	309.1c 315.1c	58-397	T2b
VEJ1	16178c 16183c 16189c 16234t 16324c	15997-16409	73g 195c 263g	315.1c	60-381	U8b1b
VEJ2	16126c 16294t 16296t 16304c	15997-16409	73g 263g	309.1c 315.1c	35-383	T2b
VEJ3	16094c 16178c 16183c 16189c 16234t 16324c	15997-16406	73g 195c 263g	315.1c	49-381	U8b1b1
VEJ4	16223t 16292t	16013-16409	73g 119c 189g 195c 204c 207a 263g	309.1c 315.1c	35-397	W1b
VEJ5	16069t 16126c	16014-16405				J
VEJ6	16066g 16129a 16183c 16189c 16234t 16274a	15997-16409				U8b1a1
VEJ7	16094c 16178c 16183c 16189c 16234t 16324c	16013-16409	73g 195c 263g	315.1c	51-397	U8b1b1
VEJ8	-	16016-16409	263g	309.1c 315.1c	34-397	H
FEB1	16311c	16006-16406				HV
FEB2	16126c 16294t 16296t 16304c	16004-16401				T2b
FEB3	-	15997-16406				H
FEB4	16270t 16274a 16311c 16362c	16012-16409				U5b
FEB5	16093c	16003-16409				H
BAL 1	16093c 16224c 16311c	16045-16401	73g 195c 263g	315.1c	34-382	K1a1
BAL 2	n.d.	-				n.d.
BAL 3	16126c 16189c 16294t 16296t	16046-16401	73g 263g	309.1c 315.1c	46-394	T2f
BAL 4	16126c 16147t 16294t 16296t 16297c 16304c	16057-16401				T2b23
BAL 5	16069t 16126c	16013-16401				J
BAL 6	16093c 16224c 16311c	16034-16409	73g 114t 263g	315.1c	59-397	K1a1
BAL 7	16069t 16126c	16032-16401	73g 185a 228a 263g 295t	309.1c 315.1c	35-382	J1c

BAL 8	16093c 16224c 16311c	16057-16401	73g 114t 263g	315.1c	60-384	K1a1
BAL 9	16304c	16046-16401	263g	315.1c	34-397	H5
BAL10	16093c 16224c 16234t 16311c	16046-16409	73g 199c 263g	315.1c	46-376	K1a1b1a
BAL 11	16093c 16224c 16234t 16311c	16051-16402	73g 199c 263g	315.1c	34-397	K1a1b1a
BAL 12	16298c	16046-16409				HV0
BAL 13	16304c	16025-16409	263g	315.1c	34-397	H5
BAL 14	16069t 16126c 16283t	16020-16409				J
BAL 15	16126c 16189c 16294t 16296t	16055-16401	73g 263g	309.1c 315.1c	61-394	T2f
BAL 16	16147A 16172c 16223t 16248t	16024-16409				N1a1a1
BAL 17	n.d.	-				n.d.
BAL 18	16192t 16270t	16020-16400				U5
BAL19	16311c	16025-16405				HV
BAL 21	16126c 16294t 16304c	16046-16409				T2b
BAL22	16069t 16126c	16007-16409	73g 185a 228a 263g 295t	315.1c	46-397	J1c
BAL24	16183c 16189c 16234t 16324c	16006-16404				U8b1b
BAL25	16093c 16224c 16311c 16319a	16056-16404				K
BAL26	16147A 16172c 16223t 16248t 16355t	16023-16401				N1a1a1
BAL27	16224c 16311c 16319a	16019-16401				K1b1a
BAM27	16147A 16172c 16223t 16248t 16355t	16057-16401				N1a1a1
BORK1	16224c 16311c 16398a	16046-16401				K
CSAT1	16147A 16172c 16193t 16223t 16248t 16355t	16034-16409				N1a1a1
CSAT2	16093c	16022-16409				H
CSAT3	16298c	16046-16409	72c 263g	309.2c 315.1c	35-397	HV0
CSAT5	16093c 16311c	16020-16401				H
CSAT6	16126c 16292t 16294t	16019-16409	73g 146c 152c 263g 279c	309.1c 315.1c	34-397	T2c1d2
CSAT7	16192t 16256t 16270t	15999-16409	72c 263g?	315.1c		U5a
CSAT8	n.d.	-				n.d.
CSAT9	16093c 16129a	16019-16401				H
CSAT11	16126c 16292t 16294t	15999-16409	73g 146c 152c 195c 263g 279c	309.1c 315.1c	34-397	T2c1d2
CSAT12	16093c 16224c 16311c	15999-16409	73g 114t? 263g	315.1c	39-397	K1a1
CSAT14	n.d.	-				n.d.
CSAT15	16183c 16189c 16223t 16278t	16021-16401				X
CSAT16	16093c 16224c 16311c	16022-16401	73g 195c 263g	315.1c	34-397	K1
CSAT17	16298c	16020-16401	72c 263g	309.2c 315.1c	45-397	HV0
CSAT18	16129a	15997-16409	263g?	309.1c 315.1c		H
CSAT19	-	15999-16409				H
CSAT20	16147A 16172c 16223t 16248t 16355t	15998-16401				N1a1a1
CSAT21	16304c	15999-16409				H5
CSAT22	16126c 16292t 16294t	15999-16409	73g 146c 152c 195c 263g 279c	309.1c 315.1c	36-397	T2c1d2
CSAT23	n.d.	-				n.d.
CSAT24	16086c 16147a 16172c 16223t 16248t 16320t 16355t	15999-16401				N1a1a1a
CSAT25	16126c 16294t 16296t 16304c	15999-16409				T2b
CSAT26	16126c 16147t 16294t 16296t 16297c 16304c	15999-16409				T2b23
CSAT27	16126c 16292t 16294t	15999-16409	73g 146c 152c 263g 279c	309.1c 315.1c	36-397	T2c1d2

CSAT29	16147A 16172c 16223t 16248t 16355t	16001-16409				N1a1a1
CSAT30	16129a 16171g 16192t 16270t	15999-16409	73g 150t 263g	309.1c 315.1c	34-397	U5b
CSAT31	16129a 16171g 16192t 16270t	15999-16409	73g 150t 263g	309.1c 315.1c	35-397	U5b
MORT1	n.d.	-				n.d.
MORT2	16304c	15997-16409				H5
MORT3	n.d.	-				n.d.
MORT4	16311c	15996-16401				HV
MORT5	16086c 16147a 16172c 16223t 16248t 16320t 16355t	16021-16401				N1a1a1a
MORT6	16093c 16224c 16311c	15999-16409	73g? 114t 263g ?	315.1c	34-397	K1a1
MORT7	16147A 16172c 16223t 16248t 16320t 16355t	16046-16402				N1a1a1a
MORT8	16069t 16126c	16020-16402				J
MORT9	16093c 16311c	15999-16409	263g	315.1c	35-397	H
MORT11	16093c	16019-16409	152c 263g	315.1c	35-397	H
MORT12	16147A 16172c 16223t 16248t 16355t	16000-16409				N1a1a1
MORT13	16093c 16311c	16045-16409				H
MORT15	16224c 16311c	15999-16409	73g 152c 195c 263g	315.1c	36-397	K
MORT16	16093c	16020-16401				H
MORT18	16093c 16224c 16311c	15999-16409	73g 114t 263g	315.1c	35-397	K1a1
MORT19	16224c 16311c	15999-16409	73g 146c 152c 263g	309.1c 315.1c	36-397	K1c
MORT21	16093c 16224c 16311c	16018-16404				K
MORT25	-	15997-16409	263g	315.1c	36-397	H
KON 7	16126c 16189c 16294t 16296t	16048-16401				T2f
KON 2	16224c 16311c 16398a	16057-16401				K
VEJ9	-	16017-16409	263g	309.1c 315.1c	39-397	H
VEJ10	16223t 16292t	15999-16409	73g 119c 152c 189g 195c 204c 207a 263g	309.1c 315.1c	34-397	W1b
VEJ11	16223t 16292t	15999-16409	73g 119c 152c 189g 195c 204c 207a 263g	309.1c 315.1c	34-397	W1b
VEJ12	16172c 16189c 16234t 16311c 16352c	15999-16409				U8b1
KEFP1	16157c 16311c	16002-16409				HV
KEFP2	16069t 16126c 16193t 16278t	16012-16409				J2b1a
KEFP3	16126c 16163g 16186t 16189c 16294t	16027-16409				T1a
KEFP4	16291t	16016-16409	263g	315.1c	37-390	H
KEFP5	16126c 16294t 16296t 16301t	16000-16409				T2
KEFP6	n.d.	-				n.d.
KEFP7	n.d.	-				n.d.
KEFP13	16291t	16022-16409	263g	315.1c	35-394	H
KEFP8	16311c	16036-16392				HV
KEFP9	n.d.	16045-16409				n.d.
KEFP10	16126c 16163g 16186t 16189c 16294t	16047-16400				T1a
KEFP11	16126c 16187t 16189c 16294t 16296t	16026-16409				T2f
KEFP12	16192t 16270t	16039-16409				U5
KEFP14	16093c 16129a	16039-16409				H
TOLM1	n.d.	-				n.d.
TOLM2	16126c 16292t 16294t	16013-16409				T2c1
M6-116.12	16126c 16189c 16294t 16295t 16296t	16046-16401				T2f

M6-116.2	16093c 16126c 16294t 16296t 16304c	16003-16409	T2b
M6-116.3	16126c 16294t 16296t 16304c	16006-16409	T2b
M6-116.8	16126c 16147t 16294t 16296t 16297c 16304c 16362c	16057-16401	T2b23
M6-116.10	16304c	16057-16401	H5
KON 6	16093c 16189c 16270t	16056-16401	U5b1
ALE13	16093c 16192t 16256t 16270t 16399g	16058-16401	U5a1

Supplementary Table 4. Summary of the GenoCoRe22 mtDNA multiplex results.

Detailed results of the GenoCoRe22 SNP assay are summarized to a consensus profile of reproduced alleles and orientated to the forward direction (L-strand).

		mtDNA coding region SNP (Geno Core22) consensus																					
		R9	L3'4	K	A	U	W	C	T	R0/preHV	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
Ind. name in the lab	Geno CoRe22	G13928	T3594	A10550	T4248	A11467	G8994	A13263	G13368	A11719	G4580	A8280	C6371	T10238	T10034	T7028	C5178	T14766	T12705	C10873	A12612	A2758	C10400
	ancestral	g	t	a	t	a	g	a	g	a	g	a	c	t	t	t	c	t	t	c	a	a	c
	derived	c/a	c	g	c	g	a	g	a	g	a	g	t	c	c	c	a	c	c	t	g	g	t
STANKO	U	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 1	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 2	K	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAM 3	W	G	-	A	T	a	A	A	G	A	G	A	C	T	T	t	C	t	T	T	A	G	C
BAM 4	K	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAM 5	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 6	V	G	c	A	T	A	G	A	G	G	A	A	C	T	T	T	C	C	C	T	A	G	C
BAM 7	K	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAM 8	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 9	K	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 10	H	G	c	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAM 11	H	G	c	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAM 12	U	G	c	A	T	G	G	A	g	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAM 13	X	G	C	A	T	A	G	A	G	A	G	A	T	T	T	T	C	T	T	T	A	G	C
BAM 14	J	G	c	A	T	A	G	A	g	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BAM 15	X	G	C	A	T	A	G	A	G	A	G	A	T	T	T	T	C	T	T	T	A	G	C
BAM 16	K	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	A	C	T	A	G	C
BAM 17	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 18	J	G	C	A	T	A	G	A	G	A	G	A	C	t	T	T	C	T	C	T	G	G	C
BAM 19	K	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAM 20	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	t	C	t	C	T	A	G	C
BAM 21	T	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAM 22	N1	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
BAM 23	X	G	C	A	T	A	G	A	g	A	G	A	T	T	T	T	C	T	t	T	A	G	C
BAM 24	K	G	c	G	T	G	G	A	g	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAM 25	N1	G	c	A	T	A	G	A	g	A	G	A	C	C	T	T	C	t	t	T	A	G	C
BAM 26	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
LGCS1	W	G	C	A	T	A	A	A	g	A	G	A	C	T	T	T	C	T	T	T	A	G	C
LGCS2	N1	G	c	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
LGCS3	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
LGCS4	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6_116.1	U	G	c	A	T	G	G	A	G	A	G	A	C	T	T	t	C	-	C	T	A	G	C

		R9	L3'4	K	A	U	W	C	T	R0/preH	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
M6_116.4	K	G	C	G	T	G	G	A	-	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6_116.9	T	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VINK 1	K	G	C	G	T	G	G	A	g	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VINK 2	J	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
VINK 3	HV	G	C	A	T	A	G	A	-	G	G	A	C	T	T	T	C	C	C	T	A	G	C
VINK 4	J?	g	c	A	T	a	G	a	g	g	G	a	C	T	T	t/c	C	t	C	T	G	G	C
VINK 5	K	G	C	G	T	G	G	A	g	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VINK 6	T?	G	-	A	T	A	G	A	a	A	G	A	C	T	T	t/c	C	T	C	T	A	G	C
VINJ 1	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VINJ 2	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VINJ 3	V	G	C	A	T	A	G	A	g	G	A	A	C	T	T	T	C	C	C	T	A	G	C
VINJ 4	V	G	C	A	T	A	G	A	g	G	A	A	C	T	T	T	C	C	C	T	A	G	C
VUKG 1	J	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
VUKG 2	n.d.	g	-	A	t	a	g	a	G	a	-	a	C	T	t	c	C	c	c	t	g	g	c
VUKG 3	J	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
VUKG 4	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BSZ 1	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BSZ 2	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BSZ 3	J	G	c	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	t/c
BSZ 4	n.d.	G	-	A	T	A	G	A	G	a/g	G	A	C	T	T	t/c	C	-	C	T	A	G	C
BSZ 5	N1	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
BSZ 6	n.d.	g	c	a	t	a	g	a	g	a	-	a	c	c	t	t/c	c	t/c	c	t	a	g	c
BSZ 7	n.d.	G	-	A	T	A	G	A	G	A	a/g	A	C	T	T	T	C	c	t/c	T	A	G	C
BSZ 8	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BSZ 9	J	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BSZ 10	n.d.	-	-	-	-	-	-	-	-	g	-	-	C	-	-	-	-	-	C	-	-	-	-
BSZ 11	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BSZ 12	J	G	-	A	T	A	G	A	G	A	G	A	C	T	T	t/c	C	T	C	T	G	G	C
BSZ 13	T	G	c	A	T	A	G	A	A	A	G	A	C	-	T	t	C	T	C	T	A	G	c
BSZ 14	V	G	c	A	T	A	G	A	G	G	A	a	C	t	T	T	C	C	C	T	A	G	C
BSZ 15	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BSZ 19	HV	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
BSZ 20	J	G	c	A	t	A	G	A	G	a	G	a	C	t	T	T	C	T	C	T	G	G	C
BSZ 21	J	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BSZ 22	K	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BSZ 23	n.d.	G	-	A	T	a	g	A	a/g	A	g	A	C	T	T	c	C	c	c	T	A	G	C
BSZ 24	V	G	C	A	T	A	G	A	G	G	A	A	C	T	T	T	C	C	C	T	A	G	C
BSZ 25	n.d.	G	c	A	t	a/g	G	A	G	a/g	G	A	C	t	T	T	C	t/c	C	T	A	G	C
BSZ 26	n.d.	g	-	A	t	a/g	G	a	A	a	g	a	C	T	T	t/c	c	t/c	C	T	A	g	C
BAB 1	n.d.	g	-	A	t	-	g	-	g	g	g	-	C	c/t	-	c	-	-	c	t	-	G	c
BAB 2	n.d.	g	-	-	t	g	G	A	g	-	g	a	C	t	t	c/t	c	c	c	T	a	a	c

		R9	L3'4	K	A	U	W	C	T	R0/	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
BAB 3	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAB 4	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAB 5	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAB 6	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BÖVÖ 1	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BÖVÖ 2	J	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BÖVÖ 3	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BÖVÖ 4	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BÖVÖ 5	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 1	U	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 2	V	G	C	A	T	A	G	A	G	G	A	A	C	T	T	T	C	C	C	T	A	G	C
BUD 3	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 4	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 5	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 6	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 7	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BUD 9	U	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 10	H	G	c	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BUD 11	n.d.	G	C	A	T	A	A	A	G	A	G	A	C	T	T	T	C	C	t/c	T	A	G	C
BUD 12	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 13	H	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BUD 14	K	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BUD 15	H	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
HARG 1	T	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
HARG 2	N1	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
HARG 3	T	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
HARG 4	K	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
HARG 5	T	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KON 1	H	G	c	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
KON 3	T	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KON 4	T	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KON 5	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
SZEH 4	N1	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
SZEH 9	N1	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
TOLM 3	T	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
TOLM 4	K	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
NITR2	K	G	-	G	t	G	G	A	G	A	G	A	C	T	T	T	c	T	C	T	A	G	C
NITR7	K	G	-	G	t	G	G	A	G	A	G	a	C	T	T	T	C	T	C	T	A	G	C
NITR11	N	G	-	a	t	A	G	A	G	A	G	a	C	T	T	T	C	t	t	T	A	G	C
NITR14	U	G	-	A	t	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
NITR26	n.d.	-	-	a	t	A	g	a	G	-	-	-	C	-	t	c	C	c	c	-	g	-	-
VEGI1	**T**	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI2	**J**	G	c	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
VEGI3	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI4	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI5	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI6	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI7	**K**	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C

		R9	L3'4	K	A	U	W	C	T	R0/	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
VEGI8	**N1**	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
VEGI9	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI10	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI11	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI12	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI13	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI14	**K?**	G	c	g/a	t	G	G	A	G	A	G	a	C	t	t	t/c	C	t	C	T	a	G	C
VEGI15	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
VEGI16	**U**	G	c	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI17	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI18	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI19	**N1**	G	c	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
VEGI20	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
VEGI21	**K**	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI22	**K**	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI23	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEGI24	n.d.	G	-	A	T	A	G	A	G	g/a	G	A	C	t/c	T	t/c	C	t/c	C	T	A	G	C
VEGI25	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
SEKU1	**K**	G	-	G	T	G	G	A	G	A	G	a	C	T	T	T	C	T	C	T	A	G	C
SEKU2	**U?**	G		A	-	G	G	A	G	A	g	a	C	T	T	-	C	T	C	T	A	g	C
SEKU3	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
SEKU4	**T**	G	-	A	T	A	G	A	A	A	G	-	C	T	T	T	C	T	C	T	A	G	C
SEKU5	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
SEKU6	**H**	G	-	A	t	A	G	A	G	G	G	a	C	T	T	C	C	C	C	T	A	G	C
SEKU7	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
SEKU8	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	c	T	C	T	A	G	C
SEKU9	n.d.	G	-	a	t	-	g	a	a	a	g	a	C	-	t	t	c	t	c	-	a	g	C
SEKU10	**K**	G	-	G	-	G	G	A	G	A	G	a	C	T	T	T	C	T	C	T	A	G	C
SEKU11	**J**	G	-	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
SZEH1	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	t	C	C	C	C	T	A	G	C
SZEH2	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
SZEH3	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
SZEH5	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
SZEH6	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
SZEH7	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
SZEH8	**J**	G	-	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
NEK1	**T?**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	t	C	T	A	G	C
NEB1	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
NEB2	**X??**	G	-	A	T	A	G	A	G	A	G	A	c/t	T	T	T	C	T	T	T	A	G	C
RADOV1	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BICS1	**HV**	G	c	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
BICS2	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BICS3	**B,C: J**	G	c	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BICS4	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BICS5	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
FAGA1	**HV**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
FAGA2	**H**	G	-	A	t	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C

		R9	L3'4	K	A	U	W	C	T	R0/	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
FAGA3	**H**	G	-	A	t	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
ALE1	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE2	**H**	G	C	A	T	a /g	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
ALE3	**H**	G	C	A	T	A	G	A	-	G	G	A	C	T	T	C	C	C	C	T	A	G	C
ALE4	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE5	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE6	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
ALE7	**H**	G	C	A	T	A	G	A	-	G	G	A	C	T	T	C	C	C	C	T	A	G	C
ALE8	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE9	**H?**	G	c	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
ALE10	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE11	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
ALE12	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
ALE14	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE15	**N1**	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
ALE16	**N1**	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
ALE17	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE18	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE19	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEM1.	**T**	G	c	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEM2.	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
VEM3.	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ1	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ2	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ3	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ4	**W**	G	C	A	T	A	A	A	G	A	G	A	C	T	T	T	C	T	T	T	A	G	C
VEJ5	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
VEJ6	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ7	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ8	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
FEB1	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
FEB2	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
FEB3	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
FEB4	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
FEB5	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAL 1	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 2	**H**	G	C	A	T	A	G	A	-	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAL 3	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 4	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 5	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BAL 6	**K**	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAL 7	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BAL 8	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 9	**H**	G	c	A	T	a	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
BAL10	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 11	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 12	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
BAL 13	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C

		R9	L3'4	K	A	U	W	C	T	R0/	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
BAL 14	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
BAL 15	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL 16	**N1**	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
BAL 17	n.d.	G	-	A	T	a	G	A	A	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAL 18	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL19	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
BAL 21	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	t	C	T	A	G	C
BAL22	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	-	C	T	G	G	C
BAL24	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL25	**K**	G	C	G	T	G	G	A	g	A	G	A	C	T	T	T	C	T	C	T	A	G	C
BAL26	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	t	C	T	A	G	C
BAL27	**K?**	G	-	G	T	g/a	G	A	G	A	g/a	A	C	T	T	T	C	c/t	C	T	A	g	C
BAM27	**N1**	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
BORK1	**K**	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT1	**N1**	G	C	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
CSAT2	**H**	G	c	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
CSAT3	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
CSAT5	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
CSAT6	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT7	**U**	G	-	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT8	**U**	G	-	A	t	G	G	A	G	A	g	a	C	t	T	T	C	T	C	T	A	G	C
CSAT9	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
CSAT11	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT12	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT14	n.d.	G	-	A	-	a/g	G	a	G	A	g	-	C	-	T	T	C	T	C	T	G	G	C
CSAT15	**X??**	G	-	A	T	A	G	A	G	A	G	a	c/t	T	T	T	C	T	T	T	A	G	C
CSAT16	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT17	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	c	C	C	T	A	G	C
CSAT18	**H**	G	-	A	t	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
CSAT19	**H**	G	-	A	t	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
CSAT20	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
CSAT21	**H**	G	-	A	-	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
CSAT22	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT23	n.d.	G	-	a	-	a	g	a	g	g	a	a	c	t	t	t	c	c	c	t	a	g	C
CSAT24	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
CSAT25	**T**	G	-	A	T	A	G	A	A	A	G	A	C	T	T	t/c	C	T	C	T	A	G	C
CSAT26	**T**	G	-	A	-	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT27	**T**	G	-	A	-	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
CSAT29	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
CSAT30	**U**	G	-	A	T	G	G	A	G	A	G	-	C	T	T	T	C	T	C	T	A	G	C
CSAT31	**U**	G	-	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
MORT1	**T**	G	-	A	-	-	G	A	A	A	-	-	C	-	-	T	-	T	C	T	A	G	C
MORT2	**H**	G	-	A	-	A	G	A	G	G	G	a	C	T	T	C	c	C	C	T	A	g	C
MORT3	**J?**	G	-	A	-	A	G	A	g	A	G	A	c	T	T	T	C	T	C	T	G	G	C
MORT4	**HV**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
MORT5	**N1**	G	-	A	t	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
MORT6	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C

		R9	L3'4	K	A	U	W	C	T	R0/	V	B	X	N1	I	H	D	HV	R	N	J	L2'6	M
MORT7	**N1**	G	-	A	T	A	G	A	G	A	G	A	C	C	T	T	C	T	T	T	A	G	C
MORT8	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
MORT9	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
MORT11	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
MORT12	**N1**	G	C	A	T	A	G	a	g	A	G	A	C	C	T	T	C	t	t	T	A	G	C
MORT13	**H**	G	-	A	t	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
MORT15	**K**	G	C	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
MORT16	**H**	G	-	A	t	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
MORT18	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
MORT19	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
MORT21	**K**	G	-	G	t	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
MORT25	**H**	G	-	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
KON 7	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KON 2	**K**	G	c	G	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
VEJ9	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
VEJ10	**W**	G	C	A	T	A	A	A	G	A	G	A	C	T	T	T	C	T	T	T	A	G	C
VEJ11	**W**	G	C	A	T	A	A	A	G	A	G	A	C	T	T	T	C	T	T	T	A	G	C
VEJ12	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KEFP1	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
KEFP2	**J**	G	C	A	T	A	G	A	G	A	G	A	C	T	T	T	C	T	C	T	G	G	C
KEFP3	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KEFP4	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
KEFP5	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KEFP6	n.d.	G	-	A	T	A	G	A	G	g/ a	G	A	C	T	T	C t	C	c/ t	C	T	a /g	G	c/ t
KEFP7	**K**	G	-	G	T	G	G	A	G	A	G	A	C	T	T	c/ t	C	T	C	T	A	G	C
KEFP13	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
KEFP8	**HV**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	T	C	C	C	T	A	G	C
KEFP9	**HV**	G	C	A	T	A	G	A	G	g/ a	G	A	C	T	T	T	C	C	C	T	A	G	C
KEFP10	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KEFP11	**T**	G	-	A	T	A	G	A	A	G	A	A	C	T	T	T	C	T	C	T	A	G	C
KEFP12	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
KEFP14	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
TOLM1	**H**	G	c	a	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
TOLM2	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6-116.12	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6-116.2	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6-116.3	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6-116.8	**T**	G	C	A	T	A	G	A	A	A	G	A	C	T	T	T	C	T	C	T	A	G	C
M6-116.10	**H**	G	C	A	T	A	G	A	G	G	G	A	C	T	T	C	C	C	C	T	A	G	C
KON 6	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C
ALE13	**U**	G	C	A	T	G	G	A	G	A	G	A	C	T	T	T	C	T	C	T	A	G	C

Capital letters indicate alleles that were reproduced at least two times from two different samples. Lower-case letters signify positions detected in just one sample. Hyphens indicate non- reproduced alleles or allelic dropout. Inconsistent positions are represented by slashes.

Ind. Name in the lab	NRY haplogroup	Y 25 Multiplex consensus																									Y singleplex consensus												
		M174	M231	M242	S21	M168	M45	SRY10831.1	M2	M89	M267	M343	M172	M304	M35.1	M175	M201	M170	M269	M122	M17	M96	M78	M9	M207	M216	M170	M253	M438(=P215)	P37.2	M223	M201	M285	P287	P15	S126	M69	M304	M172
	ancestral	t	g	c	c	c	g	t	a	c	t	c	t	a	g	t	g	a	t	t	g	g	c	c	a	c	a	c	a	t	g	g	g	g	c	c	t	a	t
	derived	c	a	t	t	t	a	c	g	t	g	a	g	c	c	a	t	c	c	c	c	c	t	g	g	t	c	t	g	c	a	t	c	t	t	t	c	c	g
BAM 4	G2a	T	G	C	c	T	G	C	a	T	T	c	T	a	g	T	T	A	T	t	g	G	c	c	a	c								t	t	c			
BAM 11	G2a	-	G	C	-	t	g	-	A	t	-	c	t	-	-	T	t	a	t	-	-	g	-	c	-	t									t	c			
BAM 13	G2a2b	T	G	C	C	T	G	C	A	T	T	C	T	a	g	T	T	A	T	T	G	G	C	C	A	C								t	t	t			
BAM 17	F*	T	G	C	C	T	G	C	A	T	T	C	T	A	G	T	G	A	T	T	G	G	C	C	A	C											t		
BAM 19	G2a	t	G	c	c	t	G	-	a	t	t	-	-	a	-	t	T	-	t	-	-	g	-	c	a	-								t	t				
BAM 25	F*	T	G	C	C	T	G	C	A	T	T	C	T	A	-	T	G	A	T	T	g	G	c	C	a	C											t		t
BAM 26	F*	T	G	C	C	T	G	c	a	T	T	C	t	-	g	T	G	A	T	T	g	G	c	C	a	c											t	a	t
LGCS1	G2a2b	T	G	C	c	T	G	c	a	T	T	c	T	-	-	T	T	A	t	t	g	G	c	C	a	C								t	t	t			
LGCS4	G2a2b	t	g	C	-	T	G	-	a	-	-	c	-	a	-	t	t	-	t	t	-	G	c	-	A	c/t								t	t	T			
M6-116.1	I	T	G	C	-	t	G	C	a	T	T	c	t	-	g	T	G	c	t	t	-	G	c	C	-	c	t												
VINK 5	G2a	t	g	c	c	t	g	-	-	t	t	-	t	-	-	-	T	a	t	t	g	g	-	c	-	-								t	t	c			
VINJ 4	I2a1	T	G	C	C	T	G	C	A	T	T	C	T	a	-	T	G	C	T	T	G	G	C	C	A	C				C	g								
VUKG 1	G2a	-	g	c	c	t	G	-	-	T	t	c	t	-	-	-	t	-	-	t	-	g	-	c	-	-						t		t	t	c			
BAB 5	I1	T	G	C	C	T	G	C	A	T	T	C	T	A	G	T	G	C	T	T	G	G	c	C	A	C		t	a										
BUD 4	G2a2b	T	G	C	c	T	G	c	A	T	T	c	t	a	-	t	T	A	t	T	-	g	c	C	a	c						t			t	T			
TOLM 3	F*	T	G	C	c	T	G	C	A	T	T	C	t	-	-	T	G	A	t	t	g	G	C	c	-	c													t
TOLM 4	G2a2b	T	G	C	c	t	G	C	a	t	t	c	t	-	-	T	t	a	t	t	g	g	c	-	a	c									t	t			
VEGI3	FR?	t	a/g	c	c	t	G	-	-	T	t	c	t	-	-	t	G	a	t	t	g	g	-	c	a	c/t													
SEKU1	G2a	T	G	C	c	T	G	c	-	T	-	c	-	-	g	T	t	a	T	t	-	g	c	C	a	-								t	t	c			
BICS3	I2	T	G	C	C	T	G	C	A	T	T	C	T	a	g	T	G	C	T	T	G	G	C	C	A	C			g	t									
BICS4	E1b1b1a1	T	G	C	C	T	G	C	A	C	T	C	t	A	c	T	G	A	T	T	G	C	T	C	A	C													
ALE14	J2	T	G	C	c	T	G	-	a	T	T	c	G	c	-	T	G	A	t	T	g	G	c	c	A	C													g
ALE15	I2a1	T	G	C	C	T	G	c	A	T	T	C	t	-	g	T	G	C	t	t	-	G	c	C	a	c			g	c	g								
ALE16	C	T	G	C	c	T	G	c	A	C	T	C	t	a	-	T	G	A	t	T		G	c	C	A	T													
VEJ2	C	T	G	C	c	T	G	-	A	C	t	C	-	a	g	t	g	a	T	t	g	g	c	C	a	T													
VEJ4	E/E1b1b1a1?	T	G	C	c	T	G	C	A	c	T	C	T	A	c	T	G	A	T	T	g	C	t	C	a	c													
VEJ5	G2a	T	G	C	C	T	G	C	A	T	T	C	T	a	-	T	T	A	T	T	g	G	C	C	A	C						g			t	c			
FEB3	J2	T	G	C	c	T	G	c	A	T	T	C	g	-	-	T	G	A	t	T	g	G	-	C	a	c												c	g
BAL 3	F*	T	G	C	c	T	G	C	A	T	T	C	t	-	g	T	G	A	T	T	g	G	-	-	A	C											t		t
BAL 9	F*	T	G	C	C	T	G	c	A	T	T	C	T	A	g	T	G	A	T	T	g	G	c	g	A	C									c		t		
BAL 11	F*	T	G	C	c	T	G	c	A	T	T	C	T	a	g	T	G	A	T	T	G	G	c	C	A	C											t		
BAL 12	F*	T	G	C	C	t	G	-	a	T	T	C	t	-	-	T	G	A	t	T	g	-	c	C	a	c											t		
BAL 13	F*	T	G	C	c	t	G	C	a	T	t	C	t	a	-	T	G	A	t	T	-	G	c	C	-	C									c		t		
BAL25	G2a	T	G	C	C	T	G	c	A	T	T	C	T	a	g	T	T	A	T	T	G	G	c	C	A	C								t	t	c/t			
CSAT18	G2a?	t	G	c	c	T	a/g	-	-	T	t	-	-	-	-	T	t	A	T	t	-	g	-	c	a	c									t	c			
CSAT19	F*	T	G	C	c	T	G	c	-	T	t	c	T	a	-	t	G	A	T	t	g	G	c	C	a	C											t		
CSAT24	G2a	T	G	C	C	T	G	C	A	T	T	c	t	-	g	T	T	A	T	t	g	G	c	c	a	C								t	t	c			
CSAT25	I2	T	G	C	C	T	G	C	A	T	T	C	t	a	g	T	G	C	T	T	g	G	c	C	A	C			g	c/t									
VEJ9	C	T	G	c	c	t	G	c	a	C	t	C	T	-	-	t	G	A	t	T	-	g	-	C	a	T													
VEJ12	F*	T	g	C	c	-	G	C	A	T	t	c	t	-	G	t	g	a	t	t	g	G	-	C	-	c											t		
M6-116.8	R1b	T	G	c	C	t	a/g	-	a	T	t	A	t	a	-	t	g	a	-	T	G	G	-	G	g	c													
M6-116.10	I2a2	T	G	C	C	T	G	C	A	T	T	C	T	A	g	T	G	C	T	T	G	G	C	C	A	C			g	t	a								
KON 6	R1b1a2	T	G	C	C	T	A	C	A	T	T	A	T	a	g	T	G	A	C	T	G	G	C	G	G	C													

Supplementary Table 5. Summary of the positive NRY multiplex and singleplex PCR results.

Y-chromosomal multiplex and singleplex PCR results: capital letters indicate alleles that were reproduced at least three times, lower-case letters signify alleles that were reproduced only once or twice. Hyphens indicate allelic dropout. Inconsistent positions are represented by slashes. Red letters indicate profiles, which were not considered as reproduced, and therefore not used for the population genetic analysis.

Supplementary Table 6. Comparative prehistoric data.

If the cited publication contained only uncalibrated BP dates, I performed a calibration in OxCal 4.2.3 software, IntCal 13 calibration curve. (Bronk Ramsey 2013; Reimer et al. 2013). These dates are italicised.

Abbreviation	Date (as given in the reference)	Country	Archaeological site	n aDNA data	References
Hunter-gatherer Central/ North (HGCN)	~6,850 cal BC	Germany	Bad Dürrenberg	1	Bramanti et al. 2009
	~6,700 cal BC	Germany	Hohlenstein-Stadel	2	Bramanti et al. 2009
	~13,400 cal BC	Germany	Hohler Fels	1	Bramanti et al. 2009
	~2,250 cal BC	Poland	Drestwo	1	Bramanti et al. 2009
	4,000-3,000 cal BC	Poland	Dudka	2	Bramanti et al. 2009
	Mesolithic	Lithunia	Donkalnis	1	Bramanti et al. 2009
	4,450-4,200 cal BC	Lithunia	Kretuonas	2	Bramanti et al. 2009
	~6,350 cal BC	Lithunia	Spiginas	1	Bramanti et al. 2009
	15,540-14,830 cal BC	Germany	Oberkassel	1	Fu et al. 2013
	6,104 +- 127 cal BC	Luxemburg	Reuland-Loschbour	1	Fu et al. 2013
	9,210-8,652 cal BC	Germany	Blätterhöhle	5	Bollongino et al. 2013
Hunter-gatherer South-west (HGSW)	Magdalenian	Spain	La Chora	1	Hervella et al. 2012
	Magdalenian	Spain	La Pasiega	1	Hervella et al. 2012
	Magdalenian (10,360 BC)	Spain	Erralla	1	Hervella et al. 2012
	5,650-5,470 cal BC	Spain	Aizpea	1	Hervella et al. 2012
	5,920-5,730 cal BC	Spain	La Brana	2	Sanchez-Quinto et al. 2012
	8,030-6,410 cal BC	Portugal	Toledo	1	Chandler et al. 2005, Chandler 2003
	5,990-5,715 cal BC	Portugal	Arapouco	2	Chandler et al. 2005, Chandler 2003
	5,065-4,715 cal BC	Portugal	Cabeço das Amoreiras	1	Chandler et al. 2005, Chandler 2003
	5,215-4,805 cal BC	Portugal	Cabeço de Pez	3	Chandler et al. 2005, Chandler 2003
	5,770-5,230 cal BC	Portugal	Poças de São Bento	1	Chandler et al. 2005, Chandler 2003
Hunter-gatherer East (HE)	~5,000 BC	Russia	Popovo	2	Der Sarkissan et al. 2013
	~5,500 BC	Russia	Yuzhnyy Oleni Ostrov	9	Der Sarkissan et al. 2013
	~7,800 cal BC	Russia	ChekalinoI	1	Bramanti et al. 2009
	8,000-7,000 cal BC	Russia	Lebyazhinka	1	Bramanti et al. 2009
Linear Pottery culture-Central-Europe (LBK)	5,500-4,775 cal BC	Austria	Asparn Schletz 2	1	Haak et al. 2005
		Czech Republic	Vedrovice	6	Bramanti 2008
		Germany	Flomborn	6	Haak et al. 2005
		Germany	Schwetzingen	4	Haak et al. 2005
		Germany	Vaihingen	1	Haak et al. 2005
		Germany	Seehausen	1	Haak et al. 2005
		Germany	Derenburg	22	Haak et al. 2005

		Germany	Halberstadt	31	Haak et al. 2005, Brandt et al. 2013
		Germany	Karsdorf	23	Brandt et al. 2013
		Germany	Naumburg	4	Brandt et al. 2013
		Germany	Oberwiederstedt I, Unterwiederstedt	8	Haak et al. 2005, Brandt et al. 2013
		Germany	Eilsleben	1	Haak et al. 2005
Rössen culture (RSC)	4,625-4,250 cal BC	Germany	Esperstedt	1	Brandt et al. 2013
		Germany	Halberstadt, Sonntagsfeld	1	Brandt et al. 2013
		Germany	Oberwiederstedt III, Schrammhoehe	8	Brandt et al. 2013
		Germany	Oberwiederstedt IV, Arschkerbe Ost	1	Brandt et al. 2013
		Germany	Wittmar	6	Lee et al. 2013
Schöninger group (SCG)	4,100-3,950 cal BC	Germany	Salzmünde, Schiebzig	33	Brandt et al. 2013
Baalberge culture (BAC)	3,950-3,400 cal BC	Germany	Esperstedt	1	Brandt et al. 2013
		Germany	Halle-Queis	1	Brandt et al. 2013
		Germany	Karsdorf	2	Brandt et al. 2013
		Germany	Quedlinburg VII 2	7	Brandt et al. 2013
		Germany	Quedlinburg VII 3	1	Brandt et al. 2013
		Germany	Quedlinburg IX	6	Brandt et al. 2013
		Germany	Salzmünde, Schiebzig	1	Brandt et al. 2013
Salzmünde culture (SMC)	3,400-3,100/3,025 cal BC	Germany	Esperstedt	1	Brandt et al. 2013
		Germany	Salzmünde, Schiebzig	28	Brandt et al. 2013
Bernburg culture (BEC)	3,100-2,650 cal BC	Germany	Benzingerode I	17	Brandt et al. 2013
Corded Ware culture (CWC)	2,800-2,200/2,050 cal BC	Germany	Esperstedt	12	Brandt et al. 2013
		Germany	Eulau	12	Haak et al. 2008, Brandt et al. 2013
		Germany	Oberwiederstedt II	4	Brandt et al. 2013
		Germany	Quedlinburg VII 2	1	Brandt et al. 2013
		Germany	Quedlinburg XII	1	Brandt et al. 2013
		Germany	Benzingerode, Heimburg	1	Brandt et al. 2013
Bell Beaker culture (BBC)	2,500-2,200/2,050 cal BC	Germany	Alberstedt	1	Brandt et al. 2013
		Germany	Benzingerode, Heimburg	6	Brandt et al. 2013
		Germany	Eulau	3	Brandt et al. 2013
		Germany	Karsdorf	3	Brandt et al. 2013
		Germany	Quedlinburg VII 2	7	Brandt et al. 2013
		Germany	Quedlinburg XII	3	Brandt et al. 2013
		Germany	Rothenschirmbach	5	Brandt et al. 2013
		Germany	Kromsdorf	6	Lee et al. 2012
Unetice culture (UC)	2,200-1,550 cal BC	Germany	Alberstedt	1	Brandt et al. 2013
		Germany	Benzingerode, Heimburg	9	Brandt et al. 2013
		Germany	Esperstedt	11	Brandt et al. 2013
		Germany	Eulau	19	Brandt et al. 2013
		Germany	Karsdorf	12	Brandt et al. 2013

		Germany	Leau 2	3	Brandt et al. 2013
		Germany	Plötzkau 3	8	Brandt et al. 2013
		Germany	Quedlinburg VII 2	14	Brandt et al. 2013
		Germany	Quedlinburg VIII	6	Brandt et al. 2013
		Germany	Quedlinburg XII	1	Brandt et al. 2013
		Germany	Quedlinburg XIV	1	Brandt et al. 2013
		Germany	Röcken 2	9	Brandt et al. 2013
(Epi)Cardial culture (CAR)	5,329-4,999 cal BC	Spain	Chaves	1	Gamba et al. 2012
	5,475-5,305 cal BC	Spain	Can Sadurni	7	Gamba et al. 2012
	4,250-3,700 cal BC	Spain	Sant Pau del Camp	3	Gamba et al. 2012
	5,000-4,500 cal BC	Spain	Avelaner cave	7	Lacan et al. 2011b
Neolithic in Basque Country and Navarra (NQB)	*3,290-3,210 BC*	Spain	Fuente Hoz	6	Hervella et al. 2012
	3,335 BC	Spain	Marizulo	1	Hervella et al. 2012
	4,235-3,235 BC	Spain	Los Cascajos	27	Hervella et al. 2012
	4,140-4,100 BC	Spain	Paternanbidea	9	Hervella et al. 2012
Neolithic in Portugal (NPO)	3,930-3,590 cal BC	Portugal	Algar do Bom Santo	3	Chandler et al. 2005, Chandler 2003
	5,234-5,060 cal BC	Portugal	Gruta do Caldeirão	9	Chandler et al. 2005, Chandler 2003
	3,500-3,000 BC	Portugal	Perdigões	6	Chandler et al. 2005, Chandler 2003
Treilles culture (TRE)	3,030–2,890 cal BC	France	Treilles	29	Lacan et al. 2011a
Funnel Beaker culture (FBC)	~3,200-2,950 cal BC	Germany	Ostorf	7	Bramanti et al. 2009
	3,500-2,500 BC	Sweden	Frälsegården	3	Malmström et al. 2009, Skoglund et al. 2012
Pitted Ware culture (PWC)	2,800-2,200 BC	Sweden	Ajvide	11	Malmström et al. 2009, Skoglund et al. 2012
		Sweden	Fridtorp	4	Malmström et al. 2009, Skoglund et al. 2012
	3,300-,2400 BC	Sweden	Ire	4	Malmström et al. 2009, Skoglund et al. 2012
Bronze Age Siberia (BAS)	1,800–1,100 BC	Russia	Solenoozernaïa I,IV	6	Keyser et al. 2009
		Russia	Tatarka	2	Keyser et al. 2009
		Russia	Oust-Abakansty	2	Keyser et al. 2009
		Russia	Bogratsky	1	Keyser et al. 2009
Bronze Age Kazakhstan (BAK)	1,300-1,100 BC	Kazakhstan	Ak-Mustafa	1	Lalueza-Fox et al. 2004
		Kazakhstan	Izmaylovka	2	Lalueza-Fox et al. 2004
		Kazakhstan	Oi-Zhaylau-III	2	Lalueza-Fox et al. 2004
		Kazakhstan	Vodokhranilische	3	Lalueza-Fox et al. 2004
Eneolithic North Pontic steppe (ENL)	~4,500-3,000 BC	Bulgaria	Smyadovo	4	Wilde et al. 2014
	~3,500-2,000 BC	Bulgaria	Durankulak	1	Wilde et al. 2014
		Ukraine	Mayaki	2	Wilde et al. 2014
		Ukraine	Molyukhov Bugor	2	Wilde et al. 2014
Yamnaya culture (YAM)	~3,000-2,500 BC	Bulgaria	Benkovski	1	Wilde et al. 2014
		Russia	Kalinovka I	1	Wilde et al. 2014
		Ukraine	Mayaki	3	Wilde et al. 2014
		Russia	Nikolaevka III	2	Wilde et al. 2014

		Russia	Olennii	2	Wilde et al. 2014
		Bulgaria	Ovchartsi	1	Wilde et al. 2014
		Russia	Peschanyi	1	Wilde et al. 2014
		Ukraine	Pestchanka II	1	Wilde et al. 2014
		Russia	Podlesnyj	2	Wilde et al. 2014
		Bulgaria	Popovo	3	Wilde et al. 2014
		Bulgaria	Riltsi	1	Wilde et al. 2014
		Ukraine	Kirovograd Sugokleya	4	Wilde et al. 2014
		Moldova	Tetcani	1	Wilde et al. 2014
		Ukraine	Vinogradnoe	3	Wilde et al. 2014
Catacomb culture (CAT)	~2,700-2,500 BC	Ukraine	Krasnorechensk	1	Wilde et al. 2014
		Ukraine	Lisichansk	3	Wilde et al. 2014
		Ukraine	Shakhta Stepnaya	1	Wilde et al. 2014
		Russia	Temrta III, V	5	Wilde et al. 2014
	~2,500-2,000 BC	Ukraine	Nevskoe	2	Wilde et al. 2014
		Ukraine	Novozvanovka II	2	Wilde et al. 2014
		Russia	Peschanyi	4	Wilde et al. 2014
		Ukraine	Kirovograd Sugokleya	1	Wilde et al. 2014
		Russia	Temrta III	2	Wilde et al. 2014
		Moldova	Tetcani	1	Wilde et al. 2014
		Ukraine	Vinogradnoe	2	Wilde et al. 2014

abbr.	n	C	D	Z	H	HV	V	I	J	K	N	N1a	R	T	T1	T2	U	U2	U3	U4	U5	U5a	U5b	U8	W	X
HGCN	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10.5	5.26	0	10.5	0	21.1	52.6	0	0	0
HGSW	14	0	0	0	42.9	0	0	0	0	0	14.3	0	0	0	0	0	0	0	0	7.14	0	0	35.7	0	0	0
HGE	13	21.4	0	0	7.14	0	0	0	0	0	0	0	0	0	0	0	0	14.3	0	28.6	0	21.4	0	0	0	0
STA	44	0	0	0	6.82	2.27	6.82	0	11.4	27.3	0	6.82	0	0	2.27	20.5	0	0	2.27	2.27	0	0	0	0	4.55	6.82
LBKT	42	0	0	0	30.8	2.56	2.56	0	7.69	12.8	0	10.3	0	0	2.56	25.6	0	2.56	0	0	0	2.56	0	0	0	0
VIN	31	0	0	0	3.23	3.23	0	0	9.68	29	0	12.9	0	0	0	19.4	0	3.23	9.68	0	0	6.45	3.23	0	0	0
SOP	37	0	0	0	27	10.8	0	0	13.5	13.5	0	10.8	0	0	0	13.5	0	0	2.7	0	0	2.7	0	2.7	0	2.7
LGY	82	0	0	0	22	7.32	0	0	7.32	18.3	0	12.2	0	0	0	18.3	0	0	0	0	1.22	1.22	3.66	6.1	1.22	1.22
BL	13	0	0	0	23.1	7.69	0	0	7.69	7.69	0	0	0	0	7.69	23.1	0	0	0	0	0	0	0	7.69	15.4	0
LBK	108	0	0	0	16.7	4.63	4.63	0	12	20.4	0	12	0	0	0	22.2	0	0	0.93	0	0	1.85	0.93	0	2.78	0.93
RSC	17	0	0	0	29.4	23.5	5.88	0	0	11.8	0	5.88	0	0	0	11.8	0	0	0	0	0	0	5.88	0	0	5.88
SCG	33	0	0	0	15.2	3.03	0	0	15.2	30.3	0	3.03	0	0	0	12.1	0	0	0	0	0	0	6.06	3.03	9.09	3.03
BAC	19	0	0	0	26.3	5.26	0	0	5.26	10.5	0	5.26	0	0	5.26	21.1	0	0	0	0	0	0	5.26	5.26	0	10.5
SMC	29	0	0	0	31	3.45	3.45	0	20.7	10.3	0	6.9	0	0	0	6.9	0	0	10.3	0	0	0	3.45	0	0	3.45
BEC	17	0	0	0	23.5	0	5.88	0	0	17.6	0	0	0	0	0	11.8	0	0	0	0	0	11.8	17.6	0	5.88	5.88
CWC	44	0	0	0	22.7	2.27	0	2.27	9.09	13.6	0	0	0	0	6.82	11.4	0	2.27	0	6.82	0	9.09	4.55	0	2.27	6.82
BBC	35	0	0	0	40	0	0	2.86	2.86	5.71	0	0	0	0	5.71	5.71	0	2.86	0	5.71	0	14.3	5.71	0	8.57	0
UC	94	0	0	0	21.3	2.13	3.19	12.8	6.38	7.45	0	0	1.06	0	2.13	6.38	2.13	6.38	0	1.06	0	12.8	2.13	3.19	4.26	5.32
NPO	18	0	0	0	66.7	0	5.56	0	0	0	0	0	0	0	0	0	5.56	0	0	0	0	16.7	5.56	0	0	0
CAR	20	0	0	0	25	0	0	0	0	30	20	0	0	0	0	10	0	0	0	0	0	0	10	0	0	5
NBQ	43	0	0	0	44.2	2.33	0	2.33	4.65	9.3	0	0	0	0	0	2.33	25.6	0	0	0	0	0	6.98	0	0	2.33
TRE	29	0	0	0	20.7	6.9	3.45	0	20.7	6.9	0	0	0	0	0	6.9	3.45	0	0	0	0	3.45	13.8	0	0	13.8
FBC	10	0	0	0	10	0	0	0	20	10	0	0	0	0	0	30	0	0	0	0	0	10	20	0	0	0
PWC	19	0	0	0	0	10.5	5.26	0	0	5.26	0	0	0	0	0	5.26	0	0	0	42.1	0	15.8	15.8	0	0	0
BAS	11	0	0	9.09	9.09	0	0	0	0	9.09	0	0	0	0	9.09	9.09	0	9.09	0	27.3	0	18.2	0	0	0	0
BAK	8	0	0	0	12.5	12.5	0	12.5	0	0	0	0	0	0	25	12.5	0	0	0	0	0	12.5	12.5	0	0	0
ENL	9	0	0	0	33.3	0	0	0	0	0	0	0	0	0	0	22.2	0	0	0	0	0	33.3	0	0	11.1	0
YAM	26	0	3.85	0	19.2	0	0	3.85	3.85	7.69	0	3.85	0	7.69	11.5	3.85	3.85	3.85	0	3.85	0	15.4	0	0	3.85	3.85
CAT	24	0	0	0	25	4.17	0	4.17	8.33	0	0	0	8.33	0	8.33	0	0	0	0	29.2	0	12.5	0	0	0	0

Supplementary table 7. Relative mtDNA haplogroup frequencies for the PCA with 29 prehistoric population (29POP).

For culture information, see Supplementary Table S6

	HGCN	HGSW	HGE	STA	LBKT	VIN	SOP	LGY	BL	LBK	RSC	SCG	BAC	SMC	BEC	CWC	BBC	UC	NPO	CAR	NBQ	TRE	PWC	FBC	BAS	BAK	ENL	YAM	CAT
HGCN	*	*0.07306*	*0.02247*	*0.00000*	*0.00000*	*0.00000*	*0.00000*	*0.00000*	*0.00000*	*0.00010*	*0.00000*	*0.00000*	*0.00000*	*0.00000*	*0.01614*	*0.00010*	*0.00040*	*0.00010*	*0.00109*	*0.00000*	*0.00000*	*0.00010*	*0.00099*	*0.00277*	*0.01168*	*0.01762*	*0.07989*	*0.00000*	*0.0001*
HGSW	0.0502	*	*0.45451*	*0.00149*	*0.00079*	*0.00416*	*0.00050*	*0.00297*	*0.00079*	*0.00099*	*0.00485*	*0.00059*	*0.01733*	*0.00000*	*0.18572*	*0.04861*	*0.02871*	*0.03782*	*0.20820*	*0.00752*	*0.00000*	*0.01040*	*0.02267*	*0.00653*	*0.33353*	*0.04514*	*0.03227*	*0.01792*	*0.0381*
HGE	0.0995	-0.0046	*	*0.00119*	*0.00109*	*0.00238*	*0.00040*	*0.00178*	*0.00188*	*0.00059*	*0.02960*	*0.00020*	*0.03614*	*0.00030*	*0.09356*	*0.02435*	*0.02089*	*0.01317*	*0.18691*	*0.00317*	*0.00020*	*0.00525*	*0.05277*	*0.01950*	*0.48282*	*0.10593*	*0.08504*	*0.01762*	*0.0129*
STA	0.1878	0.1003	0.1015	*	*0.26760*	*0.36759*	*0.37689*	*0.43332*	*0.39818*	*0.36432*	*0.28255*	*0.77982*	*0.74408*	*0.13811*	*0.14375*	*0.05009*	*0.00307*	*0.00000*	*0.00030*	*0.06039*	*0.00050*	*0.09445*	*0.00188*	*0.21394*	*0.10197*	*0.09682*	*0.02980*	*0.05564*	*0.0001*
LBKT	0.1805	0.1226	0.1324	0.0049	*	*0.67102*	*0.84110*	*0.30690*	*0.22869*	*0.85041*	*0.28007*	*0.13504*	*0.44134*	*0.31640*	*0.12197*	*0.01614*	*0.00733*	*0.00000*	*0.00040*	*0.00723*	*0.00149*	*0.03713*	*0.00188*	*0.28027*	*0.07118*	*0.11722*	*0.13632*	*0.07811*	*0.0006*
VIN	0.1528	0.0972	0.1049	0.0014	-0.0093	*	*0.61489*	*0.60430*	*0.16385*	*0.74616*	*0.31769*	*0.44916*	*0.47718*	*0.34779*	*0.22047*	*0.01812*	*0.00673*	*0.00020*	*0.00069*	*0.04425*	*0.00485*	*0.03524*	*0.00604*	*0.24948*	*0.12088*	*0.11147*	*0.08950*	*0.05663*	*0.0006*
SOP	0.1618	0.0958	0.1014	0.001	-0.012	-0.007	*	*0.72399*	*0.30482*	*0.73349*	*0.65637*	*0.33462*	*0.59766*	*0.86328*	*0.25641*	*0.08177*	*0.03623*	*0.00376*	*0.00020*	*0.00842*	*0.01406*	*0.24532*	*0.00089*	*0.24027*	*0.09583*	*0.07554*	*0.10781*	*0.19790*	*0.0025*
LGY	0.1509	0.0886	0.1007	-0.0008	0.0025	-0.0052	-0.0067	*	*0.21463*	*0.47322*	*0.29393*	*0.54371*	*0.69171*	*0.16721*	*0.19919*	*0.02168*	*0.00733*	*0.00000*	*0.00059*	*0.03881*	*0.00673*	*0.03475*	*0.00406*	*0.17602*	*0.08197*	*0.08969*	*0.04524*	*0.06079*	*0.001*
BL	0.2126	0.1109	0.1083	0.0002	0.0124	0.0223	0.006	0.0137	*	*0.15731*	*0.15197*	*0.28552*	*0.92011*	*0.08771*	*0.11831*	*0.33215*	*0.07613*	*0.07603*	*0.00000*	*0.00554*	*0.00337*	*0.20137*	*0.00020*	*0.22978*	*0.17870*	*0.41956*	*0.08890*	*0.79259*	*0.0018*
LBK	0.1806	0.1239	0.1335	0.0008	-0.0089	-0.0089	-0.007	-0.0012	0.0199	*	*0.16177*	*0.28106*	*0.37798*	*0.23572*	*0.07831*	*0.00366*	*0.00139*	*0.00000*	*0.00069*	*0.01544*	*0.00109*	*0.02515*	*0.00158*	*0.19077*	*0.04198*	*0.07227*	*0.05168*	*0.02505*	*0.0006*
RSC	0.1856	0.0773	0.063	0.0077	0.0073	0.0063	-0.0103	0.0062	0.0235	0.016	*	*0.14523*	*0.48936*	*0.28304*	*0.48381*	*0.16018*	*0.14434*	*0.14494*	*0.00812*	*0.01554*	*0.04495*	*0.40343*	*0.00089*	*0.04445*	*0.26631*	*0.03762*	*0.04237*	*0.16404*	*0.0031*
SCG	0.1837	0.0992	0.1119	-0.0104	0.0148	-0.0017	0.0031	-0.0033	0.009	0.0038	0.0189	*	*0.70072*	*0.22057*	*0.32155*	*0.13563*	*0.01069*	*0.00871*	*0.00000*	*0.16295*	*0.02584*	*0.19364*	*0.00287*	*0.20295*	*0.07376*	*0.05198*	*0.02089*	*0.06851*	*0.0003*
BAC	0.1464	0.0568	0.055	-0.013	-0.0016	-0.0041	-0.007	-0.0101	-0.0284	0.0006	-0.002	-0.0111	*	*0.23879*	*0.50817*	*0.88130*	*0.23493*	*0.31106*	*0.00149*	*0.09415*	*0.02168*	*0.35779*	*0.00307*	*0.59222*	*0.51549*	*0.67736*	*0.14434*	*0.92694*	*0.0011*
SMC	0.1975	0.1099	0.1225	0.0139	0.0042	0.0022	-0.014	0.0101	0.0338	0.0063	0.0083	0.0103	0.01	*	*0.11444*	*0.03386*	*0.00703*	*0.00248*	*0.00000*	*0.01188*	*0.00515*	*0.23918*	*0.00000*	*0.17097*	*0.04930*	*0.03435*	*0.04475*	*0.03297*	*0.0031*
BEC	0.072	0.0221	0.0448	0.0196	0.0234	0.0136	0.0092	0.0117	0.0324	0.0297	-0.0032	0.0049	-0.0046	0.0248	*	*0.63439*	*0.53084*	*0.40600*	*0.01901*	*0.07316*	*0.06980*	*0.38174*	*0.06217*	*0.27700*	*0.45332*	*0.21691*	*0.44708*	*0.31413*	*0.0213*
CWC	0.1034	0.0365	0.0516	0.019	0.0308	0.0331	0.0151	0.0207	0.0047	0.0347	0.0145	0.0117	-0.0164	0.025	-0.0089	*	*0.58865*	*0.14098*	*0.01406*	*0.01406*	*0.01267*	*0.62984*	*0.00693*	*0.41738*	*0.74696*	*0.34680*	*0.14464*	*0.87031*	*0.018*
BBC	0.0994	0.0445	0.0557	0.0473	0.0419	0.0438	0.0227	0.0333	0.0316	0.0497	0.0157	0.0376	0.009	0.0354	-0.0045	-0.0041	*	*0.27888*	*0.05712*	*0.00079*	*0.01238*	*0.18671*	*0.01871*	*0.10732*	*0.59727*	*0.10464*	*0.23493*	*0.51064*	*0.0913*
UC	0.0964	0.0337	0.0518	0.0416	0.0554	0.0479	0.029	0.0334	0.0264	0.0533	0.0135	0.0261	0.0039	0.0365	0.0007	0.0069	0.0032	*	*0.00792*	*0.00188*	*0.02049*	*0.17048*	*0.00030*	*0.02841*	*0.24374*	*0.08692*	*0.04336*	*0.16167*	*0.0034*
NPO	0.1696	0.021	0.0314	0.1099	0.1155	0.1052	0.0888	0.0897	0.1449	0.1136	0.0763	0.1183	0.0842	0.1017	0.0645	0.0467	0.0304	0.043	*	*0.00277*	*0.00000*	*0.00139*	*0.03099*	*0.00158*	*0.54737*	*0.00416*	*0.00871*	*0.00614*	*0.2397*
CAR	0.2298	0.0956	0.11	0.034	0.0734	0.0425	0.059	0.034	0.1093	0.054	0.0733	0.0171	0.0328	0.0577	0.0443	0.0482	0.077	0.049	0.1054	*	*0.00703*	*0.02386*	*0.00485*	*0.04970*	*0.06326*	*0.00743*	*0.00396*	*0.00129*	*0.0003*
NBQ	0.246	0.1389	0.1707	0.059	0.0554	0.0484	0.0294	0.0302	0.0888	0.0456	0.0321	0.0287	0.0391	0.0422	0.0279	0.0273	0.0273	0.0175	0.1319	0.0604	*	*0.02327*	*0.00010*	*0.00198*	*0.00693*	*0.00149*	*0.00020*	*0.00832*	*0.0004*
TRE	0.154	0.0723	0.0909	0.0197	0.0326	0.0344	0.0071	0.0242	0.0182	0.032	0.0014	0.0109	0.0028	0.009	0.0025	-0.0063	0.0098	0.008	0.089	0.0523	0.0336	*	*0.00109*	*0.24908*	*0.15127*	*0.05257*	*0.05247*	*0.24057*	*0.0165*
PWC	0.1479	0.0708	0.0629	0.0892	0.091	0.0727	0.0779	0.0669	0.1366	0.0859	0.0861	0.0814	0.0768	0.0935	0.0432	0.0537	0.0439	0.0685	0.0677	0.0912	0.1134	0.0951	*	*0.02089*	*0.27661*	*0.00980*	*0.02366*	*0.00218*	*0.1353*
FBC	0.1597	0.1244	0.1125	0.0185	0.0107	0.0142	0.0138	0.0212	0.0201	0.0203	0.0608	0.0212	-0.0122	0.0248	0.0158	0.0003	0.0334	0.0507	0.1721	0.0725	0.1355	0.0162	0.0923	*	*0.29651*	*0.43867*	*0.40075*	*0.34422*	*0.0226*
BAS	0.0954	0.0073	-0.0067	0.0325	0.0445	0.032	0.0301	0.035	0.0223	0.054	0.0116	0.0372	-0.0052	0.0481	-0.0017	-0.0175	-0.0096	0.0102	-0.0105	0.0519	0.0732	0.0248	0.0128	0.0127	*	*0.69696*	*0.29918*	*0.66736*	*0.2536*
BAK	0.1248	0.0878	0.0656	0.0457	0.0415	0.0433	0.0488	0.046	-0.0019	0.0584	0.0766	0.0632	-0.0204	0.0855	0.0267	0.005	0.0408	0.0385	0.1679	0.1399	0.1794	0.0684	0.1326	-0.0045	-0.0288	*	*0.47619*	*0.72072*	*0.0061*
ENL	0.0668	0.0919	0.0804	0.073	0.0361	0.0486	0.0367	0.0588	0.055	0.0635	0.0691	0.0827	0.0316	0.0712	-0.0048	0.0278	0.017	0.0478	0.144	0.1568	0.1785	0.0652	0.1068	-0.0019	0.0127	-0.011	*	*0.29294*	*0.0261*
YAM	0.1126	0.0553	0.0633	0.0248	0.0221	0.0296	0.0093	0.0215	-0.0196	0.032	0.0165	0.0227	-0.0216	0.0312	0.0052	-0.0131	-0.0035	0.0085	0.0644	0.0933	0.0418	0.0082	0.0805	0.0049	-0.0142	-0.0252	0.0099	*	*0.0145*
CAT	0.1531	0.0552	0.0798	0.0874	0.0775	0.0758	0.0486	0.0653	0.0955	0.079	0.0609	0.0793	0.0655	0.0559	0.0505	0.0351	0.0187	0.04	0.0146	0.1014	0.0681	0.0489	0.0242	0.0825	0.014	0.1362	0.0891	0.0431	*

Supplementary Table 8. F_{st} values and p values (italicised) for 29 prehistoric populations.

		HGCN	HGSW	HGE	STA	LBKT	VIN	SOP	LGY	BL	LBK	RSC	SCG	BAC	SMC	BEC	CWC	BBC	UC	NPO	CAR	NBQ	TRE	PWC	FBC	BAS	BAK	ENL	EBA-YA	EBA-Cat
HGCN	19	0	0.0529	0.1104	0.2312	0.2295	0.1803	0.1931	0.1778	0.2701	0.2204	0.2279	0.225	0.1716	0.2462	0.0776	0.1153	0.1103	0.1067	0.2042	0.2984	0.3262	0.182	0.1736	0.19	0.1055	0.1426	0.0715	0.1269	0.1808
HGSW	14	0.0529	0	0	0.1115	0.1498	0.1077	0.1059	0.0972	0.1248	0.1414	0.0838	0.1102	0.0603	0.1234	0.0226	0.0379	0.0466	0.0349	0.0214	0.1057	0.1613	0.078	0.0761	0.1421	0.0074	0.0962	0.1013	0.0585	0.0584
HGE	13	0.1104	0	0	0.1129	0.1611	0.1172	0.1129	0.112	0.1214	0.154	0.0672	0.126	0.0582	0.1396	0.0469	0.0544	0.0589	0.0547	0.0324	0.1236	0.2059	0.0999	0.0671	0.1268	0	0.0702	0.0874	0.0676	0.0868
STA	44	0.2312	0.1115	0.1129	0	0.0119	0.0014	0.001	0	0.0002	0.0008	0.0078	0	0	0.0141	0.02	0.0193	0.0497	0.0434	0.1235	0.0352	0.0627	0.0201	0.0979	0.0189	0.0336	0.0479	0.0787	0.0255	0.0958
LBKT	42	0.2295	0.1498	0.1611	0.0119	0	0	0	0.0086	0.0125	0	0.0122	0.0264	0.0028	0.006	0.0327	0.0374	0.0487	0.0647	0.1398	0.1033	0.0752	0.0382	0.1151	0.0131	0.0529	0.0417	0.0325	0.0228	0.09
VIN	31	0.1803	0.1077	0.1172	0.0014	0	0	0	0	0.0228	0	0.0063	0	0	0.0023	0.0137	0.0342	0.0458	0.0503	0.1176	0.0443	0.0508	0.0356	0.0784	0.0144	0.0331	0.0453	0.0511	0.0305	0.082
SOP	37	0.1931	0.1059	0.1129	0.001	0	0	0	0	0.0061	0	0	0.0031	0	0	0.0093	0.0153	0.0233	0.0299	0.0975	0.0627	0.0303	0.0071	0.0844	0.014	0.0311	0.0513	0.0381	0.0094	0.0511
LGY	82	0.1778	0.0972	0.112	0	0.0086	0	0	0	0.0139	0	0.0063	0	0	0.0102	0.0118	0.0211	0.0344	0.0345	0.0986	0.0352	0.0312	0.0248	0.0717	0.0216	0.0362	0.0482	0.0625	0.022	0.0699
BL	13	0.2701	0.1248	0.1214	0.0002	0.0125	0.0228	0.0061	0.0139	0	0.0203	0.024	0.0091	0	0.035	0.0335	0.0047	0.0326	0.0271	0.1694	0.1227	0.0975	0.0185	0.1581	0.0205	0.0228	0	0.0582	0	0.1056
LBK	108	0.2204	0.1414	0.154	0.0008	0	0	0	0	0.0203	0	0.0162	0.0038	0.0006	0.0063	0.0306	0.0359	0.0523	0.0562	0.1281	0.0571	0.0477	0.033	0.094	0.0207	0.0571	0.062	0.0678	0.0331	0.0858
RSC	17	0.2279	0.0838	0.0672	0.0078	0.0122	0.0063	0	0.0063	0.024	0.0162	0	0.0193	0	0.0084	0	0.0147	0.016	0.0136	0.0826	0.0791	0.0331	0.0014	0.0943	0.0647	0.0118	0.083	0.0742	0.0168	0.0649
SCG	33	0.225	0.1102	0.126	0	0.0264	0	0.0031	0	0.0091	0.0038	0.0193	0	0	0.0104	0.0049	0.0119	0.0391	0.0267	0.1342	0.0174	0.0295	0.0111	0.0886	0.0217	0.0386	0.0674	0.0901	0.0232	0.0862
BAC	19	0.1716	0.0603	0.0582	0	0.0028	0	0	0	0	0.0006	0	0	0	0.0101	0	0	0.009	0.004	0.092	0.0339	0.0407	0.0028	0.0831	0	0	0	0.0326	0	0.0701
SMC	29	0.2462	0.1234	0.1396	0.0141	0.006	0.0023	0	0.0102	0.035	0.0063	0.0084	0.0104	0.0101	0	0.0254	0.0256	0.0367	0.0378	0.1132	0.0612	0.0441	0.0091	0.1032	0.0254	0.0506	0.0935	0.0767	0.0322	0.0592
BEC	17	0.0776	0.0226	0.0469	0.02	0.0327	0.0137	0.0093	0.0118	0.0335	0.0306	0	0.0049	0	0.0254	0	0	0	0.0007	0.0689	0.0463	0.0287	0.0025	0.0452	0.0161	0	0.0274	0	0.0052	0.0532
CWC	44	0.1153	0.0379	0.0544	0.0193	0.0374	0.0342	0.0153	0.0211	0.0047	0.0359	0.0147	0.0119	0	0.0256	0	0	0	0.0069	0.049	0.0507	0.028	0	0.0567	0.0003	0	0.005	0.0286	0	0.0364
BBC	35	0.1103	0.0466	0.0589	0.0497	0.0487	0.0458	0.0233	0.0344	0.0326	0.0523	0.016	0.0391	0.009	0.0367	0	0	0	0.0032	0.0313	0.0834	0.028	0.0099	0.0459	0.0345	0	0.0426	0.0173	0	0.0191
UC	94	0.1067	0.0349	0.0547	0.0434	0.0647	0.0503	0.0299	0.0345	0.0271	0.0562	0.0136	0.0267	0.004	0.0378	0.0007	0.0069	0.0032	0	0.0449	0.0515	0.0178	0.0081	0.0735	0.0534	0.0103	0.0401	0.0502	0.0086	0.0416
NPO	18	0.2042	0.0214	0.0324	0.1235	0.1398	0.1176	0.0975	0.0986	0.1694	0.1281	0.0826	0.1342	0.092	0.1132	0.0689	0.049	0.0313	0.0449	0	0.1178	0.1519	0.0977	0.0726	0.2078	0	0.2018	0.1683	0.0688	0.0148
CAR	20	0.2984	0.1057	0.1236	0.0352	0.1033	0.0443	0.0627	0.0352	0.1227	0.0571	0.0791	0.0174	0.0339	0.0612	0.0463	0.0507	0.0834	0.0515	0.1178	0	0.0643	0.0551	0.1003	0.0782	0.0548	0.1626	0.186	0.1029	0.1129
NBQ	43	0.3262	0.1613	0.2059	0.0627	0.0752	0.0508	0.0303	0.0312	0.0975	0.0477	0.0331	0.0295	0.0407	0.0441	0.0287	0.028	0.028	0.0178	0.1519	0.0643	0	0.0348	0.1279	0.1568	0.0789	0.2186	0.2172	0.0436	0.073
TRE	29	0.182	0.078	0.0999	0.0201	0.0382	0.0356	0.0071	0.0248	0.0185	0.033	0.0014	0.0111	0.0028	0.0091	0.0025	0	0.0099	0.0081	0.0977	0.0551	0.0348	0	0.1051	0.0165	0.0255	0.0735	0.0698	0.0083	0.0514
PWC	19	0.1736	0.0761	0.0671	0.0979	0.1151	0.0784	0.0844	0.0717	0.1581	0.094	0.0943	0.0886	0.0831	0.1032	0.0452	0.0567	0.0459	0.0735	0.0726	0.1003	0.1279	0.1051	0	0.1017	0.0129	0.1529	0.1196	0.0875	0.0248
FBC	10	0.19	0.1421	0.1268	0.0189	0.0131	0.0144	0.014	0.0216	0.0205	0.0207	0.0647	0.0217	0	0.0254	0.0161	0.0003	0.0345	0.0534	0.2078	0.0782	0.1568	0.0165	0.1017	0	0.0128	0	0	0.005	0.09
BAS	11	0.1055	0.0074	0	0.0336	0.0529	0.0331	0.0311	0.0362	0.0228	0.0571	0.0118	0.0386	0	0.0506	0	0	0	0.0103	0	0.0548	0.0789	0.0255	0.0129	0.0128	0	0	0.0129	0	0.0142
BAK	8	0.1426	0.0962	0.0702	0.0479	0.0417	0.0453	0.0513	0.0482	0	0.062	0.083	0.0674	0	0.0935	0.0274	0.005	0.0426	0.0401	0.2018	0.1626	0.2186	0.0735	0.1529	0	0	0	0	0	0.1577
ENL	9	0.0715	0.1013	0.0874	0.0787	0.0325	0.0511	0.0381	0.0625	0.0582	0.0678	0.0742	0.0901	0.0326	0.0767	0	0.0286	0.0173	0.0502	0.1683	0.186	0.2172	0.0698	0.1196	0	0.0129	0	0	0.01	0.0978
YAM	26	0.1269	0.0585	0.0676	0.0255	0.0228	0.0305	0.0094	0.022	0	0.0331	0.0168	0.0232	0	0.0322	0.0052	0	0	0.0086	0.0688	0.1029	0.0436	0.0083	0.0875	0.005	0	0	0.01	0	0.045
CAT	24	0.1808	0.0584	0.0868	0.0958	0.09	0.082	0.0511	0.0699	0.1056	0.0858	0.0649	0.0862	0.0701	0.0592	0.0532	0.0364	0.0191	0.0416	0.0148	0.1129	0.073	0.0514	0.0248	0.09	0.0142	0.1577	0.0978	0.045	0

Supplementary Table 9. Slatkin matrix for the MDS with 29 prehistoric populations.

Supplementary Table 10. Haplogroup frequencies of 20 prehistoric populations.

For culture/population information, see Supplementary Table 3, 6.

Culture	n	H	H5	HV	HV0	V	I	J	K	N	N1a	R	T1	T2	U	U2	U3	U4	U5	U5a	U5b	U8	W	X
HGCN	19	0	0	0	0	0	0	0	0	0	0	0	0	0	10.5	5.26	0	10.5	0	21.1	52.6	0	0	0
HGCN*	18	0	0	0	0	0	0	0	0	0	0	0	0	0	11.1	5.56	0	11.1	0	22.2	50	0	0	0
HGSW	14	42.86	0	0	0	0	0	0	0	14.3	0	0	0	0	0	0	0	7.14	0	0	35.7	0	0	0
HGSW*	13	46.15	0	0	0	0	0	0	0	15.4	0	0	0	0	0	0	0	7.69	0	0	30.8	0	0	0
STA	44	4.545	2.27	0	2.27	6.82	0	11.4	27.3	0	6.82	0	2.27	20.5	0	0	2.27	2.27	0	0	0	0	4.55	6.82
STA*	40	2.5	2.5	0	2.5	7.5	0	10	27.5	0	7.5	0	2.5	20	0	0	2.5	2.5	0	0	0	0	5	7.5
LBKT	42	23.81	4.76	2.38	0	2.38	0	7.14	16.7	0	9.52	0	2.38	23.8	0	4.76	0	0	0	2.38	0	0	0	0
LBKT*	37	24.32	5.41	2.7	0	2.7	0	5.41	16.2	0	10.8	0	2.7	21.6	0	5.41	0	0	0	2.7	0	0	0	0
VIN	31	3.226	0	0	3.23	0	0	9.68	29	0	12.9	0	0	19.4	0	3.23	9.68	0	0	6.45	3.23	0	0	0
VIN*	28	3.571	0	0	3.57	0	0	10.7	28.6	0	14.3	0	0	14.3	0	3.57	10.7	0	0	7.14	3.57	0	0	0
SOP	37	13.51	13.5	2.7	8.11	0	0	13.5	13.5	0	10.8	0	0	13.5	0	0	2.7	0	0	2.7	0	2.7	0	2.7
SOP*	32	12.5	9.38	3.13	9.38	0	0	12.5	15.6	0	9.38	0	0	15.6	0	0	3.13	0	0	3.13	0	3.13	0	3.13
LGY	82	17.07	4.88	3.66	3.66	0	0	7.32	18.3	0	12.2	0	0	18.3	0	0	0	0	1.22	1.22	3.66	6.1	1.22	1.22
LGY*	65	18.46	4.62	4.62	3.08	0	0	6.15	16.9	0	15.4	0	0	15.4	0	0	0	0	1.54	1.54	3.08	6.15	1.54	1.54
BL	13	23.08	0	7.69	0	0	0	7.69	7.69	0	0	0	7.69	23.1	0	0	0	0	0	0	0	7.69	15.4	0
BL*	11	18.18	0	9.09	0	0	0	9.09	9.09	0	0	0	9.09	27.3	0	0	0	0	0	0	0	9.09	9.09	0
LBK	108	14.81	1.85	4.63	0	4.63	0	12	20.4	0	12	0	0	22.2	0	0	0.93	0	0	1.85	0.93	0	2.78	0.93
LBK*	86	16.28	2.33	3.49	0	2.33	0	10.5	20.9	0	14	0	0	22.1	0	0	1.16	0	0	2.33	1.16	0	2.33	1.16
RSC	17	17.65	11.8	0	23.5	5.88	0	0	11.8	0	5.88	0	0	11.8	0	0	0	0	0	0	5.88	0	0	5.88
RSC*	15	20	13.3	0	13.3	6.67	0	0	13.3	0	6.67	0	0	13.3	0	0	0	0	0	0	6.67	0	0	6.67
SCG	33	15.15	0	3.03	0	0	0	15.2	30.3	0	3.03	0	0	12.1	0	0	0	0	0	0	6.06	3.03	9.09	3.03

SCG*	25	16	0	4	0	0	0	12	28	0	4	0	0	12	0	0	0	0	0	0	8	4	8	4
BAC	19	26.32	0	5.26	0	0	0	5.26	10.5	0	5.26	0	5.26	21.1	0	0	0	0	0	0	5.26	5.26	0	10.5
SMC	29	13.79	17.2	3.45	0	3.45	0	20.7	10.3	0	6.9	0	0	6.9	0	0	10.3	0	0	0	3.45	0	0	3.45
SMC*	20	15	5	5	0	5	0	20	15	0	10	0	0	10	0	0	5	0	0	0	5	0	0	5
BEC	17	17.65	5.88	0	0	5.88	0	0	17.6	0	0	0	0	11.8	0	0	0	0	0	11.8	17.6	0	5.88	5.88
BEC*	13	15.38	7.69	0	0	7.69	0	0	15.4	0	0	0	0	7.69	0	0	0	0	0	7.69	23.1	0	7.69	7.69
CWC	44	20.45	2.27	0	2.27	0	2.27	9.09	13.6	0	0	0	6.82	11.4	0	2.27	0	6.82	0	9.09	4.55	0	2.27	6.82
CWC*	36	25	2.78	0	2.78	0	2.78	5.56	11.1	0	0	0	5.56	11.1	0	2.78	0	8.33	0	8.33	5.56	0	2.78	5.56
BBC	35	31.43	8.57	0	0	0	2.86	2.86	5.71	0	0	0	5.71	5.71	0	2.86	0	5.71	0	14.3	5.71	0	8.57	0
BBC*	29	27.59	6.9	0	0	0	3.45	3.45	6.9	0	0	0	6.9	6.9	0	3.45	0	3.45	0	13.8	6.9	0	10.3	0
UC	94	18.09	3.19	2.13	0	3.19	12.8	6.38	7.45	0	0	1.06	2.13	6.38	2.13	6.38	0	1.06	0	12.8	2.13	3.19	4.26	5.32
UC*	77	18.18	2.6	2.6	0	2.6	9.09	7.79	9.09	0	0	1.3	2.6	7.79	2.6	7.79	0	1.3	0	11.7	2.6	2.6	3.9	3.9
NPO	18	66.67	0	0	0	5.56	0	0	0	0	0	0	0	0	5.56	0	0	0	0	16.7	5.56	0	0	0
NPO*	12	50	0	0	0	8.33	0	0	0	0	0	0	0	0	8.33	0	0	0	0	25	8.33	0	0	0
CAR	20	25	0	0	0	0	0	0	30	20	0	0	0	10	0	0	0	0	0	0	10	0	0	5
CAR*	16	31.25	0	0	0	0	0	0	25	18.8	0	0	0	6.25	0	0	0	0	0	0	12.5	0	0	6.25
NBQ	43	44.19	0	2.33	0	0	2.33	4.65	9.3	0	0	0	0	2.33	25.6	0	0	0	0	0	6.98	0	0	2.33
NBQ*	25	32	0	4	0	0	4	8	8	0	0	0	0	4	24	0	0	0	0	0	12	0	0	4

abbr.	n	HGCN	HGCN*	HGSW	HGSW*	STA	STA*	LBKT	LBKT*	VIN	VIN*	SOP	SOP*	LGY	LGY*	BL	BL*	LBKC	LBKC*	RSC	RSC*	SCG	SCG*	BAC	SMC	SMC*	BEC	BEC*	CWC	CWC*	BBC	BBC*	UC	UC*	NPO	NPO*	CAR	CAR*	NBQ	NBQ*
HGCN	19	0	0	0.0504	0.0753	0.2291	0.2259	0.2141	0.2012	0.1773	0.172	0.1896	0.1786	0.1763	0.1758	0.2549	0.2703	0.2173	0.2054	0.2202	0.2006	0.2163	0.1871	0.1643	0.2417	0.2137	0.0751	0.0684	0.111	0.1011	0.1054	0.0971	0.0998	0.0991	0.1908	0.1042	0.2761	0.2645	0.3178	0.22557
HGCN*	18	0	0	0.0404	0.063	0.2144	0.211	0.2009	0.188	0.1647	0.1591	0.1765	0.1655	0.1651	0.1642	0.2332	0.2468	0.2059	0.194	0.2014	0.1818	0.2015	0.1722	0.1494	0.2262	0.1979	0.0649	0.0576	0.0997	0.0896	0.0921	0.0842	0.0894	0.0883	0.1705	0.0875	0.2588	0.247	0.2971	0.20791
HGSW	14	0.0504	0.0404	0	0	0.1115	0.105	0.1371	0.1231	0.1067	0.0974	0.1054	0.0975	0.0966	0.0931	0.1228	0.1371	0.1395	0.1292	0.0838	0.0659	0.1081	0.0816	0.0592	0.1232	0.0991	0.0226	0	0.0364	0.0228	0.046	0.0356	0.0342	0.0297	0.0225	0.0175	0.0967	0.0782	0.1613	0.10278
HGSW*	13	0.0753	0.063	0	0	0.1002	0.0932	0.1312	0.1172	0.1003	0.091	0.0995	0.0916	0.0902	0.0864	0.11	0.1231	0.1329	0.1227	0.0744	0.0563	0.0999	0.0735	0.0513	0.1154	0.0905	0.0241	0	0.0327	0.0194	0.043	0.0323	0.0294	0.0244	0.0128	0.0159	0.0853	0.0674	0.1566	0.09894
STA	44	0.2291	0.2144	0.1115	0.1002	0	0	0.0048	0.0043	0.0014	0.0061	0.0011	0	0	0.0039	0.0011	0	0.0007	0.002	0.008	0	0	0	0	0.0142	0	0.02	0.0168	0.0193	0.0254	0.0517	0.0386	0.0446	0.033	0.1247	0.1024	0.0259	0.0383	0.0634	0.03638
STA*	40	0.2259	0.211	0.105	0.0932	0	0	0.0073	0.0059	0.0007	0.0043	0.0025	0	0	0.0024	0.0019	0	0.002	0.0027	0.0066	0	0	0	0	0.0165	0	0.0192	0.0148	0.0195	0.0252	0.0517	0.0381	0.043	0.0322	0.1222	0.1002	0.0242	0.0353	0.0657	0.03756
LBKT	42	0.2141	0.2009	0.1371	0.1312	0.0048	0.0073	0	0	0	0	0	0	0.0024	0.003	0.0119	0.0107	0	0	0.007	0	0.0145	0.01	0	0.0042	0	0.0232	0.0255	0.0308	0.0327	0.044	0.0357	0.0583	0.0467	0.128	0.0918	0.0659	0.0786	0.0583	0.03524
LBKT*	37	0.2012	0.188	0.1231	0.1172	0.0043	0.0059	0	0	0	0	0	0	0	0	0.0103	0.0112	0	0	0	0	0.012	0.006	0	0.0021	0	0.0178	0.0181	0.0266	0.0269	0.036	0.0282	0.0498	0.0398	0.1165	0.0798	0.0616	0.0724	0.0487	0.02722
VIN	31	0.1773	0.1647	0.1067	0.1003	0.0014	0.0007	0	0	0	0	0	0	0	0	0.0231	0.0225	0	0	0.0064	0	0	0	0	0.0025	0	0.0135	0.0104	0.0333	0.0363	0.0459	0.0357	0.0502	0.0421	0.1163	0.0758	0.0358	0.0434	0.0516	0.02514
VIN*	28	0.172	0.1591	0.0974	0.091	0.0061	0.0043	0	0	0	0	0	0	0	0	0.0282	0.0307	0	0	0.0068	0	0	0	0	0.0035	0	0.0144	0.0087	0.0319	0.0339	0.0416	0.0317	0.0423	0.0364	0.1106	0.0688	0.028	0.0334	0.0446	0.01919
SOP	37	0.1896	0.1765	0.1054	0.0995	0.0011	0.0025	0	0	0	0	0	0	0	0	0.0061	0.0113	0	0	0	0	0.0029	0	0	0	0	0.0091	0.0046	0.0146	0.0144	0.024	0.0172	0.0306	0.0233	0.0965	0.0611	0.0482	0.0548	0.0301	0.01291
SOP*	32	0.1786	0.1655	0.0975	0.0916	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0021	0	0.008	0.0082	0.0202	0.0123	0.0263	0.0182	0.0926	0.0561	0.0439	0.0511	0.0252	0.00698
LGY	82	0.1763	0.1651	0.0966	0.0902	0	0	0.0024	0	0	0	0	0	0	0	0.0143	0.0172	0	0	0.0063	0	0	0	0	0.0103	0	0.0117	0.0084	0.0211	0.0226	0.0373	0.0285	0.0365	0.03	0.1	0.0729	0.0253	0.0319	0.0316	0.01669
LGY*	65	0.1758	0.1642	0.0931	0.0864	0.0039	0.0024	0.003	0	0	0	0	0	0	0	0.022	0.0268	0	0	0.0025	0	0.0036	0	0	0.0086	0	0.014	0.0084	0.0271	0.0271	0.0384	0.0301	0.0351	0.0309	0.0985	0.0704	0.0318	0.0367	0.0342	0.0188
BL	13	0.2549	0.2332	0.1228	0.11	0.0011	0.0019	0.0119	0.0103	0.0231	0.0282	0.0061	0	0.0143	0.022	0	0	0.0204	0.0177	0.0242	0.0048	0.0077	0	0	0.0357	0.0101	0.0333	0.0256	0.005	0.0081	0.0333	0.0135	0.0278	0.0159	0.1627	0.1253	0.1092	0.1203	0.0986	0.04616
BL*	11	0.2703	0.2468	0.1371	0.1231	0	0	0.0107	0.0112	0.0225	0.0307	0.0113	0	0.0172	0.0268	0	0	0.0218	0.0199	0.0348	0.0129	0.011	0.0015	0	0.0432	0.0133	0.0415	0.0367	0.0091	0.0158	0.0491	0.0259	0.0378	0.0226	0.1903	0.1489	0.1252	0.139	0.1319	0.06419
LBK	108	0.2173	0.2059	0.1395	0.1329	0.0007	0.002	0	0	0	0	0	0	0	0	0.0204	0.0218	0	0	0.0158	0.0015	0.004	0.0011	0.0005	0.0062	0	0.0299	0.031	0.0353	0.0392	0.0539	0.0452	0.0572	0.0485	0.1278	0.0983	0.0434	0.0557	0.0474	0.03065
LBK*	86	0.2054	0.194	0.1292	0.1227	0.002	0.0027	0	0	0	0	0	0	0	0	0.0177	0.0199	0	0	0.0127	0	0.0048	0.0004	0	0.0055	0	0.0275	0.0268	0.0341	0.0368	0.0493	0.0403	0.0523	0.0446	0.1209	0.0908	0.0475	0.0583	0.0452	0.02808
RSC	17	0.2202	0.2014	0.0838	0.0744	0.008	0.0066	0.007	0	0.0064	0.0068	0	0	0.0063	0.0025	0.0242	0.0348	0.0158	0.0127	0	0	0.0192	0.0062	0	0.0085	0	0	0	0.014	0.0081	0.0177	0.0116	0.0148	0.0091	0.0821	0.0502	0.0652	0.0714	0.0334	0.01488
RSC*	15	0.2006	0.1818	0.0659	0.0563	0	0	0	0	0	0	0	0	0	0	0.0048	0.0129	0.0015	0	0	0	0.0024	0	0	0	0	0.0013	0	0.0053	0	0.0029	0	0.0728	0.0431	0.049	0.0558	0.0195	0.00047
SCG	33	0.2163	0.2015	0.1081	0.0999	0	0	0.0145	0.012	0	0	0.0029	0	0	0.0036	0.0077	0.011	0.004	0.0048	0.0192	0.0024	0	0	0	0.0109	0	0.0049	0.001	0.0121	0.0172	0.041	0.029	0.0287	0.02	0.1321	0.0976	0.0021	0.0138	0.0304	0.00728
SCG*	25	0.1871	0.1722	0.0816	0.0735	0	0	0.01	0.006	0	0	0	0	0	0	0	0.0015	0.0011	0.0004	0.0062	0	0	0	0	0.0042	0	0	0	0.0025	0.0057	0.0274	0.0155	0.0147	0.0076	0.112	0.077	0.0066	0.0149	0.0198	0
BAC	19	0.1643	0.1494	0.0592	0.0513	0	0	0	0	0	0	0	0	0	0	0	0	0.0005	0	0	0	0	0	0	0.0103	0	0	0	0	0	0.0117	0	0.0065	0	0.0908	0.0614	0.0259	0.0274	0.0411	0.00924
SMC	29	0.2417	0.2262	0.1232	0.1154	0.0142	0.0165	0.0042	0.0021	0.0025	0.0035	0	0	0.0103	0.0086	0.0357	0.0432	0.0062	0.0055	0.0085	0	0.0109	0.0042	0.0103	0	0	0.0254	0.0205	0.0253	0.0276	0.0391	0.0335	0.039	0.0307	0.114	0.0804	0.0413	0.0475	0.0444	0.02297
SMC*	20	0.2137	0.1979	0.0991	0.0905	0	0	0	0	0	0	0	0	0	0	0.0101	0.0133	0	0	0	0	0	0	0	0	0	0.0111	0.0044	0.0125	0.0153	0.0307	0.0217	0.0233	0.0167	0.1011	0.0653	0.0262	0.0315	0.031	0.00597
BEC	17	0.0751	0.0649	0.0226	0.0241	0.02	0.0192	0.0232	0.0178	0.0135	0.0144	0.0091	0.0021	0.0117	0.014	0.0333	0.0415	0.0299	0.0275	0	0	0.0049	0	0	0.0254	0.0111	0	0	0	0	0	0	0.0023	0	0.0687	0.0212	0.0369	0.0408	0.0293	0.00238
BEC*	13	0.0684	0.0576	0	0	0.0168	0.0148	0.0255	0.0181	0.0104	0.0087	0.0046	0	0.0084	0.0084	0.0256	0.0367	0.031	0.0268	0	0	0.001	0	0	0.0205	0.0044	0	0	0	0	0	0	0	0	0.054	0.0106	0.0283	0.0287	0.0206	0
CWC	44	0.111	0.0997	0.0364	0.0327	0.0193	0.0195	0.0308	0.0266	0.0333	0.0319	0.0146	0.008	0.0211	0.0271	0.005	0.0091	0.0353	0.0341	0.014	0.0013	0.0121	0.0025	0	0.0253	0.0125	0	0	0	0	0	0	0.008	0	0.047	0.022	0.0356	0.0358	0.0279	0.00689
CWC*	36	0.1011	0.0896	0.0228	0.0194	0.0254	0.0252	0.0327	0.0269	0.0363	0.0339	0.0144	0.0082	0.0226	0.0271	0.0081	0.0158	0.0392	0.0368	0.0081	0	0.0172	0.0057	0	0.0276	0.0153	0	0	0	0	0	0	0.0019	0	0.03	0.0061	0.0391	0.0376	0.0222	0.00351
BBC	35	0.1054	0.0921	0.046	0.043	0.0517	0.0517	0.044	0.036	0.0459	0.0416	0.024	0.0202	0.0373	0.0384	0.0333	0.0491	0.0539	0.0493	0.0177	0.0053	0.041	0.0274	0.0117	0.0391	0.0307	0	0	0	0	0	0	0.0053	0	0.026	0	0.0649	0.0621	0.0323	0.01594
BBC*	29	0.0971	0.0842	0.0356	0.0323	0.0386	0.0381	0.0357	0.0282	0.0357	0.0317	0.0172	0.0123	0.0285	0.0301	0.0135	0.0259	0.0452	0.0403	0.0116	0	0.029	0.0155	0	0.0335	0.0217	0	0	0	0	0	0	0	0	0.028	0	0.0524	0.0497	0.0301	0.01009
UC	94	0.0998	0.0894	0.0342	0.0294	0.0446	0.043	0.0583	0.0498	0.0502	0.0423	0.0306	0.0263	0.0365	0.0351	0.0278	0.0378	0.0572	0.0523	0.0148	0.0029	0.0287	0.0147	0.0065	0.039	0.0233	0.0023	0	0.008	0.0019	0.0053	0	0	0	0.0436	0.0172	0.0346	0.0316	0.0204	0.00571
UC*	77	0.0991	0.0883	0.0297	0.0244	0.033	0.0322	0.0467	0.0398	0.0421	0.0364	0.0233	0.0182	0.03	0.0309	0.0159	0.0226	0.0485	0.0446	0.0091	0	0.02	0.0076	0	0.0307	0.0167	0	0	0	0	0	0	0	0	0.0355	0.0117	0.0268	0.0249	0.0159	0.00092
NPO	18	0.1908	0.1705	0.0225	0.0128	0.1247	0.1222	0.128	0.1165	0.1163	0.1106	0.0965	0.0926	0.1	0.0985	0.1627	0.1903	0.1278	0.1209	0.0821	0.0728	0.1321	0.112	0.0908	0.114	0.1011	0.0687	0.054	0.047	0.03	0.026	0.028	0.0436	0.0355	0	0	0.0959	0.0822	0.1523	0.12489
NPO*	12	0.1042	0.0875	0.0175	0.0159	0.1024	0.1002	0.0918	0.0798	0.0758	0.0688	0.0611	0.0561	0.0729	0.0704	0.1253	0.1489	0.0983	0.0908	0.0502	0.0431	0.0976	0.077	0.0614	0.0804	0.0653	0.0212	0.0106	0.022	0.0061	0	0	0.0172	0.0117	0	0	0.083	0.0711	0.11	0.08045
CAR	20	0.2761	0.2588	0.0967	0.0853	0.0259	0.0242	0.0659	0.0616	0.0358	0.028	0.0482	0.0439	0.0253	0.0318	0.1092	0.1252	0.0434	0.0475	0.0652	0.049	0.0021	0.0066	0.0259	0.0413	0.0262	0.0369	0.0283	0.0356	0.0391	0.0649	0.0524	0.0346	0.0268	0.0959	0.083	0	0	0.0204	0
CAR*	16	0.2645	0.247	0.0782	0.0674	0.0383	0.0353	0.0786	0.0724	0.0434	0.0334	0.0548	0.0511	0.0319	0.0367	0.1203	0.139	0.0557	0.0583	0.0714	0.0558	0.0138	0.0149	0.0274	0.0475	0.0315	0.0408	0.0287	0.0358	0.0376	0.0621	0.0497	0.0316	0.0249	0.0822	0.0711	0	0	0.0259	0
NBQ	43	0.3178	0.2971	0.1613	0.1566	0.0634	0.0657	0.0583	0.0487	0.0516	0.0446	0.0301	0.0252	0.0316	0.0342	0.0986	0.1319	0.0474	0.0452	0.0334	0.0195	0.0304	0.0198	0.0411	0.0444	0.031	0.0293	0.0206	0.0279	0.0222	0.0323	0.0301	0.0204	0.0159	0.1523	0.11	0.0204	0.0259	0	0
NBQ*	25	0.2256	0.2079	0.1028	0.0989	0.0364	0.0376	0.0352	0.0272	0.0251	0.0192	0.0129	0.007	0.0167	0.0188	0.0462	0.0642	0.0307	0.0281	0.0149	0.0005	0.0073	0	0.0092	0.023	0.006	0.0024	0	0.0069	0.0035	0.0159	0.0101	0.0057	0.0009	0.1249	0.0805	0	0	0	0

Supplementary Table 11. Slatkin matrix for the MDS with 20 prehistoric populations (38 datasets).

Supplementary Table 12. Results of the AMOVA analysis with 15 prehistoric populations (15POP).

	among groups					within groups, among populations				
Arrangements	**sum of squares**	**Va**	**% of variation**	**Fct**	**p**	**sum of squares**	**Vb**	**% of variation**	**Fsc**	**p**
(STA_LBKT_VIN_LBK_SOP_LGY_BL_RSC_SCG_BA	27.	0.09	3.	0.03	0.00050+-	37.	0.00	0.	0.00	0.37040+-
C_SMC)+(BEC_UC_BBC_CWC)	704	226	21	207	0.00026	73	301	1	108	0.00447
(STA_LBKT_VIN_LBK_SOP_LGY_BL_RSC_SCG_BA	25.	0.08	3.	0.03	0.00307+-	39.	0.00	0.	0.00	0.24287+-
C_SMC_BEC)+(UC_BBC_CWC)	724	881	09	087	0.00054	71	686	24	246	0.00398
(STA_LBKT_VIN_LBK_SOP_LGY_RSC_SCG_BAC_	27.	0.08	3.	0.03	0.00079+-	37.	0.00	0.	0.00	0.33733+-
SMC)+(BL_BEC_UC_BBC_CWC)	44	786	06	058	0.00027	995	351	12	126	0.00460
(STA_LBKT_VIN_LBK_SOP_LGY_BL_SCG_BAC_	27.	0.08	3.	0.03	0.00050+-	37.	0.00	0.	0.00	0.35554+-
SMC)+(RSC_BEC_UC_BBC_CWC)	711	803	06	064	0.00022	723	299	1	107	0.00501
(STA_LBKT_VIN_LBK_SOP_LGY_BL_RSC_SCG_	27.	0.08	2.	0.02	0.00079+-	38.	0.00	0.	0.00	0.31990+-
SMC)+(BAC_BEC_UC_BBC_CWC)	281	589	99	991	0.00030	153	381	13	137	0.00478
(STA_LBKT_VIN_LBK_SOP_LGY_RSC_SCG_BAC_	25.	0.08	2.	0.02	0.00168+-	39.	0.00	0.	0.00	0.23347+-
SMC_BEC)+(BL_UC_BBC_CWC)	548	428	93	933	0.00040	886	716	25	257	0.00435
(STA_LBKT_VIN_LBK_SOP_LGY_BL_SCG_SMC)+	27.	0.08	2.	0.02	0.00069+-	38.	0.00	0.	0.00	0.32267+-
(RSC_BAC_BEC_UC_BBC_CWC)	413	317	9	9	0.00029	021	355	12	127	0.00504
(STA_LBKT_VIN_LBK_LGY_BL_RSC_SCG_BAC_	22.	0.06	2.	0.02	0.00554+-	42.	0.01	0.	0.00	0.09921+-
SMC)+(SOP_BEC_UC_BBC_CWC)	683	504	27	275	0.00075	751	26	44	451	0.00288
(STA_LBKT_VIN_LBK_SOP_BL_RSC_SCG_BAC_	18.	0.04	1.	0.01	0.02149+-	47.	0.02	0.	0.00	0.02495+-
SMC)+(LGY_BL_BEC_UC_BBC_CWC)	08	387	54	541	0.00136	354	147	75	766	0.00134
(STA_LBKT_VIN_LBK_SOP_BL_RSC_SCG_BAC_	17.	0.04	1.	0.01	0.02356+-	47.	0.02	0.	0.00	0.02426+-
SMC)+(LGY_BEC_UC_BBC_CWC)	901	364	53	533	0.00147	534	186	77	78	0.00141
(STA_LBKT_VIN_LBK_SOP_LGY_BL_RSC_SCG_BA	27.	0.09	3.	0.03	0.00050+-	37.	0.00	0.	0.00	0.37040+-
C_SMC)+(BEC_UC_BBC_CWC)	704	226	21	207	0.00026	73	301	1	108	0.00447

Summary of shared haplotypes (%)

	HGCN 19	STA 44	VIN 31	LBKT 42	LBK 108	SOP 37	LGY 82	RSC 17	BL 13	SCG 33	BAC 19	SMC 29	BEC 17	CWC 44	BBC 35	UC 94
HGCN	100	2.273	6.452	0	1.852	0	2.439	5.882	0	0	0	0	11.76	6.818	2.8571	4.2553
STA		100	45.16	61.9	55.56	54.05	51.22	70.59	38.46	60.61	36.84	62.07	41.18	43.18	42.857	37.234
VIN			100	59.52	52.78	40.54	42.68	52.94	7.692	45.45	21.05	48.28	35.29	29.55	22.857	20.213
LBKT				100	62.04	56.76	50	47.06	23.08	54.55	36.84	62.07	29.41	31.82	37.143	27.66
LBK					100	54.05	67.07	47.06	38.46	72.73	52.63	51.72	47.06	43.18	34.286	36.17
SOP						100	59.76	64.71	7.692	54.55	36.84	65.52	35.29	38.64	34.286	27.66
LGY							100	76.47	38.46	75.76	47.37	65.52	52.94	50	42.857	32.979
RSC								100	7.692	24.24	15.79	31.03	17.65	13.64	31.429	14.894
BL									100	27.27	21.05	17.24	17.65	15.91	28.571	14.894
SCG										100	36.84	51.72	35.29	29.55	28.571	26.596
BAC											100	24.14	17.65	22.73	22.857	18.085
SMC												100	35.29	31.82	34.286	25.532
BEC													100	20.45	34.286	23.404
CWC														100	60	38.298
BBC															100	23.404
UC																100

Supplementary Table 13a. Shared haplotype analysis with 16 prehistoric populations.

Summary of ancestral shared haplotypes (%)

		HGCN	STA	VIN	LBKT	LBK	SOP	LGY	RSC	BL	SCG	BAC	SMC	BEC	CWC	BBC	UC
HGCN	19	100															
STA	44	2.27	97.73														
VIN	31	6.45	48.39	45.161													
LBKT	42	0	61.9	9.5238	28.6												
LBK	108	1.85	55.56	6.4815	10.2	25.9											
SOP	37	0	54.05	5.4054	10.8	5.41	24.324										
LGY	82	2.44	51.22	3.6585	9.76	12.2	1.2195	19.5									
RSC	17	5.88	70.59	0	0	11.8	0	0	11.76								
BL	13	0	38.46	0	7.69	0	0	0	0	53.85							
SCG	33	0	60.61	0	9.09	3.03	0	3.03	0	3.03	21.21						
BAC	19	0	36.84	0	10.5	10.5	0	0	5.263	0	0	36.84					
SMC	29	0	62.07	10.345	3.45	3.45	3.4483	0	0	3.448	3.448	0	10.34				
BEC	17	11.8	41.18	0	0	0	0	0	0	0	0	0	0	47.06			
CWC	44	6.82	43.18	0	2.27	0	6.8182	0	0	0	0	0	0	0	40.91		
BBC	35	2.86	42.86	0	0	0	2.8571	0	0	0	0	0	0	0	20	31.43	
UC	94	4.26	36.17	0	3.19	0	0	0	0	2.128	0	0	0	0	7.447	6.383	40.43

Supplementary Table 13b. Ancestral shared haplotype analysis with 16 prehistoric populations.

Supplementary Table 14a-b. F_{st} and p values for the Transdanubian populations, LBK, and hunter-gatherers.

A

F_{st} **values**		HGCN	STA	LBKT	VIN	SOP	LGY	BL	LBKC
HGCN	19	0							
STA	44	0.18636	0						
LBKT	42	0.17629	0.00481	0					
VIN	31	0.15036	0.0013	-0.00936	0				
SOP	37	0.15937	0.00104	-0.01227	-0.00696	0			
LGY	82	0.14983	-0.00078	0.00241	-0.00513	-0.00673	0		
BL	13	0.20304	0.00111	0.01173	0.02247	0.00601	0.01411	0	
LBK	108	0.17847	0.00073	-0.00888	-0.00881	-0.00728	-0.0012	0.01998	0

B

P values

	HGCN	STA	LBKT	VIN	SOP	LGY	BL	LBK
HGCN	*	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
STA	*0.00000*	*	0.14573+-0.0027	0.35739+-0.0043	0.37373+-0.0044	0.43392+-0.0050	0.38343+-0.0045	0.35561+-0.0050
LBKT	*0.00000*	*0,4122618*	*	0.46857+-0.0041	0.80952+-0.0043	0.16196+-0.0033	0.24275+-0.0042	0.62231+-0.0048
VIN	*0.00000*	*0,5964467*	*0,63056*	*	0.60786+-0.0046	0.59608+-0.0050	0.15771+-0.0041	0.75369+-0.0046
SOP	*0.00000*	*0,5964467*	*0,80952*	*0,7260283*	*	0.73577+-0.0041	0.30413+-0.0040	0.75012+-0.0049
LGY	*0.00000*	*0,63056*	*0,4122618*	*0,7260283*	*0,7816044*	*	0.20364+-0.0042	0.47292+-0.0055
BL	*0.00000*	*0,5964467*	*0,5228462*	*0,4122618*	*0,5964467*	*0,47516*	*	0.15454+-0.0039
LBK	*0.00000*	*0,5964467*	*0,7260283*	*0,7816044*	*0,7816044*	*0,63056*	*0,4122618*	*

Genetic distances were calculated from HVS-I sequences (np 16056-16400). Genetic distance p values were post hoc adjusted to correct for multiple comparison by Benjamin and Hochberg method (italicised). Population/culture abbreviations: hunter-gatherers in Central and North Europe (HGCN), LBK in Central Europe (LBK), Starčevo (STA), LBK in Transdanubia (LBKT), Vinča (VIN), Sopot (SOP), Lengyel (LGY), Balaton-Lasinja (BL).

Supplementary Table 15a-b. ASHA with the Transdanubian populations and the Central European LBK population.

A

Summary of shared haplotypes (%)

	HGCN 19	STA 44	VIN 31	LBKT 42	SOP 37	LGY 82	BL 13
HGCN	100.00	2.27	6.45	0.00	0.00	2.44	0.00
STA		100.00	48.39	61.90	54.05	51.22	38.46
VIN			100.00	59.52	40.54	42.68	7.69
LBKT				100.00	56.76	50.00	23.08
SOP					100.00	59.76	7.69
LGY						100.00	38.46
BL							100.00

B

Summary of ancestral shared haplotypes (%)

	HGCN	STA	VIN	LBKT	SOP	LGY	BL
HGCN	100.00						
STA	2.27	97.73					
VIN	6.45	48.39	45.16				
LBKT	0.00	61.90	9.52	28.57			
SOP	0.00	54.05	5.41	10.81	29.73		
LGY	2.44	46.34	4.88	9.76	7.32	31.71	
BL	0.00	38.46	0.00	7.69	0.00	0.00	61.54

Population/culture abbreviations: hunter-gatherers in Central and North Europe (HGCN), Starčevo (STA), LBK in Transdanubia (LBKT), Vinča (VIN), Sopot (SOP), Lengyel (LGY), Balaton-Lasinja (BL).

Supplementary Table 16. AMOVA with the Transdanubian populations and the Central European LBK.

The best 10 arrangements are listed for the two-group compositions, out of the 50 tested arrangements.

	among groups					within groups, among populations				
Arrangement	sum of squares	Va	% of variation	Fct	p	sum of squares	Vb	% of variation	Fsc	p
(STA_LGY_BL)+(LBK_LBKT_VIN_SOP)	4.713	0.01483	0.5	0.00504	0.02861+-0.00144	12.314	-0.01093	-0.37	-0.0037	0.72663+-0.00416
(LBKT_BL)+(LBK_STA_VIN_SOP_LGY)	3.766	0.01186	0.4	0.00403	0.09861+-0.00295	13.261	-0.00583	-0.2	-0.002	0.62040+-0.00439
(STA_SOP_LGY_BL)+(LBK_LBKT_VIN)	4.285	0.0111	0.38	0.00377	0.05931+-0.00225	12.742	-0.00904	-0.31	-0.0031	0.67743+-0.00395
(STA_LBK_VIN_LGY)+(SOP_BL_LBKT)	4.044	0.01076	0.37	0.00365	0.09030+-0.00280	12.983	-0.00717	-0.24	-0.0025	0.64505+-0.00424
(STA_LGY)+(LBK_LBKT_VIN_SOP_BL)	4.01	0.00988	0.34	0.00336	0.14990+-0.00390	13.018	-0.00777	-0.26	-0.0027	0.64406+-0.00454
(STA_VIN_LGY_BL)+(LBK_LBKT_SOP)	3.82	0.00762	0.26	0.00259	0.11693+-0.00297	13.207	-0.00687	-0.23	-0.0023	0.64129+-0.00473
(STA_VIN_LGY)+(LBK_LBKT_SOP_BL)	3.768	0.00728	0.25	0.00248	0.14109+-0.00368	13.259	-0.00661	-0.22	-0.0023	0.61020+-0.00491
(LBKT_SOP)+(LBK_VIN_STA_LGY_BL)	3.434	0.00594	0.2	0.00202	0.19525+-0.00370	13.593	-0.00463	-0.16	-0.0016	0.58396+-0.00528
(STA_SOP_LGY)+(LBK_LBKT_VIN_BL)	3.407	0.00457	0.16	0.00155	0.17980+-0.00389	13.62	-0.00496	-0.17	-0.0017	0.57406+-0.00476
(LBK_LBKT)+(STA_VIN_SOP_LGY_BL)	3.32	0.00427	0.15	0.00145	0.23574+-0.00397	13.707	-0.00472	-0.16	-0.0016	0.59356+-0.00437

AMOVA based on HVS-I sequences (np 16056-16400) was performed between the Carpathian Basin populations and the Central European LBK. Whole datasets were used for the analysis (LBKT n=39, without the Nitra site). The "among groups" and "within groups" variance, F_{ct}, F_{sc}, F_{st}, and significant values (p) were computed. Populations were arranged into different groups (models) and AMOVA was conducted for each arrangement. Populations within one group are indicated in brackets separated by underscores (_). The groups within a model are separated by +. The colouring indicates significant variance and F_{ct}-F_{sc} values. Best models are highlighted with yellow colour.

Abbreviations: LBK in Central Europe (LBK), Starčevo (STA), LBK in Transdanubia (LBKT), Vinča (VIN), Sopot (SOP), Lengyel (LGY), Balaton-Lasinja (BL) populations/cultures.

	n	H	H5	HV	HV0	V	J	K	N1a	T1	T2	U	U2	U3	U4	U5	U5a	U5b	U8	W	X
STA	44	4.5455	2.2727	0	2.2727	6.8182	11.364	27.273	6.8182	2.2727	20.455	0	0	2.2727	2.2727	0	0	0	0	4.5455	6.8182
STA*	40	2.5	2.5	0	2.5	7.5	10	27.5	7.5	2.5	20	0	0	2.5	2.5	0	0	0	0	5	7.5
VIN	31	3.2258	0	0	3.2258	0	9.6774	29.032	12.903	0	19.355	0	3.2258	9.6774	0	0	6.4516	3.2258	0	0	0
VIN*	28	3.5714	0	0	3.5714	0	10.714	28.571	14.286	0	14.286	0	3.5714	10.714	0	0	7.1429	3.5714	0	0	0
LBKT south	23	30.435	0	4.3478	0	0	13.044	17.391	17.391	0	17.391	0	0	0	0	0	0	0	0	0	0
LBKT south*	21	33.333	0	4.7619	0	0	9.5238	19.048	19.048	0	14.286	0	0	0	0	0	0	0	0	0	0
LBKT north	19	15.79	10.526	0	0	5.2632	0	15.79	0	5.2632	31.579	0	10.526	0	0	0	5.2632	0	0	0	0
LBKT north*	16	12.5	12.5	0	0	6.25	0	12.5	0	6.25	31.25	0	12.5	0	0	0	6.25	0	0	0	0
SOP	37	13.514	13.514	2.7027	8.1081	0	13.514	13.514	10.811	0	13.514	0	0	2.7027	0	0	2.7027	0	2.7027	0	2.7027
SOP*	32	12.5	9.375	3.125	9.375	0	12.5	15.625	9.375	0	15.625	0	0	3.125	0	0	3.125	0	3.125	0	3.125
LGY south	41	12.195	7.3171	4.878	2.439	0	12.195	31.707	14.634	0	9.7561	0	0	0	0	2.439	0	0	2.439	0	0
LGY south*	31	9.6774	6.4516	6.4516	3.2258	0	9.6774	29.032	19.355	0	9.6774	0	0	0	0	3.2258	0	0	3.2258	0	0
LGY north	41	21.951	2.439	2.439	4.878	0	2.439	4.878	9.7561	0	26.829	0	0	0	0	0	2.439	7.3171	9.7561	2.439	2.439
LGY north*	33	27.273	3.0303	3.0303	3.0303	0	3.0303	3.0303	12.121	0	21.212	0	0	0	0	0	3.0303	6.0606	9.0909	3.0303	3.0303
BL	13	23.077	0	7.6923	0	0	7.6923	7.6923	0	7.6923	23.077	0	0	0	0	0	0	0	7.6923	15.385	0
BL*	11	18.182	0	9.0909	0	0	9.0909	9.0909	0	9.0909	27.273	0	0	0	0	0	0	0	9.0909	9.0909	0
HGCN	19	0	0	0	0	0	0	0	0	0	0	10.526	5.2632	0	10.526	0	21.053	52.632	0	0	0

	n	H	H5	HV	HV0	V	J	K	N1a	N1b	R	T1	T2	U	U2	U3	U4	U5	U5a	U5b	U8	W	X
HGCN	19	0	0	0	0	0	0	0	0	0	0	0	0	10.526	5.2632	0	10.526	0	21.053	52.632	0	0	0
STA	44	4.5455	2.2727	0	2.2727	6.8182	11.364	27.273	6.8182	0	0	2.2727	20.455	0	0	2.2727	2.2727	0	0	0	0	4.5455	6.8182
VIN	31	3.2258	0	0	3.2258	0	9.6774	29.032	12.903	0	0	0	19.355	0	3.2258	9.6774	0	0	6.4516	3.2258	0	0	0
LBKsouth	23	30.435	0	4.3478	0	0	13.044	17.391	17.391	0	0	0	17.391	0	0	0	0	0	0	0	0	0	0
LBKTnorth	19	15.79	10.526	0	0	5.2632	0	15.79	0	0	0	5.2632	31.579	0	10.526	0	0	0	5.2632	0	0	0	0
SOP	37	13.514	13.514	2.7027	8.1081	0	13.514	13.514	10.811		0	0	13.514	0	0	2.7027	0	0	2.7027	0	2.7027	0	2.7027
LGYsouth	41	12.195	7.3171	4.878	2.439	0	12.195	31.707	14.634	0	0	0	9.7561	0	0	0	0	2.439	0	0	2.439	0	0
LGYnorth	41	21.951	2.439	2.439	4.878	0	2.439	4.878	9.7561	0	0	0	26.829	0	0	0	0	0	2.439	7.3171	9.7561	2.439	2.439
BL	13	23.077	0	7.6923	0	0	7.6923	7.6923	0	0	0	7.6923	23.077	0	0	0	0	0	0	0	7.6923	15.385	0
Körös	16	18.75	0	0	0	6.25	6.25	37.5	0	0	0	0	25	0	0	0	0	0	0	0	0	0	6.25
AVK-Szatm	34	14.706	8.8235	5.8824	0	2.9412	2.9412	23.529	8.8235	0	2.9412	0	11.765	0	0	0	5.8824	0	2.9412	2.9412	2.9412	0	2.9412
AVK-Szakalhat	51	13.725	3.9216	0	1.9608	3.9216	27.451	11.765	3.9216	0	0	3.9216	21.569	0	0	0	1.9608	0	0	3.9216	0	1.9608	0
AVK-Ti/Bükk	39	10.256	0	0	2.5641	2.5641	7.6923	10.256	10.256	0	0	5.1282	17.949	0	0	0	2.5641	0	7.6923	5.1282	2.5641	0	15.385
AVK-Esztar	20	15	5	5	5	10	10	30	0	0	5	0	5	0	0	0	0	0	5	0	5	0	0
Tisza	48	10.417	0	4.1667	4.1667	0	29.167	12.5	4.1667	2.0833	0	2.0833	18.75	0	0	0	0	0	0	2.0833	8.3333	0	2.0833

Supplementary Table 17a-b. Relative mtDNA haplogroup frequencies of the Transdanubian regional groups and the Alföld cultures.

AVK referred to the same as ALBK: Alföld *Linearbandkeramik* culture.

Supplementary Table 18a-b. Slatkin matrix for the MDS with the Transdanubian regional groups and with the Alföld datasets.

AVK referred to the same as ALBK: Alföld *Linearbandkeramik* culture.

A

	STA	STA*	LBKT-south	LBKT-south*	LBKT-north	LBKT-north*	VIN	VIN*	SOP	SOP*	LGY-south	LGY-south*	LGY-north	LGY-north*	BL	BL*
STA	0	0	0.01243	0.0194	0.01209	0.02025	0.00135	0.00613	0.00096	0	0.00578	0.00806	0.01253	0.0109	0.00113	0
STA*	0	0	0.01261	0.01797	0.01547	0.02409	0.00066	0.00434	0.00235	0	0.00347	0.0053	0.01252	0.01003	0.00185	0
LBKT-south	0.01243	0.01261	0	0	0.04539	0.04903	0	0	0	0	0.01267	0	0.01402	0.00493	0.02345	0.02759
LBKT-south*	0.0194	0.01797	0	0	0.06331	0.06979	0	0	0	0	0.0096	0	0.02217	0.01069	0.03511	0.0422
LBKT-north	0.01209	0.01547	0.04539	0.06331	0	0	0.01206	0.03318	0.01432	0.00466	0.0564	0.05909	0.01152	0.01562	0.02728	0.01763
LBKT-north*	0.02025	0.02409	0.04903	0.06979	0	0	0.02073	0.0414	0.01367	0.00457	0.08078	0.07796	0.00545	0.01043	0.02061	0.01063
VIN	0.00135	0.00066	0	0	0.01206	0.02073	0	0	0	0	0	0	0.01327	0.0065	0.02308	0.02249
VIN*	0.00613	0.00434	0	0	0.03318	0.0414	0	0	0	0	0	0	0.01917	0.01089	0.02823	0.03071
SOP	0.00096	0.00235	0	0	0.01432	0.01367	0	0	0	0	0.01519	0.00339	0	0	0.00688	0.01213
SOP*	0	0	0	0	0.00466	0.00457	0	0	0	0	0.00982	0.0004	0	0	0	0
LGY-south	0.00578	0.00347	0.01267	0.0096	0.0564	0.08078	0	0	0.01519	0.00982	0	0	0.04206	0.03313	0.06019	0.063
LGY-south*	0.00806	0.0053	0	0	0.05909	0.07796	0	0	0.00339	0.0004	0	0	0.03188	0.01971	0.0528	0.05758
LGY-north	0.01253	0.01252	0.01402	0.02217	0.01152	0.00545	0.01327	0.01917	0	0	0.04206	0.03188	0	0	0	0
LGY-north*	0.0109	0.01003	0.00493	0.01069	0.01562	0.01043	0.0065	0.01089	0	0	0.03313	0.01971	0	0	0.00316	0.00674
BL	0.00113	0.00185	0.02345	0.03511	0.02728	0.02061	0.02308	0.02823	0.00688	0	0.06019	0.0528	0	0.00316	0	0
BL*	0	0	0.02759	0.0422	0.01763	0.01063	0.02249	0.03071	0.01213	0	0.063	0.05758	0	0.00674	0	0

B

	STA	**LBKT-south**	**LBKT-north**	**VIN**	**SOP**	**LGY-south**	**LGY-north**	**BL**	**Körös**	**AVK-Szatmar**	**AVK-Szakalhat**	**AVK-Tiszadob/ Bükk**	**AVK-Esztar**	**Tisza**
STA	0	0.01239	0.01209	0.00131	0.00094	0.00578	0.01251	0.00108	0	0	0.02436	0.01148	0.011	0.01013
LBKT-south	0.01239	0	0.04539	0	0	0.01266	0.01401	0.02342	0.06682	0.01162	0.04497	0.02145	0.05127	0.02339
LBKT-north	0.01209	0.04539	0	0.01199	0.0143	0.05639	0.01154	0.02724	0.00446	0.02801	0.01671	0.03714	0.05344	0.04297
VIN	0.00131	0	0.01199	0	0	0	0.01323	0.02294	0.01514	0	0.04218	0.024	0.01053	0.02842
SOP	0.00094	0	0.0143	0	0	0.01519	0	0.00685	0.04609	0	0.01549	0.00911	0.0181	0.01215
LGY-south	0.00578	0.01266	0.05639	0	0.01519	0	0.04207	0.06009	0.01084	0	0.08057	0.05215	0.00889	0.0487
LGY-north	0.01251	0.01401	0.01154	0.01323	0	0.04207	0	0	0.04778	0.01304	0.02674	0.00254	0.04225	0.02488
BL	0.00108	0.02342	0.02724	0.02294	0.00685	0.06009	0	0	0.05924	0.0195	0.01529	0	0.04235	0
Körös	0	0.06682	0.00446	0.01514	0.04609	0.01084	0.04778	0.05924	0	0.01747	0.06526	0.06159	0.03644	0.05392
AVK-Szatmar	0	0.01162	0.02801	0	0	0	0.01304	0.0195	0.01747	0	0.06233	0.0156	0	0.04087
AVK-Szakalhat	0.02436	0.04497	0.01671	0.04218	0.01549	0.08057	0.02674	0.01529	0.06526	0.06233	0	0.04582	0.07821	0.00222
AVK-Tiszadob/ Bükk	0.01148	0.02145	0.03714	0.024	0.00911	0.05215	0.00254	0	0.06159	0.0156	0.04582	0	0.0435	0.0288
AVK-Esztar	0.011	0.05127	0.05344	0.01053	0.0181	0.00889	0.04225	0.04235	0.03644	0	0.07821	0.0435	0	0.04588
Tisza	0.01013	0.02339	0.04297	0.02842	0.01215	0.0487	0.02488	0	0.05392	0.04087	0.00222	0.0288	0.04588	0

Supplementary Table 19. Fst values and p values of the regional Transdanubian groups and populations.

P values are italicised. Significant p values (<0.05) are marked by green colour.

	STA	LBKTsou	LBKTno	VIN	SOP	LGYsou	LGYno	BL
STA	*	*0.20137+-0.0043*	*0.20562+-0.0036*	*0.36660+-0.0050*	*0.37759+-0.0050*	*0.24839+-0.0040*	*0.12751+-0.0033*	*0.38491+-0.0045*
LBKTsou	0.01228	*	*0.06732+-0.0028*	*0.81220+-0.0037*	*0.69894+-0.0042*	*0.20939+-0.0042*	*0.18266+-0.0038*	*0.20147+-0.0041*
LBKTno	0.01195	0.04342	*	*0.23354+-0.0041*	*0.18117+-0.0040*	*0.02594+-0.0016*	*0.20483+-0.0038*	*0.15414+-0.0042*
VIN	0.00135	-0.01884	0.01191	*	*0.60697+-0.0052*	*0.61707+-0.0047*	*0.15127+-0.0040*	*0.15870+-0.0033*
SOP	0.00096	-0.012	0.01412	-0.00689	*	*0.12821+-0.0034*	*0.63489+-0.0051*	*0.29720+-0.0049*
LGYsou	0.00575	0.01251	0.05339	-0.00785	0.01497	*	*0.01238+-0.0012*	*0.03960+-0.0020*
LGYno	0.01237	0.01382	0.01139	0.0131	-0.00583	0.04036	*	*0.50361+-0.0051*
BL	0.00113	0.02291	0.02655	0.02256	0.00683	0.05677	-0.00459	*

Supplementary Table 20. AMOVA analysis of the Transdanubian regional groups.

The best five arrangements are listed.

	among groups					within groups, among populations				
Arrangement	**sum of squares**	**Va**	**% of variation**	**Fct**	**p**	**sum of squares**	**Vb**	**% of variation**	**Fsc**	**p**
(STA_LGYsou_VIN_LBKTsou_SOP)+(LGYno_BL_LBKTno)	8.083	0.04651	1.54	0.0154	0.01802+-0.0013	19.326	0.0087	0.29	0.0029	0.29515+-0.00453
(LGYsou_VIN_LBKTsou)+(STA_SOP_LGYno_BL_LBKTno)	8.523	0.0454	1.51	0.0151	0.03327+-0.0019	18.886	0.0063	0.21	0.0021	0.33614+-0.00498
(STA_LGYsou_VIN_LBKTsou)+(LGYno_BL_SOP_LBKTno)	8.315	0.04132	1.37	0.0138	0.02723+-0.0016	19.094	0.0075	0.25	0.0025	0.31198+-0.00472
(LGYsou_VIN_LBKTsou_SOP)+(STA_LGYno_BL_LBKTno)	7.756	0.03554	1.18	0.0118	0.05921+-0.0026	19.653	0.0106	0.35	0.0036	0.27713+-0.00465
(LBKT-no_STA_VIN_LGY-sou)+(LGYno_BL_SOP_LBKT-sou)	6.862	0.02715	0.91	0.0091	0.08366+-0.0028	20.547	0.0155	0.52	0.0052	0.20436+-0.00372

	STA	LBKT-south	LBKT-north	VIN	SOP	LGY-south	LGY-north	BL	Körös	ALBK-Szatmar	ALBK-Szakalhat	ALBK-Tiszadob/Bükk	ALBK-Esztar	Tisza
STA	*	*0.20255+-0.0039*	*0.20533+-0.0044*	*0.36145+-0.0052*	*0.37650+-0.0050*	*0.25304+-0.0042*	*0.12821+-0.0031*	*0.38650+-0.0053*	*0.56529+-0.0049*	*0.54361+-0.0052*	*0.03425+-0.0017*	*0.14801+-0.0037*	*0.22077+-0.0035*	*0.16573+-0.0036*
LBKT-south	*0.01224*	*	*0.07247+-0.0025*	*0.80824+-0.0042*	*0.69449+-0.0048*	*0.20602+-0.0036*	*0.17691+-0.0036*	*0.19642+-0.0040*	*0.05237+-0.0022*	*0.22196+-0.0042*	*0.02307+-0.0016*	*0.11791+-0.0029*	*0.05851+-0.0028*	*0.10148+-0.0034*
LBKT-north	*0.01195*	*0.04342*	*	*0.22414+-0.0039*	*0.18592+-0.0034*	*0.02554+-0.0016*	*0.21325+-0.0045*	*0.15256+-0.0037*	*0.32254+-0.0045*	*0.09118+-0.0028*	*0.15484+-0.0040*	*0.04594+-0.0021*	*0.05326+-0.0020*	*0.03109+-0.0019*
VIN	*0.0013*	*-0.01887*	*0.01184*	*	*0.60311+-0.0048*	*0.61529+-0.0049*	*0.15187+-0.0035*	*0.16187+-0.0040*	*0.22414+-0.0039*	*0.64241+-0.0046*	*0.01505+-0.0010*	*0.06890+-0.0025*	*0.23998+-0.0041*	*0.04821+-0.0023*
SOP	*0.00094*	*-0.01198*	*0.0141*	*-0.00693*	*	*0.13306+-0.0033*	*0.63113+-0.0046*	*0.29334+-0.0044*	*0.03970+-0.0017*	*0.45392+-0.0053*	*0.09385+-0.0030*	*0.18731+-0.0041*	*0.13652+-0.0038*	*0.13741+-0.0031*
LGY-south	*0.00574*	*0.0125*	*0.05338*	*-0.00789*	*0.01496*	*	*0.00921+-0.0009*	*0.04316+-0.0020*	*0.25790+-0.0038*	*0.67320+-0.0043*	*0.00020+-0.0001*	*0.00465+-0.0006*	*0.26017+-0.0047*	*0.00535+-0.0008*
LGY-north	*0.01236*	*0.01382*	*0.01141*	*0.01306*	*-0.00581*	*0.04038*	*	*0.49599+-0.0051*	*0.03812+-0.0019*	*0.14870+-0.0035*	*0.02574+-0.0015*	*0.32888+-0.0049*	*0.02822+-0.0015*	*0.03307+-0.0021*
BL	*0.00108*	*0.02288*	*0.02651*	*0.02243*	*0.0068*	*0.05669*	*-0.00456*	*	*0.07425+-0.0029*	*0.16612+-0.0041*	*0.20562+-0.0038*	*0.60588+-0.0043*	*0.07712+-0.0031*	*0.46035+-0.0043*
Körös	*-0.00877*	*0.06263*	*0.00444*	*0.01492*	*0.04406*	*0.01073*	*0.0456*	*0.05592*	*	*0.19107+-0.0040*	*0.02138+-0.0016*	*0.02158+-0.0016*	*0.12504+-0.0036*	*0.02980+-0.0015*
ALBK-Szatmar	*-0.00448*	*0.01149*	*0.02725*	*-0.00852*	*-0.00159*	*-0.00866*	*0.01287*	*0.01913*	*0.01717*	*	*0.00188+-0.0004*	*0.11900+-0.0032*	*0.77240+-0.0045*	*0.01208+-0.0011*
ALBK-Szakalhat	*0.02378*	*0.04303*	*0.01644*	*0.04047*	*0.01526*	*0.07456*	*0.02605*	*0.01506*	*0.06126*	*0.05867*	*	*0.00277+-0.0005*	*0.00327+-0.0006*	*0.32393+-0.0043*
ALBK-Tiszadob/Bükk	*0.01135*	*0.021*	*0.03581*	*0.02344*	*0.00903*	*0.04957*	*0.00254*	*-0.00944*	*0.05801*	*0.01536*	*0.04382*	*	*0.02534+-0.0015*	*0.02366+-0.0015*
ALBK-Esztar	*0.01088*	*0.04877*	*0.05073*	*0.01042*	*0.01777*	*0.00881*	*0.04053*	*0.04063*	*0.03516*	*-0.01386*	*0.07254*	*0.04169*	*	*0.02604+-0.0016*
Tisza	*0.01003*	*0.02285*	*0.0412*	*0.02763*	*0.01201*	*0.04644*	*0.02428*	*-0.0033*	*0.05116*	*0.03926*	*0.00221*	*0.028*	*0.04387*	*

Supplementary Table 21. F_{st} values and p values from the Transdanubian regional groups and the Alföld populations.

P values are in the upper-right part of the table (italicized). Significant p values (>0.05) are marked with grey colour.

	among groups					within groups, among populations				
	sum of squares	Va	% of variation	Fct	p	sum of squares	Vb	% of variation	Fsc	p
(STA_VIN_LBKTsou_KÖR_LGYsou_ALBKszatm_ALBKesz)+(SOP_LGYno_LBKTno_ALBKszak_ALBKtibu_TIS_BL)	16.368	0.05542	1.92	0.01918	0.0050+-0.00022	45.24	0.03048	1.05	0.01075	0.02099+-0.00129
(STA_VIN_LBKTsou_KÖR_LGYsou_ALBKszatm_ALBKesz)+(LGYno_LBKTno_ALBKszak_ALBKtibu_TIS_BL)	15.804	0.05251	1.82	0.01818	0.00119+-0.0038	45.804	0.03205	1.11	0.0113	0.01554+-0.00113
(STA_VIN__KÖR_LGYsou_ALBKszatm_ALBKesz_LBKTno_)+(SOP_LGYno_LBKTsou_ALBKszak_ALBKtibu_TIS_BL)	15.539	0.05152	1.78	0.011784	0.00139+-0.00035	46.07	0.03265	1.13	0.01151	0.01267+-0.00099
(STA_SOP_VIN_LBKTsou_KÖR_LGYsou_ALBKszatm_ALBKesz_LBKTno_)+(LGYno_ALBKszak_ALBKtibu_TIS_BL)	15.253	0.050777	1.76	0.01758	0.00089+-0.00028	46.356	0.0336	1.16	0.01184	0.01297+-0.00117
(VIN_LBKTsou_KÖR_ALBKszatm_ALBKesz_LBKTno_)+(SOP_STA_LGYno_ALBKszak_ALBKtibu_TIS_BL)	14.206	0.04656	1.61	0.01613	0.00327+-0.00055	47.403	0.03598	1.25	0.01267	0.00812+-0.00088
(SOP_VIN_LBKTsou_KÖR_LGYsou_ALBKszatm_ALBKesz_LBKTno_)+(STA_LGYno_ALBKszak_ALBKtibu_TIS_BL)	13.723	0.04224	1.47	0.01465	0.00495+-0.00074	47.886	0.03745	1.3	0.01318	0.00604+-0.00077
(KÖR_LBKTno_ALBKszak)+(SOP_LGYno_LBKTsou_ALBKtibu_TIS_BL_LGYsou_ALBKszatm_ALBKesz_STA_VIN)	10.201	0.04022	1.39	0.01391	0.03129+-0.00172	51.407	0.04684	1.62	0.01643	0.00079+-0.00027

Supplementary Table 22. AMOVA analysis with the Transdanubian regional groups and the Alföld datasets.

The best five arrangements are listed.

Supplementary Table 23. Genetic distance mapping with the prehistoric mtDNA results.

country	Kazakhstan	Tunisia	Tunisia	Morocco	Morocco	Morocco	Libya	Libya	Algeria	Egypt	
region		Tunisia, south	Tunisia, north	Morocco, central	Morocco, northwest	Morocco, northeast	Libya, south	Libya, north			
population											
abbr.	Kazakhs	Tunisians	Tunisians	Moroccans	Moroccans	Moroccans	Libyans	Libyans	Algerians	Egyptians	
n all	KAZ	TUNs	TUNn	MORc	MORnw	MORne	LIBs	LIBn	ALG	EGY	
n random selected	457	263	311	268	325	431	129	269	129	491	
latitude	140	140	140	140	140	140	129	140	129	140	
longitude	48.019573	33.455206	36.818741	31.633311	35.576203	34.686667	21.566448	26.335100	32.489047	26.820553	
Fst values	66.923684	9.767869	10.165977	-8.000010	-5.368437	-1.911397	24.833216	17.228331	3.678534	30.802498	
STA	0.09078	0.04489	0.03503	0.04677	0.03261	0.04168	0.0768	0.03291	0.07566	0.02699	
LBKT	0.11382	0.07793	0.05617	0.07458	0.05264	0.07218	0.09727	0.06196	0.11009	0.04945	
VIN	0.11186	0.06771	0.05842	0.06831	0.0565	0.06919	0.08878	0.062	0.10497	0.04793	
SOP	0.08284	0.05	0.03202	0.04584	0.02801	0.04663	0.06633	0.03922	0.07949	0.02963	
LGY	0.09232	0.05292	0.03714	0.05006	0.03905	0.051	0.06627	0.04679	0.08375	0.03269	
BL	0.08625	0.02483	0.0091	0.02574	0.00518	0.01987	0.0544	0.00139	0.0491	-0.00421	
references	Comas et al. 1998 (55), Comas et al. 2004 (20), Gokcumen et al. 2008 (136), Irwin et al. 2010 (246)	Ennafaa et al. 2011 (80), Fadhlaoui-Zid et al. 2004 (153), Loueslati et al. 2006 (30)	Cherni et al. 2009 (248), Turchi et al. 2009 (63)	Harich et al. 2010 (81), Falachi et al. 2006 (52), Coudray et al. 2008 (53), Brakez et al. 2001 (50), Rhouda et al. 2009 (32)	Rhouda et al. 2009 (248), Pinto et al. 1996 (18), Rando et al. 1998 (59)	Rhouda et al. 2009 (188), Coudray et al. 2008 (164), Rhouda et al. 2009 (79)	Ottoni et al. 2009 (129)	Fadhlaoui-Zid et al. 2011 (269)	Corte-Real et al. 1996 (85), Plaza et al. 2003 (44)	Krings et al. 1999 (87), Rowold et al. 2007 (115), Saunier et al. 2009 (265), Stevanovich et al. 2003 (24)	

Russia	Russia	Russia	Russia	Uzbekistan	Uzbekistan	Uzbekistan	Turkmenistan	Tajikistan	Kyrgyzstan	**country**
Russia, Ivanovo Oblast, Vladimir	Russia, Belgorod Oblast	Russia, Bashkortostan Rebublic	Russia, Arkhangelsk Oblast	Uzbekistan, south	Uzbekistan, northwest	Uzbekistan, northeast				**region**
Russians	Russians	Russians	Russians	Uzbeks	Uzbeks	Uzbeks	Turkmens	Tajiks	Kyrgyzs	**population**
YAR	BGO	BSH	ARK	UZBs	UZBnw	UZBne	TUK	TAJ	KYR	**abbr.**
154	242	83	155	160	203	122	352	292	351	**n all**
140	140	83	140	140	140	122	140	140	140	**n random selected**
57.929669	50.710693	54.231217	63.285280	38.898623	43.804133	41.266667	38.969719	38.861034	41.204380	**latitude**
39.347329	37.753338	56.164526	42.588419	66.046353	59.445799	69.216667	59.556278	71.276093	74.766098	**longitude**
0.02945	0.03543	0.06273	0.02744	0.05303	0.04295	0.05253	0.05787	0.05033	0.09093	**STA**
0.03114	0.03265	0.07339	0.04496	0.06735	0.06199	0.06565	0.0803	0.07868	0.11314	**LBKT**
0.03633	0.04906	0.06366	0.04716	0.0694	0.0615	0.0713	0.08322	0.07386	0.11215	**VIN**
0.01172	0.01364	0.03675	0.02661	0.04024	0.03772	0.03933	0.05246	0.05381	0.07993	**SOP**
0.02291	0.03099	0.04213	0.02729	0.05182	0.04696	0.05381	0.06461	0.05643	0.0917	**LGY**
0.00891	0.01462	0.06554	0.01736	0.05381	0.03373	0.04041	0.05342	0.04402	0.08443	**BL**
Malyarchuk et al. 2004 (113), Morozova et al. unpublished (41)	Balanovsky et al. unpublished (242)	Belyaeva et al. 2003 (83)	Belyaeva et al. 2003 (75), Tonks et al. unpublished (80)	Comas et al. 2004 (20), Devor et al. 2009 (27), Irwin et al. 2010 (71), Quintana-Murci et al. 2004 (42)	Chaix et al. 2007 (40), Comas et al. 2004 (20), Devor et al. 2009 (2), Irwin et al. 2010 (141)	Devor et al. 2009 (18), Irwin et al. 2010 (104)	Chaix et al. 2007 (51), Comas et al. 2004 (20), Irwin et al. 2010 (240), Quintana-Murci et al. 2004 (41)	Comas et al. 2004 (20), Derenko et al. 2007 (44), Irwin et al. 2010 (228)	Comas et al. 1998 (93), Comas et al. 2004 (20), Irwin et al. 2010 (238)	**references**

Fst values

Russia	Russia	Russia	Russia	Russia	Russia	Russia	Russia	Russia	Russia	**country**
				Russia, Smolensk Oblast	Russia, Rostow Oblast	Russia, Novgorod Oblast	Russia, North Caucasian District	Russia, Kostrona Oblast, Kursk Oblast	Russia, Kaluga Oblast, Oryol Oblast, Tula Oblast	**region**
Khants, Mansi	Karelians	Kalmyks	Altaians	Russians	Russians	Russians	Russians	Russians	Russians	**population**
KHA	KAR	KMY	ALT	SMO	ROS	NOV	CAU	TAM	TUL	**abbr.**
277	595	106	406	180	124	190	37	231	184	**n all**
140	140	106	140	140	124	140	37	140	140	**n random selected**
62.228706	63.753419	46.567684	50.618192	54.988299	47.685325	58.823193	43.402330	52.641659	54.163768	**latitude**
70.641006	33.979229	45.773161	86.219931	32.667738	41.825895	33.412752	45.718747	41.421645	37.564951	**longitude**
0.06114	0.06974	0.10518	0.08952	0.02579	0.04097	0.03185	0.02435	0.03655	0.0245	**STA**
0.0722	0.08164	0.13262	0.1156	0.02619	0.04893	0.02937	0.04808	0.03724	0.02352	**LBKT**
0.08158	0.08853	0.12178	0.11052	0.04091	0.0522	0.03868	0.05378	0.04862	0.03232	**VIN**
0.04313	0.05327	0.10211	0.08381	0.00929	0.01982	0.00796	0.02732	0.01422	0.00839	**SOP**
0.06387	0.05824	0.10842	0.09467	0.02536	0.03382	0.02344	0.04016	0.02837	0.02099	**LGY**
0.04635	0.05969	0.11311	0.10012	0.0096	0.04894	0.01912	-0.00767	0.01842	0.01302	**BL**
Derbeneva et al. 2002 (98), Pimenoff et al. 2008 (169), Voevoda et al. unpublished (10)	Lappalainen et al. 2008 (512), Sajantila et al. 1995 (83)	Derenko et al. 2007 (106)	Derenko et al. 2003 (107), Derenko et al. 2007 (212), Shields et al. 1993 (16), Starikovskaya et al. 2005 (71)	Balanovsky et al. unpublished (147), Morozova et al. unpublished (33)	Kornienko et al. unpublished (124)	Grzybowski et al. 2007 (156), Morozova et al. unpublished (34)	Quintana-Murci et al. 2004 (37)	Morozova et al. unpublished (129), Orekov et al. 1999 (102)	Malyarchuk et al. 2004 (144), Morozova et al. unpublished (40)	**references**

Fst values

Iran	Georgia	Azerbaijan	Armenia	Pakistan	Pakistan	Afghanistan	Russia	Russia	Russia	**country**
Iran, north				Parkistan, south	Parkistan, north					**region**
Iranans	Georgians	Azerbaijan, Turks	Armenians	Pakisthani	Pakisthani	Afghans	Tatars	Ossetians	Nenets, Komi	**population**
IRAn	GEO	AZE	ARM	PAKs	PAKn	AFG	TAT	OSS	NEN	**abbr.**
342	261	68	29	274	327	90	197	199	86	**n all**
140	140	68	29	140	140	90	140	140	86	**n random selected**
35.696113	42.315407	40.143105	40.069099	25.894302	34.952620	33.939110	55.180236	43.045130	67.607834	**latitude**
51.423054	43.356892	47.576927	45.038189	68.524715	72.331113	67.709953	50.726394	44.287097	57.633833	**longitude**
0.03163	0.01569	0.01128	0.01678	0.0713	0.04113	0.03121	0.04309	0.04638	0.04972	**STA**
0.03917	0.0253	0.02359	0.03068	0.08251	0.05234	0.06179	0.05255	0.06036	0.07011	**LBKT**
0.04727	0.03332	0.03498	0.03951	0.07801	0.05518	0.04088	0.05623	0.06168	0.07922	**VIN**
0.01614	0.01249	0.00922	0.015	0.05326	0.02785	0.0417	0.023	0.03844	0.05024	**SOP**
0.0335	0.01765	0.01882	0.02986	0.06103	0.03668	0.03112	0.0365	0.05009	0.05862	**LGY**
0.00931	-0.00347	-0.01423	-0.01122	0.05952	0.02147	0.04683	0.03487	0.02257	0.03003	**BL**
Metspalu et al. 2004 (254), Quintana-Murci et al. 2004 (58), Schönberg et al. 2011 (30)	Alfonso-Sánchez et al. 2006 (48), Comas et al. 2000 (45), Quintana-Murci et al. 2004 (18), Reidla unpublished (124), Schönberg et al. 2011 (26)	Schönberg et al. 2011 (28), Quintana-Murci et al. 2004 (40)	Schönberg et al. 2011 (29)	Quintana-Murci et al. 2004 (274)	Quintana-Murci et al. 2004 (111), Rakha et al. 2011 (216)	Irwin et al. 2010 (90)	Malyarchuk et al. 2010 (197)	Kaldma et al. unpublished (199)	Tonks et al. unpublished (70), Voevoda et al. unpublished (16)	**references**

Fst values

Qatar	Oman	Lebanon	Kuwait	Jordan	Israel, Lebanon, Syria	Israel	Iraq	Iraq	Iran	**country**
					Druze		Iraq, Marsh		Iran, south	**region**
Qatari	Arab	Lebanese	Kuwaitis	Jordanans	DRZ	Palestinians	Marsh arabs	Iraqis	Iranans	**population**
QAT	OMA	LEB	KUW	JOR	325	PAL	MSH	IRQ	IRAs	**abbr.**
90	95	787	350	183	140	303	141	167	237	**n all**
90	95	140	140	140	Golan	140	140	140	140	**n random selected**
25.354826	21.512583	33.854721	29.311660	31.949381	35.689532	31.252973	31.536667	33.223191	31.436015	**latitude**
51.183884	55.923255	35.862285	47.481766	35.932911	0.02473	34.791462	47.672222	43.679291	49.041312	**longitude**
0.03273	0.04137	0.01888	0.0232	0.02817	0.06597	0.02284	0.0239	0.01401	0.02099	**STA**
0.05021	0.04974	0.02777	0.03778	0.03919	0.05925	0.03029	0.03829	0.02638	0.03218	**LBKT**
0.0535	0.04008	0.03127	0.04969	0.0426	0.04241	0.03122	0.04196	0.03519	0.03858	**VIN**
0.02731	0.02909	0.01008	0.02219	0.01744	0.0378	0.00709	0.01511	0.0117	0.01078	**SOP**
0.04258	0.04244	0.0178	0.03608	0.02887	0.01787	0.01993	0.02938	0.02003	0.02562	**LGY**
0.00671	0.03089	0.0042	-0.00936	0.00604	Macaulay et al. 1999 (45),	-0.0011	-0.00052	-0.01887	-0.00316	**BL**
Rowold et al. 2007 (90)	Rowold et al. 2007 (95)	Haber et al. 2011 (787)	Scheible et al. 2011 (350)	Gonzalez et al. 2008 (145), Rowold et al. 2007 (38)		Amar et al. 2007 (295), Di Rienzo and Wilson 1991 (8)	Al-Zahery et al. 2011 (141)	Al-Zahery et al. 2003 & Al-Zahery et al. 2011 (167)	Metspalu et al. 2004 (144), Nasidze et al. 2008 (77), Quintana-Murci et al. 2004 (16)	**references**

Fst values

France	France	France	France	Austria	Yemen	United Arab Emirates	Turkey	Syria	Saudi Arabia	**country**
France, Normandy, Picardie	France, Loire	France, Brittany	France, Aquitanien			Arab				**region**
French	French	French	French	Austrian	Arab	UAE	Turks	Syrian	Arab	**population**
NOM	LOI	BRI	AQU	AUS	YEM	361	TUR	SYR	SAU	**abbr.**
189	165	220	81	374	394	140	245	49	551	**n all**
140	140	140	81	140	140	United Arab Emirates	140	49	140	**n random selected**
49.524641	47.763284	48.202047	44.700222	47.516231	15.552727	53.847818	38.963745	34.802075	24.688002	**latitude**
0.882833	-0.329969	-2.932643	-0.299578	14.550072	48.516388	0.03101	35.243322	38.996815	46.722433	**longitude**
0.0173	0.0318	0.03196	0.05129	0.01478	0.02983	0.04487	0.03233	0.01411	0.03049	**STA**
0.02911	0.03107	0.03455	0.05405	0.02406	0.05631	0.05025	0.04422	0.02272	0.04563	**LBKT**
0.03022	0.03589	0.03564	0.06845	0.03393	0.04711	0.02027	0.04485	0.02679	0.05368	**VIN**
0.01055	0.01103	0.01193	0.02592	0.01421	0.03598	0.03549	0.01596	0.00649	0.02734	**SOP**
0.01518	0.0188	0.02245	0.04359	0.01914	0.04231	0.0124	0.02749	0.01624	0.04322	**LGY**
0.01039	0.0307	0.02565	0.04113	0.00656	0.01319	Alshamali et al. 2008 (235),	0.01773	-0.00515	0.01364	**BL**
Dubut et al. 2004 (38), Garcia et al. 2011 (27), Richard et al. 2007 (124)	Richard et al. 2007 (165)	Dubut et al. 2004 (59), Richard et al. 2007 (161)	Richard et al. 2007 (81)	Brandstätter et al. 2007 (259), Handt et al. 1994 (16), Parson et al. 1998 (99)	Cherný et al. 2008 (184), Kivisild et al. 2004 (113), Non et al. 2011 (50), Rowold et al. 2007 (47)		Calafell et al. 1996 (28), Comas et al. 1996 (45), Di Benedetto et al. 2001 (72), Kivisild et al. 2002 (5), Quintana-Murci et al. 2004 (48), Richards et al. 1996 (22), Schönberg et al. 2011 (25)	Vernesi et al. 2001 (49)	Abu-Amero et al. 2008 (551)	**references**

Fst values

Poland	Poland	Hungary	Czech Republic	Switzerland	Germany	Germany	Germany	France	France	**country**
Poland, Pomerania	Poland, Kuyavian-Pomerania				Germany, Saxony	Germany, Mecklenburg-Vorpommern	Germany, central		France, Poitou, Périgord-Limousin	**region**
Poles	Poles	Hungarians	Czechs	Swiss	German	German	German	Corsican	French	**population**
POM	KUY	HUN	CZE	SWZ	SAX	MEC	GERc	COR	POU	**abbr.**
251	435	367	268	225	1197	507	259	99	193	**n all**
140	140	140	140	140	140	140	140	99	140	**n random selected**
54.294425	53.164836	47.162494	49.817492	46.818188	52.264148	53.612650	51.433237	42.039604	45.833619	**latitude**
18.153116	18.483422	19.503304	15.472962	8.227512	10.526380	12.429595	7.661594	9.012893	1.261105	**longitude**
0.0409	0.0341	0.02297	0.02711	0.02541	0.02768	0.02899	0.04402	0.03671	0.02532	**STA**
0.04862	0.03599	0.02519	0.0288	0.02716	0.03199	0.03028	0.04689	0.04122	0.03915	**LBKT**
0.06004	0.05063	0.03855	0.0417	0.04028	0.04441	0.04338	0.05058	0.0497	0.03868	**VIN**
0.02456	0.01478	0.00625	0.00677	0.00806	0.01578	0.00815	0.02316	0.01813	0.01647	**SOP**
0.03455	0.02999	0.01882	0.02258	0.02308	0.0274	0.02462	0.02957	0.02723	0.02115	**LGY**
0.02473	0.02057	0.0091	0.0079	0.01136	0.0112	0.01158	0.04749	0.03192	0.03299	**BL**
Grzybowski et al. 2007 (251)	Malyarchuk et al. 2002 (435)	Bogácsi-Szabó et al. 2005 (73), Irwin et al. 2006 (201), Tömöry et al. 2007 (93)	Malyarchuk et al. 2006 (177), Vanecek et al. 2004 (91)	Dimo-Simonin et al. 2000 (151), Pult et al. 1994 (74)	Pfeiffer et al. 2001 (1197)	Poetsch et al. 2004 (299), Tetzlaff et al. 2007 (208)	Baassner et al. 1998 (50), Baasner and Madea 2000 (100), Pfeiffer et al. 1999 (109)	Falachi et al. 2006 (53), Varesi et al. 2000 (46)	Dubut et al. 2004 (69), Richard et al. 2007 (124)	**references**

Fst values

Finland	Finland	Estonia	Denmark	Ukraine	Slovenia	Slovakia	Belorussia	Belorussia	Poland	**country**
Finland, Oulu	Finland, south						Belorussia, south	Belorussia, north	Poland, Silesian	**region**
Finns	Finns	Estonians	Danes	Ukrainians	Slovenians	Slovaks	Belorussians	Belorussians	Poles	**population**
OUL	FINs	EST	DEN	UKR	SLO	SVK	BELs	BELn	SIL	**abbr.**
201	159	165	224	252	358	581	162	99	87	**n all**
140	140	140	140	140	140	140	140	99	87	**n random selected**
64.227000	62.893334	58.595272	56.263920	48.379433	46.151241	48.669026	52.164875	55.295983	50.571659	**latitude**
27.728500	27.679328	25.013607	9.501785	31.165580	14.995463	19.699024	29.133325	28.758363	19.321977	**longitude**
0.0761	0.05592	0.03986	0.03603	0.0344	0.03323	0.03071	0.036	0.04093	0.02923	**STA**
0.09837	0.07842	0.04041	0.03898	0.03999	0.03204	0.03738	0.04379	0.04247	0.03746	**LBKT**
0.09084	0.07069	0.05318	0.04601	0.05272	0.04423	0.0425	0.05596	0.04358	0.0345	**VIN**
0.06898	0.04334	0.02076	0.01568	0.0171	0.01328	0.01208	0.01893	0.01483	0.01177	**SOP**
0.0661	0.04362	0.03259	0.03194	0.0279	0.02638	0.02277	0.03259	0.0275	0.02	**LGY**
0.07332	0.05063	0.01838	0.02771	0.01646	0.00705	0.01939	0.01595	0.02788	0.02053	**BL**
Meinilä et al. 2001 (201)	Richards et al. 1996 (29), Meinilä et al. 2001 (100), Hedmann et al. 2007 (30)	Lappalainen et al. 2008 (117), Sajantila et al. 1996 (20), Sajantila et al. 1995 (28)	Mikkelsen et al. 2010 (191), Richards et al. 1996 (33)	Balanovsky et al. unpublished (234), Malyarchuk and Derenko 2001 (18)	Malyarchuk et al. 2003 (102), Metspalu et al. unpublished (128), Zupanic Pajnic et al. 2004 (128)	Lehocký et al. 2008 (374), Malyarchuk et al. 2008 (207)	Balanovsky et al. unpublished (71), Belyaeva et al. 2003 (91)	Balanovsky et al. unpublished (99)	Grzybowski et al. 2007 (87)	**references**

Fst values

Italy	Finland, Norway, Sweden	Sweden	Sweden	Norway	Norway	Lithuania	Latvia	Iceland	Finland	**country**
Italy, south	Saami	Sweden, south	Sweden, central	Norway, south	Norway, north				Finland, Western Finland	**region**
Italians	SAA	Swedes	Swedes	Norwegians	Norwegians	Lithunians	Latvians	Icelanders	Finns	**population**
ITAs	322	SWEs	SWEc	NORs	NORn	LIT	LTV	ICE	WEF	**abbr.**
322	140	273	295	289	323	342	412	984	246	**n all**
140	Inari	140	140	140	140	140	140	140	140	**n random selected**
40.637242	27.030464	58.345364	63.171192	59.913869	63.430476	55.169438	56.879635	64.963051	61.497856	**latitude**
15.802221	0.22372	15.519784	14.959180	10.752245	10.395094	23.881275	24.603189	-19.020835	23.759633	**longitude**
0.02628	0.25172	0.03707	0.03762	0.04134	0.0511	0.04289	0.03472	0.03316	0.05461	**STA**
0.02949	0.24813	0.04858	0.04504	0.02974	0.04374	0.04423	0.03738	0.02688	0.07272	**LBKT**
0.0355	0.22171	0.0561	0.04456	0.04527	0.05589	0.04452	0.04521	0.03979	0.06335	**VIN**
0.01001	0.215	0.02428	0.02334	0.01009	0.01711	0.01965	0.01578	0.00999	0.04064	**SOP**
0.02219	0.22566	0.03217	0.02818	0.02576	0.03879	0.02921	0.02618	0.02621	0.04277	**LGY**
0.00931	Delghandi et al. 1998 (61),	0.02068	0.04033	0.02671	0.04666	0.02644	0.00704	0.01599	0.04278	**BL**
Babalini et al. 2005 (136), Ottoni et al. 2009 (186)		Tillmar et al. 2010 (245), Kittles et al. 1999 (28)	Lappalainen et al. 2008 (295)	Opdal et al. 1998 (215), Passarino et al. 2002 (74)	Helgason et al. 2001 (323)	Lappalainen et al. 2008 (163), Kasperavicuite et al. 2004 (179)	Lappalainen et al. 2008 (114), Pliss et al. 2006 (298)	Helgason et al. 2000 (394), Helgason et al. 2003 (551), Sajantila et al. 1995 (39)	Meinilä et al. 2001 (100), Hedmann et al. 2007 (146)	**references**

Fst values

Spain	Portugal	Portugal	Portugal	Italy	Italy	Italy	Italy	Italy	Italy	**country**
Spain, central	Portugal, south	Portugal, north	Portugal, central	Sardinia	Italy, Tuscany	Italy, South Tyrol	Italy, Sizily	Italy, Latium	Italy, Emilia-Romagna	**region**
Spaniards	Portuguese	Portuguese	Portuguese	Italians	Italians	Italians	Italians	Italians	Italians	**population**
SPAc	PORs	PORn	PORc	SAR	TUS	STY	SIZ	LAT	EMI	**abbr.**
144	545	452	367	302	473	540	468	232	189	**n all**
140	140	140	140	140	140	140	140	140	140	**n random selected**
40.416691	38.015624	41.149968	40.211491	40.120875	43.567115	46.433666	37.397930	41.655242	44.596761	**latitude**
-3.700345	-7.865235	-8.610243	-8.429201	9.012893	10.980700	11.169330	14.658782	12.989615	11.218640	**longitude**
0.03806	0.03128	0.02006	0.01684	0.03203	0.02516	0.03273	0.017	0.03767	0.01345	**STA**
0.04378	0.0507	0.02737	0.02446	0.036	0.02927	0.03308	0.0269	0.03713	0.02203	**LBKT**
0.05092	0.04779	0.03307	0.04009	0.04771	0.04202	0.04356	0.03224	0.05158	0.02934	**VIN**
0.01711	0.02703	0.00757	0.00576	0.01171	0.00841	0.01829	0.00961	0.01764	0.00878	**SOP**
0.03062	0.03109	0.01551	0.01908	0.02538	0.01985	0.02766	0.01885	0.0301	0.01727	**LGY**
0.02849	0.00955	0.01132	0.00842	0.01438	0.01415	0.02771	-0.00309	0.01717	0.00505	**BL**
Alvarez et al. 2007 (31), Larruga et al. 2001 (38), Prieto et al. 2011 (75)	Gonzalez et al. 2003 (137), Pereira et al. 2004 (123), Pereira et al. 2010 (285)	Gonzalez et al. 2003 (84), Pereira et al. 2004 (187), Prieto et al. 2011 (181)	Gonzalez et al. 2003 (78), Pereira et al. 2004 (238), Prieto et al. 2011 (51)	Di Rienzo and Wilson 1991 (68), Falachi et al. 2006 (234)	Achilli et al. 2007 (315), Falachi et al. 2006 (61), Francalacci et al. 1996 (49), Turchi et al. 2008 (48)	Pichler et al. 2006 (217), Stencio et al. 1996 (60), Thomas et al. 2008 (263)	Cali et al. 2001 (106), Forster et al. 2002 (159), Ottoni et al. 2009 (154),Vona et al. 2001 (49)	Babalini et al. 2005 (53), Messina et al. 2010 (124), Turchi et al. 2008 (55)	Bini et al. 2003 (99), Turchi et al. 2008 (90)	**references**

Fst values

Spain	Spain	Spain	Spain	Spain	Spain	Spain	Spain	**country**
	Spain, Valencia, Murcia	Spain, Galicia	Spain, Catalonia	Spain, Castilia Y Lèon	Spain, Cantabria	Spain, Asturias	Spain, Andalusia	**region**
Basques, west	Spaniards	Spaniards	Spaniards	Spaniards	Spaniards	Spaniards	Spaniards	**population**
BASw	VAL	GAL	CAT	CYL	CTB	AST	AND	**abbr.**
236	117	496	237	107	377	101	396	**n all**
140	117	140	140	107	140	101	140	**N randomly selected**
43.220429	39.470239	42.575055	41.591159	41.835682	43.182840	43.250439	37.544271	**latitude**
-2.698387	-0.376805	-8.133856	1.520862	-4.397636	-3.987843	-5.983258	-4.727753	**longitude**
0.04992	0.04391	0.02678	0.02974	0.02793	0.04615	0.05241	0.02904	**STA**
0.06077	0.03547	0.02838	0.04171	0.03963	0.04511	0.05635	0.04308	**LBKT**
0.06804	0.04567	0.04141	0.04283	0.0454	0.05177	0.05442	0.04662	**VIN**
0.03061	0.01186	0.00768	0.0173	0.01792	0.01757	0.02685	0.01908	**SOP**
0.03862	0.02893	0.02164	0.02463	0.0236	0.02865	0.03317	0.02648	**LGY**
0.03871	0.03317	0.00749	0.0211	0.02589	0.04739	0.05506	0.01831	**BL**
Alfonso-Sánchez et al. 2008 (55), Corte-Real et al. 1996 & Richards et al. 1996 (61), Garcia et al. 2011 (91), Prieto et al. 2011 (29)	Alvarez et al. 2007 (33), Picornell et al. 2005 (42), Prieto et al. 2011 (42)	Alvarez et al. 2007 (30), Álvarez-Iglesias et al. 2009 (281), Garcia et al. 2011 (18), Gonzalez et al. 2003 (43), Prieto et al. 2011 (32), Salas et al. 1998 (92)	Alvarez et al. 2007 (41), Álvarez-Iglesias et al. 2009 (100), Corte-Real et al. 1996 (15), Garcia et al. 2011 (23), Plaza et al. 2003 (46), Prieto et al. 2011 (12)	Alvarez et al. 2007 (15), Larruga et al. 2001 (61), Prieto et al. 2011 (31)	Alvarez et al. 2007 (11), Álvarez-Iglesias et al. 2009 (134), Cordoso et al. 2010 (60), Garcia et al. 2011 (6), Maca-Meyer et al. 2003 (160), Prieto et al. 2011 (6)	Alvarez et al. 2007 (13), Garcia et al. 2011 (76), Prieto et al. 2011 (12)	Alvarez et al. 2007 (64), Casas et al. 2006 (108), Corte-Real et al. 1996 (15), Falachi et al. 2006 (66), Larruga et al. 2001 (50), Plaza et al. 2003 (49), Prieto et al. 2011 (44)	**references**

Fst values

Romania	Macedonia	Greece, Crete	Greece, south	Greece, north	Croatia	Bulgaria	Bosnia Herzegovina	Albania	Spain	
										country
Romanians	Macedonians	Greek	Greek	Greek	Croatians	Bulgarians	Bosnians	Albanians	Basques, east	**region**
ROM	MAC	CRE	GREs	GREn	CRO	BUL	BOS	ALB	BASe	**population abbr.**
105	242	392	117	362	155	879	300	84	266	**n all**
105	140	140	117	140	140	140	140	84	140	**N randomly selected**
45.943161	41.608635	35.240117	37.979180	40.639350	43.295246	42.733883	43.915886	41.153332	43.075630	**latitude**
24.966760	21.745275	24.809269	23.716647	22.944607	17.020760	25.485830	17.679076	20.168331	-2.223667	**longitude**
0.01535	0.01614	0.01989	0.02432	0.02435	0.04155	0.03418	0.03811	0.04221	0.07125	**STA**
0.0218	0.03269	0.02345	0.02845	0.02946	0.04309	0.04136	0.04018	0.04063	0.06116	**LBKT**
0.03025	0.04162	0.03956	0.04242	0.03687	0.04617	0.04647	0.0512	0.0432	0.07248	**VIN**
0.00558	0.01641	0.01012	0.01414	0.00809	0.01838	0.01911	0.01712	0.0164	0.03358	**SOP**
0.0177	0.02183	0.02395	0.02536	0.0196	0.02845	0.02826	0.0282	0.02644	0.04872	**LGY**
0.00192	-0.00765	-0.00078	0.00083	0.00343	0.05651	0.01583	0.03545	0.02528	0.06078	**BL**
Bosch et al. 2006 (105)	Bosch et al. 2006 (37), Kouvatsi et al. 2001 (17), Zimmermann et al. 2007 (188)	Forster et al. 2002 (10), Martinez et al. 2008 (178), Vernesi et al. 2001 (18), Villems unpublished (186)	Forster et al. 2002 (59), Kouvatsi et al. 2001 (28), Vernesi et al. 2001 (30)	Bosch et al. 2006 (25), Forster et al. 2002 (13), Irwin et al. 2006 (300), Kouvatsi et al. 2001 (24)	Babalini et al. 2005 (96), Harvey et al. unpublished (59)	Calafell et al. 1996 (29), Karachanak et al. 2011 (850)	Harvey et al. unpublished (159), Malyarchuk et al. 2003 (141)	Belledi et al. 2000 (42), Bosch et al. 2006 (42)	Bertranpetit et al. 1995 (45), Cardoso et al. 2011 (108), Garcia et al. 2011 (113)	**references**

Fst values

Ireland	Great Britain	Great Britain	Great Britain	Great Britain	Great Britain	Great Britain	Great Britain	Great Britain	Great Britain	Serbia	Romania	
	Wales	Scotland, Highlands	Scotland, Fife, Grampian, Tayside	Scotland, southwest	England, East Anglia	England, southwest	England, southeast	England, north	England, central			**country**
Irish	Welsh	Scots	Scots	Scots	British	British	British	British	British	Serbians	Sekler	**region**
IRE	WAL	HIG	FIF	SCOsw	ANG	ENGsw	ENGse	ENGn	ENGc	SER	SEK	**population abbr.**
299	298	207	377	333	337	188	345	402	267	159	241	**n all**
140	140	140	140	140	140	140	140	140	140	140	140	**N randomly selected**
53.412910	52.469978	57.120000	56.844282	55.865627	52.202544	51.014653	51.500152	54.250359	52.482961	44.016521	46.367001	**latitude**
-8.243890	-3.830377	-4.710000	-3.104201	-4.257223	0.131237	-3.103446	-0.126236	-1.470855	-1.893592	21.005859	25.809450	**longitude**
0.02548	0.03243	0.03248	0.04031	0.02373	0.03238	0.03616	0.01888	0.02417	0.03429	0.03064	0.02748	**STA**
0.03076	0.02976	0.03191	0.03913	0.01827	0.03731	0.0364	0.01825	0.02182	0.03376	0.03655	0.02738	**LBKT**
0.03475	0.0342	0.04384	0.05189	0.04604	0.04563	0.04318	0.03267	0.03673	0.04295	0.04069	0.04491	**VIN**
0.01136	0.01224	0.01017	0.01265	0.00599	0.01355	0.0131	0.00517	0.00612	0.01062	0.01078	0.01099	**SOP**
0.02173	0.02175	0.02754	0.03539	0.02913	0.02675	0.02632	0.01554	0.02165	0.02756	0.02297	0.02617	**LGY**
0.01851	0.02889	0.01612	0.02675	0.00654	0.03043	0.02429	0.00641	0.00764	0.02864	0.01515	0.0018	**BL**
McEvoy et al. 2004 (299)	Sykes 2006 & Helgason et al. 2001 (298)	Sykes 2006 & Helgason et al. 2001 (207)	Sykes 2006 & Helgason et al. 2001 (377)	Sykes 2006 & Helgason et al. 2001 (333)	Sykes 2006 & Helgason et al. 2001 (337)	Sykes 2006 & Helgason et al. 2001 (188)	Sykes 2006 & Helgason et al. 2001 (345)	Sykes 2006 & Helgason et al. 2001 (402)	Sykes 2006 & Helgason et al. 2001 (267)	Harvey et al. unpublished (57), Zgonjanin et al. 2010 (102)	Tömöry et al. 2007 (74), Brandstätter et al. 2007 (167)	**references**

Fst values

Supplementary Table 24. Populations and references used for the PCA with Y-chromosomal data.

Continent	Country	Population	Abbr.	n	References:
				24476	
Africa, east	Egypt	Egyptians	EGY	262	Luis et al. 2004 (147), El-Sibai et al 2009 (115)
Africa, north	Algeria	Algerian	ALG	102	Robino et al. 2008 (102)
Asia, east	China	Tibetans	TIB	110	Karafet et al. 2001 (75), Xue et al. 2006 (35)
Asia, east	China	Manchu	MAN	87	Karafet et al. 2001 (52), Xue et al. 2006 (35)
Asia, north	Russia, China	Ewenki	EWE	162	Karafet et al. 2001 (41+95), Xue et al. 2006 (26)
Asia, east	China	Oroqen	ORO	54	Karafet et al. 2001 (23), Xue et al. 2006 (31)
Asia, east	China	Uygurs	UYU	68	Karafet et al. 2001 (68)
Asia, east	Mongolia	Mongolians	MON	147	Karafet et al. 2001 (147)
Asia, east	China	Yao, She, Zhuang, Miao, Han	CHI	539	Karafet et al. 2001 (51,60), Xue et al. 2006 (134,190,70,34)
Asia, south	Pakistan	Pakistani	PAK	***1052***	Sengupta et al 2006 (176), Firasat et al. 2007 (***876***)
Asia, south	India	Indian	IND	728	Sengupta et al 2006 (728)
Asia, south	Afghanistan	Hazara	HAZ	60	Haber et al. 2012 (60)
Asia, south	Afghanistan	Pashtun	PAS	49	Haber et al. 2012 (49)
Asia, south	Afghanistan	Tajik	TAJ	56	Haber et al. 2012 (56)
Asia, central	Kazakhstan	Kazakhs	KAZ	68	Karafet et al. 2001 (30), Zerjal et al. 2002 (38)
Asia, central	Uzbekistan, Afghanistan	Uzbeks	UZB	71	Karafet et al. 2001 (54), Haber et al. 2012 (17)
Asia, central	Tajikistan	Yagnobi	YAG	31	Wells et al. 2001 (31)
Asia, north	Russia	Buryats	BUR	81	Karafet et al. 2001 (81)
Asia, north	Russia	Altaians	ALT	175	Karafet et al. 2001 (29), Kharkov et al. 2007 (146)
Asia,north	Russia	Altaian Kazakhs	AKAZ	119	Dulic et al. 2011 (119)
Asia, north	Russia	Russians	RUS	***2005***	Fechner et al. 2005, Roewer et al. 2008 (256), Mirabal et al. 2009 (106), Malyarchuk et al. 2008 (414), Balanovsky et al. 2008 (***1229***)
Asia, north	Russia	Komi	KOM	103	Mirabal et al. 2009 (103)
Asia, north	Russia	Udmurts	UDM	43	Semino et al 2000 (43)
Asia, north	Russia	Mari	MAR	46	Semino et al 2000 (46)
Asia, north	Russia	Mansi, Khanti	KHA-MAN	53	Pimentoff et al 2008 (53)
Asia, north	Russia	Nenets	NEN	54	Wells et al. 2001 (54)
Asia, southwest	Azerbaijan, Iran	Azeri	AZE	134	Nasidze et al. 2009 (72), Grugni et al. 2012 (62)
Asia, north	Russia	Ingush	ING	105	Yunusbayev et al. 2011 (105)
Asia, north	Russia	Kumyks	KUM	73	Yunusbayev et al. 2011 (73)
Asia, north	Russia	Avars, Tabarasans, Lezgins, Dargins	CANE	195	Yunusbayev et al. 2011, Zerjal 2002 (12)
Asia, north	Russia	Nogays & Kara Nogays total	NOG	163	Yunusbayev et al. 2011 (163)
Asia, north	Russia	Abazins, Adyghe, Balkars, Cherkessians,	CANW	712	Yunusbayev et al. 2011 (712)

Asia, southwest	Georgia	Kabardin, Karachays Abkhazians	ABK	162	Yunusbayev et al. 2011 (162)
Asia, north	Russia	Chechens	CHE	165	Yunusbayev et al. 2011 (165)
Asia, southwest	Armenia	Armenians	ARM	91	Yunusbayev et al. 2011 (57), Grugni et al. 2012 (34)
Asia, southwest	Georgia	Georgians	GEO	194	Semino et al 2000 (63), Battaglia et al 2009 (66), Yunusbayev et al. 2011 (65)
Asia, north, southwest	Georgia, Russia	S-N Ossetians	OSS	153	Yunusbayev et al. 2011 (153)
Asia, southwest	UAE, Qatar	Arabs	ARA	236	Cadenas et al. 2007 (164, 72)
Asia, southwest	Kuwait	Arabs	KUW	41	El-Sibai et al. 2009 (41)
Asia, southwest	Saudi-Arabia	Arabs	SAU	***173***	Abu-Amero et al. 2009 (***173***)
Asia, southeast	Oman	Arabs	OMA	***120***	Luis et al. 2004 (***120***)
Asia, southwest	Iran	Iranians	IRN	1003	Regueiro 2006 (150), El-Sibai et al. 2009 (91), Grugni et al. 2012 (769)
Asia, southwest	Iraq	Iraqis	IRQ	***362***	Sanchez et al. 2005 (64), Al-Zahery et al. 2011 (***298***)
Asia, southwest	Yemen	Yemeni	YEM	62	Cadenas et al. 2007 (62)
Asia, southwest	Jordan	Jordanians	JOR	418	Flores et al. 2005 (146), El-Sibai et al. 2009 (272)
Asia, southwest	Syria	Syrians	SYR	540	Zalloua et al. 2008, El-Sibai et al. 2009
Asia, southwest	Israel	Jews, Samarit, Druse	ISR	112	Shen et al 2004 (112)
Asia, southwest	Palestina	Palestinans	PAL	88	Zalloua et al. 2008, El-Sibai et al. 2009 (88)
Asia, southwest	Lebanon	Lebanese	LEB	916	Zallua et al. 2008b (916)
Asia, southwest	Turkey	Turks	TUR	612	Semino et al. 2000 (30), Cinnioglu et al 2004 (523), Sanchez at al. 2005 (59)
Europe, west	France	Corsicans	COR	34	Francalacci et al 2003 (34)
Europe, west	France	French	FRE	23	Semino et al 2000 (23)
Europe, west	Netherlands	Dutch	NTH	27	Semino et al 2000 (27)
Europe, west	Great Britain	British	GB	***888***	Wells et al. 2001 (26), Capelli et al 2003. (***862***)
Europe, west	Ireland	Irish	IRL	796	Moore et al 2006 (796)
Euroep, north	Denmark	Danish	DAN	150	Birón et al. 2006 (150)
Europe, north	Sweden	Swedes	SWE	305	Karlsson et al 2006 (305)
Europe, north	Finnland	Finns	FIN	356	Lappalainen et al 2006 (316), Karlson et al. 2006 (40)
Europe, north	Norway	Norwegians	NOR	72	Passarino et al. 2002 (72)
Europe, north	Sweden, Finnland, Russia	Saamis	SAA	189	Semino et al. 2000 (24) Tambets et al 2004 (127), Karlsson et al. 2006 (38)
Europe, south	Portugal	Portuguese	POR	***1459***	Adams et al. 2008 (***139***), Beleza et al. 2005 (***658***), Goncalves et al. 2005 (553), Flores et al. 2004 (109)
Europe, south	Italy	Italians	ITA	786	Semino et al 2000 (87), Capelli et al 2007 (699)
Europe, south	Italy	Sardinian	SAR	77	Semino et al 2000 (77)
Europe, south	Italy	Sicilians	SIC	287	Di Gaetano et al. 2009 (236), Francalacci et al 2003 (51)
Europe, south	Spain	Spaniards	SPA	***1375***	Semino et al 2000 (53), Maca-Meyer et al. 2003 (***81***), Flores et al. 2004 (363), Adams et al. 2008 (***878***)
Europe, south	Canary Islands/Spain	Spaniards	CAN	480	Flores et al 2003 (480)
Europe, south	Spain, France	Basques	BAS	235	Semino et al 2000 (67) Alonso et al 2005 (168)

Europe, central	Germany	Germans	GER	150	Birón et al. 2006 (150)
Europe, central	Hungary	Hungarian	HUN	416	Semino et al. 2000 (45), Behar et al. 2004 (56), Csányi et al. 2008 (100), Völgyi et al. 2009 (215)
Europe, central	Czech Republic, Slovakia	CzechSlov	CZE	120	Semino et al 2000 (45), Battaglia et al. 2009 (75)
Europe, central	Poland	Polish	POL	154	Semino et al 2000 (55), Battaglia et al. 2009 (99)
Europe, east	Ukraina	Ukrainians	UKR	142	Semino et al 2000 (50), Battaglia et al. 2009 (92)
Europe, southeast	Greece, Macedonia	Greek	GRE	514	Semino et al. 2000 (76), Bosch et al. 2006 (41), Firasat et al. 2007 (77) King et al. 2009 (171), Battaglia et al. 2009 (149)
Europe, southeast	Greece, Crete	Greek	CRE	361	Martinez et al. 2007 (168), King et al. 2009 (193)
Europe, southeast	Slovenia	Slovenians	SLO	75	Battaglia et al. 2009 (75)
Europe, southeast	Croatia	Croatian	CRO	722	Semino et al. 2000 (58), Barac et al. 2003 (348), Marjanovic et al. 2005 (90), Pericic et al. 2005, Barac et al. 2003, Rootsi et al. 2004 (108), Battaglia et al. 2009 (118)
Europe, southeast	Albania, Maceonia, Kosovo	Albanians	ALB	***313***	Semino et al 2000 (51), Pericic et al. 2005 (***113***), Bosch et al. 2006 (30), Battaglia et al. 2009 (119)
Europe, southeast	Bosnia-Herzegovina	Bosnians	BOS	***290***	Pericic et al. 2005 (***206***), Battaglia et al. 2009, Marjanovic et al. 2005 (84)
Europe, southeast	Romania	Romanians	ROM	67	Bosch et al. 2006 (67)
Europe, southeast	Bosnia-Herzegovina, Serbia	Serbian	SER	297	Pericic et al 2005 (113), Battaglia et al. 2009, Marjanovic et al. 2005 (81), Regueiro et al. 2012 (103)
Europe, southeast	Macednoia	Macedonians	MAC	***150***	Semino et al. 2000 (20), Pericic et al 2005 (***78***), Bosch et al. 2006 (52)

Note, that the given number of samples are different in some papers from the here listed ones. The recalculated values are: Abu-Amero et al. 2009 (n=175), Al-Zahery et al. 2011 (n=297), Adams et al. 2008 (n=138; 878), Balanovsky et al. 2008 (n=1228), Beleza et al. 2005 (n=657), Capelli et al 2003. (n=862), Firasat et al. 2007 (n=875), Luis et al. 2004 (n=121), Maca-Meyer et al. 2003 (n=82), Pericic et al. 2005 (n=210, 114, 79). In these studies only rough frequencies and total number (n) were published, the sum of them was not 100%, causing a rounding error in sample numbers (modified numbers are highlighted in bold, italicised).

Supplementary Table 25. GDM coordinates and F_{st} values for the NRY population genetic analysis

References	Country	Population	Region	N in the literature	coordinates		Fst values	
					latitude	longitude	STA-LBK	SOP-LGY
Hassan et al. 2008 (369)	Sudan	Sudanese		369	15.58071	32.46899	0.27598	0.16071
Sanchez et al. 2005 (201)	Somalia	Somalis		201	4.52167	46.57874	0.62098	0.50339
Semino et al. 2002 (126)	Ethiopia	Ethiopians		126	9.01530	38.73450	0.43454	0.29922
Luis et al. 2004 (147), El-Sibai et al 2009 (115)	Egypt	Egyptians		262	26.82055	30.80250	0.39419	0.25056
Robino et al. 2008 (102)	Algeria	Algerian	Northwestern Algeria	102	32.48905	3.67853	0.27755	0.15734
Karafet et al. 2001 (68)	China	Uygurs	Xinjiang	68	44.02442	87.48853	0.18548	0.11275
Sengupta et al 2006 (176), Firasat et al. 2007 (***639***)	Pakistan	Pakistani		814	28.45903	69.38306	0.22254	0.14377
Firasat et al. 2007 (237)	Pakistan	Burusho, Kalash, Pathan	Chitral district	237	36.22655	72.06702	0.19748	0.14119
Karafet et al. 2001 (29), Kharkov et al. 2007 (146), Dulic et al. 2011 (119)	Russia	Altaians, Altainan Kazakhs		294	50.61819	86.21993	0.22554	0.13951
Bíró et al. 2009 (45)	Kazakhstan	Madjar	Torgay	45	47.33882	63.71729	0.34682	0.42408
Karafet et al. 2001 (54), Haber et al. 2012 (17)	Uzbekistan, Afghanistan	Uzbeks		71	39.70719	66.92529	0.19494	0.09469
Wells et al. 2001 (31)	Tajikistan	Yagnobi	Sughd province	31	39.58452	69.09766	0.26211	0.15805
Haber et al. 2012 (56)	Afghanistan	Tajik		56	37.08148	70.60742	0.19372	0.11731
Karafet et al. 2001 (30), Zerjal et al. 2002 (38)	Kazakhstan	Kazakhs		68	48.01957	66.92368	0.32275	0.19698
Haber et al. 2012 (60)	Afghanistan	Hazara		60	36.68604	68.81671	0.22399	0.09813
Haber et al. 2012 (49)	Afghanistan	Pashtun		49	34.55181	69.17926	0.29594	0.23345
Roewer et al. 2008, Fechner et al. 2005 (43), Balanovsky et al. (259)	Russia	Russian	Arkhangel (Mezen, Pinega, Krasnoborsk)	301	63.75335	43.50256	0.272	0.20981
Roewer et al. 2008, Fechner et al. 2005 (40), Balanovsky et al. (173)	Russia	Russian	Vologda, Unzha	213	59.66774	40.07153	0.21619	0.15016
Roewer et al. 2008, Fechner et al. 2005 (43), Balanovsky et al. (107)	Russia	Russian	Smolensk, Roslavl	150	54.83866	32.04630	0.27531	0.20225
Roewer et al. 2008, Fechner et al. 2005 (43), Balanovsky et al. (73)	Russia	Russian	Tver, Kashin	116	56.71657	35.46057	0.31829	0.2478
Balanovsky et al. (132)	Russia	Russian	Novgorod, Porhov, Ostrov	132	58.39020	33.88019	0.29979	0.23341
Roewer et al. 2008, Fechner et al. 2005 (91), Balanovsky et al. (394)	Russia	Russian	Chermigov district, South Russia, Brianskaja, Tambovskaja	485	51.80862	39.16461	0.2854	0.22118
Mirabal et al. 2009 (103)	Russia	Komi	Uralic mountains	103	65.18303	54.04932	0.45459	0.38827
Semino et al 2000 (43)	Russia	Udmurts		43	57.11239	52.99463	0.24649	0.1743
Semino et al 2000 (46)	Russia	Mari		46	56.10575	47.21289	0.41555	0.3398
Pimentoff et al 2008 (53), Mirabal et al. 2009 (27)	Russia	Mansi, Khanti	Siberia/Khanty	80	62.22871	70.64101	0.58663	0.51217
Wells et al. 2001 (54)	Russia	Nenets		54	67.60783	57.63383	0.32721	0.25765
Nasidze et al. 2009 (72), Grugni et al. 2012 (62)	Azerbaijan, Iran	Azeri	Azerbaijan, Azerbaijan Gharbi	134	40.02551	48.93544	0.14762	0.07706
Yunusbayev et al. 2011 (105)	Russia	Ingush	NE Caucasus	105	43.22819	44.77673	0.60849	0.48429
Yunusbayev et al. 2011 (73)	Russia	Kumyks	NE Caucasus	73	42.96446	47.47107	0.23625	0.15061

Yunusbayev et al. 2011, Zerjal 2002 (12)	Russia	Avars, Tabarasans, Lezgins, Dargins	NE Caucasus	195	41.75082	47.48755	0.46782	0.34511
Yunusbayev et al. 2011 (163)	Russia	Nogays & Kara Nogays total		163	44.65986	45.64661	0.17847	0.09459
Yunusbayev et al. 2011 (154)	Russia	Adyghe	NW Caucasus	154	44.58264	40.07727	0.12058	0.13547
Battaglia et al. 2009 (38), Yunusbayev et al. 2011 (135)	Russia	Balkarians	NW Caucasus	173	43.48083	43.61487	0.14195	0.122
Yunusbayev et al. 2011 (126)	Russia	Cherkessians	NW Caucasus	126	44.24913	42.02185	0.15092	0.14955
Yunusbayev et al. 2011 (140)	Russia	Kabardin	NW Caucasus	140	43.73935	44.03784	0.12766	0.12553
Yunusbayev et al. 2011 (162)	Georgia	Abkhazians	South Caucasus	162	43.05685	41.44510	0.14607	0.15291
Yunusbayev et al. 2011 (165)	Russia	Chechens	NE Caucasus	165	43.12905	45.57088	0.4921	0.37225
Yunusbayev et al. 2011 (57), Grugni et al. 2012 (34)	Armenia, Iran	Armenians	South Caucasus	91	40.06910	45.03819	0.16028	0.09237
Semino et al 2000 (63), Battaglia et al 2009 (66), Yunusbayev et al. 2011 (65)	Georgia	Georgians		194	41.71803	44.76846	0.14775	0.12637
Yunusbayev et al. 2011 (153)	Georgia, Russia	S-N Ossetians		153	43.04513	44.28710	0.20214	0.2612
Luis et al. 2004 (***120***)	Oman	Arabs		121	21.51258	55.92326	0.3013	0.16855
Abu-Amero et al. 2009 (***173)***	Saudi-Arabia	Arabs		175	26.29342	43.94202	0.34647	0.21821
Grugni et al. 2012 (166)	Iran	Arab, Kurd, Lur	Khuzestan, Kurdistan, Lorestan	166	33.49560	47.59113	0.1919	0.10923
Grugni et al. 2012 (126)	Iran	Turkmen, Persian	Golestan, Khorasan	126	36.61553	57.59967	0.19559	0.12831
Regueiro et al. 2006 (117), Grugni et al. 2012 (190)	Iran	Iranians,Banardi, Afro-Iranian, Qeshmi	South Iran, Hormozgan, Qeshmi	307	27.13248	55.13956	0.17702	0.0954
Grugni et al. 2012 (100)	Iran	Persian	Fars, Isfahan, Yazd	100	32.54681	51.61047	0.20407	0.112
Grugni et al. 2012 (133)	Iran	Gilak, Mazarandi	Gilan, Mazandaran, Tehran	133	35.74651	51.42069	0.19226	0.11748
Sanchez et al. 2005 (64), Al-Zahery et al. 2011 (***298***)	Iraq	Iraquis		361	33.22319	43.67929	0.42071	0.29691
El-Sibai et al. 2009 (41)	Kuwait	Kuwaits		41	29.31166	47.48177	0.26344	0.13703
Cadenas et al. 2007 (72)	Quatar	Quatari		72	25.30927	51.53137	0.40911	0.2723
Cadenas et al. 2007 (***164***)	UAE	Arabs		164	23.42408	53.84782	0.24675	0.12891
Cadenas et al. 2007 (62)	Yemen	Yemeni		62	15.55273	48.51639	0.58826	0.43724
Flores et al. 2005 (146), El-Sibai et al. 2009 (272)	Jordan	Jordanians		418	31.35364	36.25159	0.2948	0.17156
Zalloua et al. 2008, El-Sibai et al. 2009	Syria	Syrians	all	540	34.54276	38.30768	0.32793	0.20946
Zalloua et al. 2008, El-Sibai et al. 2009 (88), Shen et al 2004 (112)	Israel	Jews, Samarit, Druse, Palestin		200	32.06977	34.78811	0.33198	0.20125
Zallua et al. 2008b (916)	Lebanon	Lebanese		916	33.85472	35.86229	0.25532	0.14245
Cinnioglu et al 2004 (133)	Turkey	Turks-1, 9	Marmara region, Istambul	133	41.00944	28.98732	0.19986	0.09152
Cinnioglu et al 2004 (112)	Turkey	Turks2,-3	Black Sea Region Ost-West	112	41.29432	36.33014	0.16846	0.09139
Cinnioglu et al 2004 (125)	Turkey	Turks-4,5	Anatolian Region eastern, Southeastern	125	37.93553	40.20831	0.20532	0.11121
Cinnioglu et al 2004 (153)	Turkey	Turks-6,7,8	Mediterranean Region, Central Anatolia	153	37.79676	29.05719	0.17438	0.09304

			Regon, Aegean Region					
Francalacci et al 2003 (34)	France	Corsicans		34	42.03960	9.01289	0.24792	0.17605
Semino et al 2000 (23)	France	French		23	46.95026	2.19067	0.27402	0.17519
Semino et al 2000 (27)	Netherlands	Dutch		27	56.26392	9.50179	0.4238	0.33116
Birón et al. 2006 (150)	Germany	Germans		150	51.54292	9.94098	0.24561	0.16573
Birón et al. 2006 (150)	Denmark	Danish		150	55.42901	10.43536	0.25624	0.18335
Capelli et al 2003 (***211***), Wells et al. 2001 (26)	Great Britain	British	Orkney, Durness, Pitlorchy	239	56.78283	-3.71503	0.46469	0.40324
Capelli et al 2003 (***273***)	Great Britain	British	Southwell, Llangefni, Norfolk	271	53.33590	-0.98410	0.45025	0.3836
Capelli et al 2003 (166)	Great Britain	British	Haverfordwest, Faversham, Cornwall	166	51.45401	-1.03436	0.60156	0.53634
Moore et al 2006 (796), Capelli et al. 2003 (119)	Ireland	Irish		915	53.41291	-8.24389	0.67519	0.63181
Karlsson et al 2006 (305)	Sweden	Swedes		305	59.31077	15.20178	0.21788	0.14789
Lappalainen et al 2006 (316), Karlson et al. 2006 (40)	Finnland	Finns		356	61.52270	23.72717	0.4368	0.37803
Passarino et al. (2002) (72)	Norway	Norwegians		72	61.39672	8.76379	0.25064	0.178
Semino et al. 2000 (24) Tambets et al 2004 (127), Karlsson et al. 2006 (38)	Sweden, Finnland, Russia	Saami all		189	68.90576	27.03046	0.29092	0.22234
Adams et al. 2008 (***139***), Beleza et al. 2005 (***658***), Goncalves et al. 2005 (553), Flores et al. 2004 (109)	Portugal	Portuguese		1457	40.21149	-8.42920	0.32638	0.25536
Capelli et al 2007 (407), Semino et al. 2000 (50)	Italy	Italian	Central-North Italy (Latium, Tuscany, Elba)	457	43.78696	11.22473	0.23845	0.16749
Capelli et al 2007 (258), Semino et al. 2000 (37)	Italy	Italian	South-Italy (Apulia, Campania, Calabria)	295	40.22502	16.68192	0.20397	0.11994
Semino et al 2000 (77)	Italy	Sardinian		77	40.12088	9.01289	0.13531	0.07673
Di Gaetano et al. 2009 (236), Francalacci et al 2003 (51)	Italy	Sicilians		287	37.39793	14.65878	0.21438	0.11635
Flores et al 2004 (81), Adams et al. 2008 (***131***)	Spain	Spaniard	Castile, Leon	212	41.83568	-4.39764	0.36596	0.29334
Flores et al 2004 (31), Adams et al. 2008 (***136***)	Spain	Spaniard	Valencia, Castilia la Mancha	167	39.44468	-0.39716	0.39363	0.32174
Flores et al 2004 (258), Adams et al. 2008 (***168***), Semino et al. 2000 (29)	Spain	Spaniard	Andalusia	455	37.54427	-4.72775	0.35878	0.28461
Flores et al 2004 (19), Adams et al. 2008 (***108***)	Spain	Spaniard	Galicia, Asturia	127	42.57506	-8.13386	0.33346	0.2511
Flores et al 2004 (36), Maca-Meyer et al. 2003 (118)	Spain	Spaniard	Cantabria	188	43.18284	-3.98784	0.30739	0.24273
Adams et al. 2008 (***80***), Semino et al. 2000 (24)	Spain	Spaniard	Catalonia	104	41.59116	1.52086	0.56689	0.49896
Flores et al 2003 (480)	Canary Islands/Spain	Spaniards		480	28.29156	-16.62913	0.29408	0.2127
Semino et al 2000 (67) Alonso et al 2005 (168)	Spain, France	Basques		235	43.30919	-1.96188	0.66407	0.60654
Battaglia et al. 2009 (75)	Slovenia	Slovenians		75	46.15124	14.99546	0.2494	0.17705
Semino et al. 2000 (45), Behar et al. 2004 (56), Csányi et al. 2008 (100), Völgyi et al. 2009 (215)	Hungary	Hungarian		416	47.16249	19.50330	0.20437	0.12856
Semino et al 2000 (45), Battaglia et al. 2009 (75)	Czech Republic, Slovakia	CzechSlov		120	49.81749	15.47296	0.25792	0.19346

Semino et al 2000 (55), Battaglia et al. 2009 (99)	Poland	Polish		154	51.82220	19.39856	0.34332	0.27338
Semino et al 2000 (50), Battaglia et al. 2009 (92)	Ukraina	Ukrainians		142	48.54571	32.28717	0.28006	0.20787
Bosch et al. 2006 (67)	Romania	Romanians		67	45.94316	24.96676	0.16962	0.08628
Csányi et al. 2008 (97)	Romania	Szeklers		97	46.36209	25.81705	0.18202	0.09258
Semino et al. 2000 (58), Pericic et al. 2005b, Barac et al. 2003, Rootsi et al. 2004 (108), Battaglia et al. 2009 (118)	Croatia	Croatian	Craoatian Mainland	284	45.80200	16.01724	0.23033	0.15507
Barac et al 2003 (91)	Croatia	Croatian	Hvar	91	43.18515	16.59554	0.33158	0.24569
Barac et al 2003 (***134***)	Croatia	Croatian	Korcula	134	42.93732	16.91551	0.22458	0.16567
Pericic et al. 2005 (***206***), Battaglia et al. 2009, Marjanovic et al. 2005 (84)	Bosnia-Herzegovina	Bosnians		294	43.91589	17.67908	0.30943	0.22439
Pericic et al 2005b (113), Battaglia et al. 2009, Marjanovic et al. 2005 (81), Regueiro et al. 2012 (103)	Bosnia-Herzegovina, Serbia	Serbien		297	44.01652	21.00586	0.20381	0.10933
Semino et al. 2000 (20), Pericic et al 2005b (***78***), Bosch et al. 2006 (52)	Macednoia	Macedonians		151	41.60864	21.74528	0.18436	0.08654
Semino et al 2000 (51), Pericic et al. 2005 (114), Bosch et al. 2006 (30), Battaglia et al. 2009 (119)	Albania, Maceonia, Kosovo	Albanians		314	41.15333	20.16833	0.23636	0.11797
Semino et al. 2000 (76), Bosch et al. 2006 (41), Firasat et al. 2007 (77) King et al. 2009 (171), Battaglia et al. 2009 (149)	Greece, Macedonia	Greek	mainland	514	39.55488	21.77325	0.19063	0.08893
Martinez et al. 2007 (168), King et al. 2009 (193)	Greece	Greek	Crete	361	35.24012	24.80927	0.21188	0.11631

16 Y haplogroups were used for the GDM analysis. Note that the number of samples (given in relative freqencies) are different in some papers from the calculated ones. The modified values (highlighted in bold, italicised) are: Abu-Amero et al. 2009 (n=175), Adams et al. 2008 (n=138,136,168,108,80), Al-Zahery et al. 2011 (n=297), Balanovsky et al. 2008 (n=258,173), Beleza et al. 2005 (n=657), Capelli et al. 2003 (n=213,271), Cadenas et al. 2007 (n=164), Firasat et al. 2007 (n=638), Luis et al. 2004 (n=121), Pericic et al. 2005 (n=210,114,79).

Supplementary Table 26. List of primers, used for the mtDNA amplifications.

Loci	Amplicon name	primer name	Sequence 5'-3'	Product Size (bp)	annealing (C°)	Reference
HVS I	2/1	L16045	TGTTCTTTCATGGGGAAGCAGATT	240	56°	Brandt et al. 2013
		H16240	GGGTGGCTTTGGAGTTGCAGTT			Brandt et al. 2013
	2/2	L16212	CCCCATGCTTACAAGCAAGTACA	236	56°	Adler et al. 2011
		H16402	GATATTGATTTCACGGAGGATGGT			Adler et al. 2011
	4/1	L15996	CTCCACCATTAGCACCCAAAGC	194	58°	Endicott et al. 2003
		H16144	TGGGTTTTTATGTACTACAGGTGGTCA			Knipper et al. 2013
	4/2	L16117	TACATTACTGCCAGCCACCAT	162	58°	Haak et al. 2005
		H16233	GCTTTGGAGTTGCAGTTGATGTGT			Haak et al. 2005
	4/3	L16209	CCCCATGCTTACAAGCAAGT	179	58°	Handt et al. 1996
		H16348	ATGGGGACGAGAAGGGATTTG			Haak et al. 2005
	4/4	L16287	CACTAGGATACCAACAAACC	162	58°	Handt et al. 1996
		H16410	GCGGGATATTGATTTCACGG			Handt et al. 1996
	6/1	L16018	AGCACCCAAAGCTAAGATTCTAATTTAAACTATT	139	56°	Knipper et al. 2013
		H16097	ATATTCATGGTGGCTGGCAGTAATGTA			Knipper et al. 2013
	6/2	L16072	GGTACCACCCAAGTATTGACTCACC	123	57°	Knipper et al. 2013
		H16144	TGGGTTTTTATGTACTACAGGTGGTCA			Knipper et al. 2013
	6/1b	L16045	TGTTCTTTCATGGGGAAGCAGATT	150	57°	Brandt et al. 2013
		H16145	TTGGGTTTTTATGTACTACAGGTGGTC			Ch. Roth unpublished
	6/3	L16122	ATTTCGTACATTACTGCCAGCCACCATGAATA	132	65°	Ch. Roth unpublished
		H16196	TAGTTGAGGGTTGATTGCTGTACTTGCTTGTAAGC			Ch. Roth unpublished
	6/4	L16178	AGTACATAAAAACCCAATCCACAT	136	50°	Ch. Roth unpublished
		H16267	GTAGGTTTGTTGGTATCCTAGTGG			Ch. Roth unpublished
	6/5	L16231	GTACAGCAATCAACCCTCAACTAT	139	50°	Ch. Roth unpublished
		H16323	CTGTAATGTGCTATGTACGGTAAA			Ch. Roth unpublished
	6/6a	L16303	GGATACCAACAAACCTACCCACCCTTAACAG	158	65°	Ch. Roth unpublished
		H16405	TGCGGGATATTGATTTCACGGAGGAT			Ch. Roth unpublished
	6/6b	L16307	GGATACCAACAAACCTACCCACCCTTAACAGTACA	158	67°	Ch. Roth unpublished
		H16402.C	TGCGGGATATTGATTTCACGGAGGATGGT			Ch. Roth unpublished
HVS II	4/1	L00034	TCTATCACCCTATTAACCACTCAC	192	58°	Haak et al. 2008
		H00177	TTAGTAAGTATGTTCGCCTGTAAT			Haak et al. 2008

4/2	L00144	CGCAGTATCTGTCTTTGATTCCTG	148	58°	Haak et al. 2008
	H00243	AAAGTGGCTGTGCAGACATTCAAT			Haak et al. 2008
4/3	L00172	ATCCTATTATTTATCGCACCTACG	204	58°	Haak et al. 2008
	H00327	TTGGCAGAGATGTGTTTAAGTGCT			Haak et al. 2008
4/4	L00274	TGTCTGCACAGCCACTTTCCACAC	174	58°	Haak et al. 2008
	H00397	AGTGCATACCGCCAAAAGATAAAA			Haak et al. 2008

Primer's names indicate forward (L-strand (L)) and reverse primer orientation (H-strand (H). Primers denote the 3’ end of the primer sequence for all HVS-I primers, the first nucleotide of the amplicon in case of the HVS-II.

Ind. Nr.	Archaeological sites to have connection with	Workarea/task in the project	HVS-I after CRS	Range HVS I	Mt Haplogroup based on HVS and coding region SNiPs	GenoCoRe 22	np 34-397 (HVR II) after CRS	Y Haplogroup
1.	all sites	worker in the lab	16126c 16163g 16186t 16189c 16264t 16294t	15997-16409	T1a		73G 152C 195C 263G 309.1C 315.1C	-
2.	all sites but Vukovar, Vinkovci, Alsónyék-Bátaszék, Balatonszárszó, Kóny, Veszprém	assistance at the sampling, worker in the lab	16209c 16311c	15997-16409	H	H	263G 315.1C	-
3.	all sites	cloning	16136c 16356c	15997-16409	U4		73G 195C 263G 309.1C 315.1C	-
4.	Alsónyék-Bátaszék	manual labour on the excavation	16221t 16291t	15997-16409	HV4a1a	HV		R1b1a2
5.	Alsónyék-Bátaszék	washwoman	16192t 16256t 16270t 16304c 16399g	15997-16409	U5a1			-
6.	Alsónyék-Bátaszék	excavator archaeologist	16126c 16294t 16296t	15998-16409	T*			-
7.	Balatonszárszó	excavator archaeologist	16293g 16311c	15998-16409	H	H		J2
8.	Balatonszárszó	excavator archaeologist	16256t 16270t 16399g	15998-16409	U5a1a			I1
9.	Balatonszárszó	manual labour on the excavation	16224c 16311c 16320t	15998-16409	K			-
10.	Balatonszárszó	manual labour on the excavation	16223t	15998-16409	H/M/N/L3	n.d.		-
11.	Balatonszárszó	manual labour on the excavation	16223t	15998-16409	H	H		-
12.	Balatonszárszó, Balatonszemes, Bölcske	anthropologist	16051g 16126c 16129C 16182C 16183C 16189c 16362c	15998-16409	U2e			-
13.	Vukovar, Vinkovci	anthropologist	16356c	16010-16383	H	H		-
14.	Vukovar, Vinkovci	anthropologist	16129a 16172c 16223t 16311c 16391a	16010-16395	I1a			R1a1
15.	Vukovar, Vinkovci	anthropologist	16287t	16036-16367	H			-
16.	Vukovar, Vinkovci	anthropologist	-	16011-16409	H/U	n.d.		-
17.	Harta-Gátőrház	anthropologist	16080g 16189c 16356c	16012-16394	H	H		-
18.	Budakeszi 4/8	anthropologist	16319a	15998-16409	H	H		-
19.	Alsónyék-Bátaszék, Balatonszárszó	assistance at the sampling	16223t 16234t 16288c 16298c 16327t 16359c	15997-16409	C5		73G 152C 249del 263G 315.2C	-
20.	Alsónyék-Bátaszék	assistance at the sampling	16260t 16327t	15999-16395	H	H		-
21.	Alósnyék-Bátaszék, Balatonszárszó	assistance at the sampling	16240g 16298c	15998-16409	V3a			E
22.	all sites but Vukovar, Vinkovci, Alsónyék-Bátaszék, Balatonszárszó, Kóny, Veszprém	assistance at the sampling	16145a 16176G 16223t 16244a 16390a	16034-16409	N1b2		152C 263G 309.2C 315.1C	R1a1
23.	all sites but Vukovar, Vinkovci, Alsónyék-Bátaszék, Balatonszárszó, Kóny, Veszprém	assistance at the sampling	-	16027-16409	H	H	146C 263G 309.2C 315.1C	I*
24.	Keszthely-Fenekpuszta	excavator archaeologist	16126C 16163G 16189C 16243C 16294T	15997-16409	T1			-
25.	Borjad Kenderföldek	excavator archaeologist	16093C 16165G 16189C 16270T	15997-16409	U5b?			-
26.	Versend, Szemely, Csapdi-T., Moragy	worker in the lab	16126c 16163g 16186t 16189c 16294	15997-16409	T1		73G 152C 195C 263G 309.1C 315.1C	-

Supplementary Table 27. HVS I-II haplotypes and Y-chromosomal haplogroups of the contributor researchers.

15.4 References to the Supplementary Tables

Abe S, Usami S, Shinkawa H, Weston MD, Overbeck LD, et al. (1998) Phylogenetic analysis of mitochondrial DNA in Japanese pedigrees of sensorineural hearing loss associated with the A1555G mutation from Swanetia. Eur J Hum Genet 6(6):563-9.

Abu-Amero KK, Hellani A, González AM, Larruga JM, Cabrera VM, et al. (2009) Saudi Arabian Y-Chromosome diversity and its relationship with nearby regions. BMC Genet 22;10:59.

Abu-Amero KK, Larruga JM, Cabrera VM, González AM. (2008) Mitochondrial DNA structure in the Arabian Peninsula. BMC Evol Biol 8:45.

Achilli A, Olivieri A, Pala M, Metspalu E, Fornarino S, et al. (2007) A Mitochondrial DNA variation of modern Tuscans supports the near eastern origin of Etruscans. Am J Hum Genet 80(4):759-68.

Adams SM, Bosch E, Balaresque PL, Ballereau SJ, Lee AC, et al. (2008)The genetic legacy of religious diversity and intolerance: Paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. Am J Hum Genet 83: 725–736.

Alfonso-Sánchez MA, Cardoso S, Martínez-Bouzas C, Peña JA, Herrera RJ, et al. (2008) Mitochondrial DNA haplogroup diversity in Basques: a reassessment based on HVI and HVII polymorphisms. Am J Hum Biol 20(2):154-64.

Alfonso-Sánchez MA, Martínez-Bouzas C, Castro A, Pen˜ a JA, Fernández- Fernández I, Herrera RJ, et al. (2006) Sequence polymorphisms of the mtDNA control region in a human isolate: The Georgians from Swanetia. J Hum Genet 51(5):429-39.

Alonso S, Flores C, Cabrera V, Alonso A, Martín P, et al. (2005) The place of the Basques in the European Y-chromosome diversity landscape Eur J Hum Genet Dec;13(12):1293-302.

Alshamali F, Brandstätter A, Zimmermann B, Parson W. (2008) Mitochondrial DNA control region variation in Dubai, United Arab Emirates. Forensic Sci Int Genet 2(1):e9-10.

Alvarez JC, Johnson DL, Lorente JA, Martinez-Espin E, Martinez-Gonzalez LJ, et al. (2007) Characterization of human control region sequences for Spanish individuals in a forensic mtDNA data set. Leg Med (Tokyo) 9(6):293-304.

Alvarez-Iglesias V, Mosquera-Miguel A, Cerezo M, Quintáns B, Zarrabeitia MT, et al. (2009) New Population and Phylogenetic Features of the Internal Variation within Mitochondrial DNA Macro-Haplogroup R0. PLoS One 4(4):e5112.

Al-Zahery N, Pala M, Battaglia V, Grugni V, Hamod MA, et al. (2011) In search of the genetic footprints of Sumerians: a survey of Y-chromosome and mtDNA variation in the Marsh Arabs of Iraq. BMC Evol Biol 11:288.

Al-Zahery N, Semino O, Benuzzi G, Magri C, Passarino G, et al. (2003) Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations. Mol Phylogenet Evol 28(3):458-72.

Amar S, Shamir A, Ovadia O, Blanaru M, Reshef A, et al. (2007) Mitochondrial DNA HV lineage increases the susceptibility to schizophrenia among Israeli Arabs. Schizophr Res 94(1-3):354-8.

Baasner A, Madea B (2000) Sequence polymorphisms of the mitochondrial DNA control region in 100 German Caucasians. J Forensic Sci 45(6):1343-8.

Baasner A, Schäfer C, Junge A, Madea B (1998) Polymorphic sites in human mitochondrial DNA control region sequences: population data and maternal inheritance. Forensic Sci Int 98(3):169-78.

Babalini C, Martínez-Labarga C, Tolk HV, Kivisild T, Giampaolo R, et al. (2005) The population history of the Croatian linguistic minority of Molise (southern Italy): a maternal view. Eur J Hum Genet 13(8):902-12.

Balanovsky O, Rootsi S, Pshenichnov A, Kivisild T, Churnosov M, et al. (2008) Two sources of the Russian patrilineal heritage in their Eurasian context. Am J Hum Genet 82: 236-250.

Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, et al. (2009) Y-chromosomal evidence of the cultural diffusion of agriculture in southeast Europe. Eur J Hum Genet 17:820–830.

Barać L, Pericić M, Klarić IM, Rootsi S, Janićijević B, et al. (2003) Y chromosomal heritage of Croatian population and its island isolates. Eur J Hum Genet 11: 535-542.

Behar DM, Garrigan D, Kaplan ME, Mobasher Z, Rosengarten D, et al. (2004) Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations Hum Genet Mar;114(4):354-65.

Beleza S, Gusmao L, Lopes A, Alves C, Gomes I, et al. (2005) Micro-phylogeographic and demographic history of Portuguese male lineages. Ann Hum Genet 70: 181-194.

Belledi M, Poloni ES, Casalotti R, Conterio F, Mikerezi I, et al. (2000) Maternal and paternal lineages in Albania and the genetic structure of Indo-European populations. Eur J Hum Genet 8(7):480-6.

Belyaeva O, Bermisheva M, Khrunin A, Slominsky P, Bebyakova N, et al. (2003) Mitochondrial DNA variations in Russian and Belorussian populations. Hum Biol 75(5):647-60.

Bertranpetit J, Sala J, Calafell F, Underhill PA, Moral P, et al. (1995) Human mitochondrial DNA variation and the origin of Basques. Ann Hum Genet 59(Pt 1):63-81.

Betty DJ, Chin-Atkins AN, Croft L, Sraml M, Easteal S (1996) Multiple independent origins of the COII/tRNA(Lys) intergenic 9-bp mtDNA deletion in aboriginal Australians. Am J Hum Genet 58(2):428-33.

Bickle P, Alasdair W et al. (2013) The First Farmers of Central Europe: Diversity in LBK Lifeways. Oxford Books. 680 p. ISBN: 9781842175309

Bini C, Ceccardi S, Luiselli D, Ferri G, Pelotti S, et al. (2003) Different informativeness of the three hypervariable mitochondrial DNA regions in the population of Bologna (Italy). Forensic Sci Int 135(1):48-52.

Bogácsi-Szabó E, Kalmár T, Csányi B, Tömöry G, Czibula A, et al. (2005) Mitochondrial DNA of ancient Cumanians: culturally Asian steppe nomadic immigrants with substantially more western Eurasian mitochondrial DNA lineages. Hum Biol 77(5):639-62.

Bosch E, Calafell F, González-Neira A, Flaiz C, Mateu E, et al. (2006) Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns Ann Hum Genet Jul;70(Pt 4):459-87.

Brakez Z, Bosch E, Izaabel H, Akhayat O, et al. (2001) Human mitochondrial DNA sequence variation in the Moroccan population of the Souss area. Ann Hum Biol 28(3):295-307.

Brandstätter A, Klein R, Duftner N, Wiegand P, Parson W (2006) Application of a quasi-median network analysis for the visualization of character conflicts to a population sample of mitochondrial DNA control region sequences from southern Germany (Ulm). Int J Legal Med 120(5):310-4.

Brandstätter A, Niederstätter H, Pavlic M, Grubwieser P, Parson W. (2007) Generating population data for the EMPOP database - an overview of the mtDNA sequencing and data evaluation processes considering 273 Austrian control region sequences as example. Forensic Sci Int 166(2-3):164-75.

Brión M, Sanchez JJ, Balogh K,Thacker C, Blanco-Verea A, et al. (2006) Analysis of 29 Y-chromosome SNPs in a single multiplex useful to predict the geographic origin of male lineages International Congress Series 1288. 13– 15.

Cadenas AM, Zhivotovsky LA, Cavalli-Sforza LL, Underhill PA, Herrera RJ (2007) Y-chromosome diversity characterizes the Gulf of Oman. Eur J Hum Genet 16, 374–386.

Calafell F, Underhill P, Tolun A, Angelicheva D, Kalaydjieva L (1996) From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. Ann Hum Genet 60(Pt 1):35-49.

Cali F, Le Roux MG, D'Anna R, Flugy A, De Leo G, et al. (2001) MtDNA control region and RFLP data for Sicily and France. Int J Legal Med 114(4-5):229-31.

Capelli C, Brisighelli F, Scarnicci F, Arredi B, Caglia' A, et al. (2007) Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter. Mol Phylogenet Evol 44: 228-239.

Capelli C, Redhead N, Abernethy JK, Gratrix F, Wilson JF, et al. (2003) A Y chromosome census of the British Isles. Curr Biol May 27;13(11):979-84.

Cardoso S, Alfonso-Sánchez MA, Valverde L, Odriozola A, Pérez-Miranda AM, et al. (2011) The maternal legacy of Basques in northern navarre: New insights into the mitochondrial DNA diversity of the Franco-Cantabrian area. Am J Phys Anthropol 145(3):480-8.

Cardoso S, Zarrabeitia MT, Valverde L, Odriozola A, Alfonso-Sánchez MA, et al. (2010) Variability of the Entire Mitochondrial DNA Control Region in a Human Isolate from the Pas Valley (Northern Spain). J Forensic Sci 55(5):1196-201.

Casas MJ, Hagelberg E, Fregel R, Larruga JM, González AM. (2006) Human mitochondrial DNA diversity in an archaeological site in al-Andalus: genetic impact of migrations from North Africa in medieval Spain. Am J Phys Anthropol 131(4):539-51.

Cerný V, Mulligan CJ, Rídl J, Zaloudková M, Edens CM, et al. (2008) Regional differences in the distribution of the sub-Saharan, West Eurasian, and South Asian mtDNA lineages in Yemen. Am J Phys Anthropol 136(2):128-37.

Chaix R, Quintana-Murci L, Hegay T, Hammer MF, Mobasher Z, et al. (2007) From social to genetic structures in central Asia. Curr Biol 17(1):43-8.

Chandler H, Sykes B, J. Zilhão. Using ancient DNA to examine genetic continuity at the Mesolithic-Neolithic transition in Portugal. in Arias P, Ontañón R, García-Moncó C (eds.), Actas dell III Congreso del Neolítico en la Península Ibérica, Santander, Monografías del Instituto internacional de Investigaciones Prehistóricas de Cantabria 1 2005: 781-86

Chandler, H. Using Ancient DNA to Link Culture and Biology in Human Populations (University of Oxford, 2003)

Chen F, Wang SY, Zhang RZ, Hu YH, Gao GF, et al. (2008) Analysis of mitochondrial DNA polymorphisms in Guangdong Han Chinese. Forensic Sci Int Genet 2(2):150-3.

Cherni L, Fernandes V, Pereira JB, Costa MD, Goios A, et al. (2009) Post-last glacial maximum expansion from Iberia to North Africa revealed by fine characterization of mtDNA H haplogroup in Tunisia. Am J Phys Anthropol 139(2):253-60.

Cherni L, Loueslati BY, Pereira L, Ennafaâ H, et al. (2005) Female gene pools of Berber and Arab neighboring communities in central Tunisia: microstructure of mtDNA variation in North Africa. Hum Biol 77(1):61-70.

Cinnioğlu C, King R, Kivisild T, Kalfoğlu E, Atasoy S, et al. (2004) Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet 114:127–148.

Comas D, Calafell F, Bendukidze N, Fañanás L, Bertranpetit J. (2000) Georgian and kurd mtDNA sequence analysis shows a lack of correlation between languages and female genetic lineages. Am J Phys Anthropol 112(1):5-16.

Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bertranpetit J (1996) Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations. Mol Biol Evol 13(8):1067-77.

Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bosch E, et al. (1998) Trading genes along the silk road: mtDNA sequences and the origin of central Asian populations. Am J Hum Genet 63(6):1824-38.

Comas D, Plaza S, Wells RS, Yuldaseva N, Lao O, et al. (2004) Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. Eur J Hum Genet 12(6):495-504.

Côrte-Real HB, Macaulay VA, Richards MB, Hariti G, Issad MS, et al. (1996) Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. Ann Hum Genet 60(Pt 4):331-50.

Coudray C, Olivieri A, Achilli A, Pala M, Melhaoui M, et al. (2009) The complex and diversified mitochondrial gene pool of Berber populations. Ann Hum Genet 73(2):196-214.

Crespillo M, Luque JA, Paredes M, Fernández R, Ramírez E, et al. (2000) Mitochondrial DNA sequences for 118 individuals from northeastern Spain. Int J Legal Med 114(1-2):130-2.

Crubézy E, Amory S, Keyser C, Bouakaze C, Bodner M, et al. (2010) Human evolution in Siberia: from frozen bodies to ancient DNA. BMC Evol Biol 10:25.

Csányi B, Bogácsi-Szabó E, Tömöry G, Czibula A, Priskin K et al. (2008) Y-chromosome analysis of ancient Hungarian and two modern Hungarian-speaking populations from the Carpathian Basin. Ann Hum Genet Jul;72(Pt 4):519-34.

De Benedetto G, Nasidze IS, Stenico M, Nigro L, Krings M, Lanzinger M, Vigilant L, Stoneking M, Pääbo S, Barbujani G. Mitochondrial DNA sequences in prehistoric human remains from the Alps. Eur J Hum Genet. 2000; 8(9):669-77

Deguilloux MF, Soler L, Pemonge MH, Scarre C, Joussaume R, Laporte L. News from the west: ancient DNA from a French megalithic burial chamber. Am J Phys Anthropol. 2011; 144(1):108-18.

Delsate D, Guinet JM, Saverwyns S. De l'ocre sur le crâne mésolithique (haplogroupe U5a) de Reuland-Loschbour (Grand-Duché de Luxembourg). Bull. Soc. Pré hist. Luxembourgeoise. 2009; 31:7–30.

Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, Koshel S, V Zaporozhchenko, Gronenborn D, Moiseyev V, Kolpakov E, Shumkin V, Alt KW, E Balanovska E, Cooper A, Haak W, The Genographic Consortium. Ancient Mitochondrial DNA Unravels Complex Human Population History in North East Europe. unpublished

Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, Koshel S, V Zaporozhchenko, Gronenborn D, Moiseyev V, Kolpakov E, Shumkin V, Alt KW, E Balanovska E, Cooper A, Haak W, The Genographic Consortium. Ancient DNA Reveals Prehistoric Gene-Flow From Siberia in the Complex Human Population History of North East Europe. PLoS Genet. 2013; 9(2):e1003296s.

Derbeneva OA, Starikovskaya EB, Wallace DC, Sukernik RI (2002) Traces of early Eurasians in the Mansi of northwest Siberia revealed by mitochondrial DNA analysis. Am J Hum Genet 70(4):1009-14.

Derbeneva OA, Starikovskaya EB, Wallace DC, Sukernik RI (2002) Traces of early Eurasians in the Mansi of northwest Siberia revealed by mitochondrial DNA analysis. Am J Hum Genet 70(4):1009-14.

Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Dambueva I, et al. (2007) Phylogeographic analysis of mitochondrial DNA in northern Asian populations. Am J Hum Genet 81(5):1025-41.

Derenko MV, Grzybowski T, Malyarchuk BA, Dambueva IK, Denisova GA, et al. (2003) Diversity of mitochondrial DNA lineages in South Siberia. Ann Hum Genet 67(Pt 5):391-411.

Devor EJ, Abdurakhmonov I, Zlojutro M, Millis MP, Galbraith JJ, et al. (2009) A Gene Flow at the Crossroads of Humanity: mtDNA Sequence Diversity and Alu Insertion Polymorphism Frequencies in Uzbekistan. Open Genomics J 2: 1-11.

Di Benedetto G, Ergüven A, Stenico M, Castrì L, Bertorelle G, et al. (2001) DNA diversity and population admixture in Anatolia. Am J Phys Anthropol 115(2):144-56.

Di Gaetano C, Cerutti N, Crobu F, Robino C, Inturri S, et al. (2009) Differential Greek and northern African migrations to Sicily are supported by genetic evidence from the Y chromosome. Eur J Hum Genet Jan;17(1):91-9.

Di Rienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci U S A 88(5):1597-601.

Dimo-Simonin N, Grange F, Taroni F, Brandt-Casadevall C, Mangin P (2000) Forensic evaluation of mtDNA in a population from south west Switzerland. Int J Legal Med 113(2):89-97.

Dubut V, Chollet L, Murail P, Cartault F, Béraud-Colomb E, et al. (2004) mtDNA polymorphisms in five French groups: importance of regional sampling. Eur J Hum Genet 12(4):293-300.

Dulik MC, Osipova LP, Schurr TG (2011) Y-Chromosome Variation in Altai Kazakhs Reveals a Common Paternal Gene Pool for Kazakhs and the Influence of Mongolian Expansions PLoS One Mar 11;6(3):e17548.

El-Sibai M, Platt DE, Haber M, Xue Y, Youhanna SC, et al. (2009) Geographical Structure of the Y-chromosomal Genetic Landscape of the Levant: A coastal-inland contrast. Ann Hum Genet 73:658-681.

Ennafaa H, Fregel R, Khodjet-El-Khil H, González AM, Mahmoudi HA, et al. (2011) Mitochondrial DNA and Y-chromosome microstructure in Tunisia. J Hum Genet 56(10):734-41.

Ermini L, Olivieri C, Rizzi E, Corti G, Bonnal R, Soares P, Luciani S, Marota I, De Bellis G, Richards MB, Rollo F. Complete mitochondrial genome sequence of the Tyrolean Iceman. Curr Biol. 2008; 18(21):1687-93.

Fadhlaoui-Zid K, Plaza S, Calafell F, Ben Amor M, Comas D, et al. (2004) Mitochondrial DNA heterogeneity in Tunisian Berbers. Ann Hum Genet 68(Pt 3):222-33.

Fadhlaoui-Zid K, Rodríguez-Botigué L, Naoui N, Benammar-Elgaaied A, Calafell F, et al. (2011) Mitochondrial DNA structure in North Africa reveals a genetic discontinuity in the Nile Valley. Am J Phys Anthropol 145(1):107-17.

Falchi A, Giovannoni L, Calo CM, Piras IS, Moral P, et al. (2006) Genetic history of some western Mediterranean human isolates through mtDNA HVR1 polymorphisms. J Hum Genet 51(1):9-14.

Fechner A, Quinque D, Rychkov S, Morozowa I, Naumova O, et al. (2008) Boundaries and clines in the West Eurasian Y-Chromosom landscape: insights from European part of Russia. Am J Phys Anthropol Sep;137 (1):41-7.

Fedorova SA, Bermisheva MA, Villems R, Maksimova NR, Khusnutdinova EK (2003) Analysis of mitochondrial DNA haplotypes in yakut population. Mol Biol (Mosk) 37(4):643-53.

Finnilä S, Lehtonen MS, Majamaa K (2001) Phylogenetic network for European mtDNA. Am J Hum Genet 68(6):1475-84.

Firasat S, Khaliq S, Mohyuddin A, Papaioannou M, Tyler-Smith C, et al. (2007) Y-chromosomal evidence for a limited Greek contribution to the Pathan population of Pakistan. Eur J Hum Genet 15, 121-126.

Flores C, Maca-Meyer N, González AM, Oefner PJ, Shen P, et al. (2004) Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. Eur J Hum Genet Oct;12(10):855-63.

Flores C, Maca-Meyer N, Larruga JM, Cabrera VM, Karadsheh N, et al. (2005) Isolates in a corridor of migrations: a high-resolution analysis of Y-chromosome variation in Jordan J Hum Genet 50: 435-441.

Forster P, Calì F, Röhl A, Metspalu E, D'Anna R, et al. (2002) Continental and subcontinental distributions of mtDNA control region types. Int J Legal Med 116(2):99-108.

Francalacci P, Bertranpetit J, Calafell F, Underhill PA (1996) Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe. Am J Phys Anthropol 100(4):443-60.

Francalacci P, Morelli L, Underhill PA, Lillie AS, Passarino G, et al. (2003) Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. Am J Phys Anthropol 121:270–279.

Fucharoen G, Fucharoen S, Horai S (2001) Mitochondrial DNA polymorphisms in Thailand. J Hum Genet 46(3):115-25.

Gamba C, Fernández E, Tirado M, Deguilloux MF, Pemonge MH, Utrilla P, Edo M, Molist M, Rasteiro R, Chikhi L, Arroyo-Pardo E. Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. Mol Ecol. 2012;21(1):45-56

García O, Fregel R, Larruga JM, Alvarez V, Yurrebaso I, et al. (2010) Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. Heredity (Edinb) 2011 Jan;106(1):37-45.

Gokcumen O, Dulik MC, Pai AA, Zhadanov SI, et al. (2008) Genetic Variation in the Enigmatic Altaian Kazakhs of South-Central Russia: Insights into Turkic Population History. Am J Phys Anthropol 136(3):278-93.

Goncalves R, Freitas A, Branco M, Rosa A, Fernandes AT, et al. (2005) Y-chromosome lineages from Portugal, Madeira and Acores record elements of Sephardim and Berber ancestry. Ann Hum Genet 69: 443-454.

González AM, Brehm A, Pérez JA, Maca-Meyer N, Flores C, et al. (2003) Mitochondrial DNA affinities at the Atlantic fringe of Europe. Am J Phys Anthropol 120(4):391-404.

González AM, Cabrera VM, Larruga JM, Tounkara A, Noumsi G, et al. (2006) Mitochondrial DNA variation in Mauritania and Mali and their genetic relationship to other Western Africa populations. Ann Hum Genet 70(Pt 5):631-57.

González AM, Karadsheh N, Maca-Meyer N, Flores C, Cabrera VM, et al. (2008) Mitochondrial DNA variation in Jordanians and their genetic relationship to other Middle East populations. Ann Hum Biol 35(2):212-31.

Goodacre S, Helgason A, Nicholson J, Southam L, Ferguson L, et al. (2005) Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. Heredity 95(2):129-35.

Grugni V, Battaglia V, Hooshiar Kashani B, Parolo S, Al-Zahery N, et al. (2012) Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians.PLoS One 7(7):e41252. Epub 2012 Jul 18.

Grzybowski T, Malyarchuk BA, Derenko MV, Perkova MA, Bednarek J, et al. (2007) Complex interactions of the Eastern and Western Slavic populations with other European groups as revealed by mitochondrial DNA analysis. Forensic Sci Int Genet 1(2):141-7.

Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, Der Sarkissian CS, Brandt G, Schwarz C, Nicklisch N, Dresely V, Fritsch B, Balanovska E, Villems R, Meller H, Alt KW, Cooper A; Members of the Genographic Consortium. (2010) Ancient DNA from European early neolithic farmers reveals their near eastern affinities. PLoS Biol. 2010; 8(11):e1000536.

Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V, Hawkesworth C, Pike AW, Meller H, Alt KW. (2008) Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. Proc Natl Acad Sci U S A. 2008; 105(47):18226-31.

Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tänzer M, Villems R, Renfrew C, Gronenborn D, Alt KW, Burger J. (2005) Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. Science. 2005; 310(5750):1016-8.

Haber M, Platt DE, Ashrafian Bonab M, Youhanna SC, Soria-Hernanz DF, et al. (2012) Afghanistan's Ethnic Groups Share a Y-Chromosomal Heritage Structured by Historical Events PLoS One 7(3):e34288.

Haber M, Youhanna SC, Balanovsky O, Saade S, Martínez-Cruz B, et al. (2011) mtDNA Lineages Reveal Coronary Artery Disease-Associated Structures in the Lebanese Population. Ann Hum Genet 2011; 76(1):1-8.

Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, et al. (1994) Molecular genetic analyses of the Tyrolean Ice Man Science 264(5166):1775-8.

Harich N, Costa MD, Fernandes V, Kandil M, Pereira JB, et al. (2010) The trans-Saharan slave trade - clues from interpolation analyses and high-resolution characterization of mitochondrial DNA lineages. BMC Evol Biol 10:138.

Harvey M, Gordon K, Owens K, Lee M, King MC MtDNA sequences from Balkan populations.

Hedman M, Brandstätter A, Pimenoff V, Sistonen P, Palo JU, et al. (2007) Finnish mitochondrial DNA HVS-I and HVS-II population data. Forensic Sci Int. 172(2-3):171-8.

Helgason A, Hickey E, Goodacre S, Bosnes V, Stefánsson K, et al. (2001) MtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. Am J Hum Genet 68(3):723-37.

Helgason A, Nicholson G, Stefánsson K, Donnelly P (2003) A reassessment of genetic diversity in Icelanders: strong evidence from multiple loci for relative homogeneity caused by genetic drift. Ann Hum Genet 67(Pt 4):281-97.

Helgason A, Sigurethardóttir S, Gulcher JR, Ward R, Stefánsson K (2000) MtDNA and the origin of the Icelanders: deciphering signals of recent population history. Am J Hum Genet 66(3):999-1016.

Hervella M, Izagirre N, Alonso S, Fregel R, Alonso A, Cabrera VM, de la Rúa C. Ancient DNA from Hunter-Gatherer and Farmer Groups from Northern Spain Supports a Random Dispersion Model for the Neolithic Expansion into Europe. PLoS one. 2012; 7(4):e34417.

Hofmann S, Jaksch M, Bezold R, Mertens S, Aholt S, et al. (1997) Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease. Hum Mol Genet 6(11):1835-46.

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, et al. (1996) mtDNA polymorphism in East Asian Populations, with special reference to the peopling of Japan. Am J Hum Genet 59(3):579-90.

Imaizumi K, Parsons TJ, Yoshino M, Holland MM (2002) A new database of mitochondrial DNA hypervariable regions I and II sequences from 162 Japanese individuals. Int J Legal Med 116(2):68-73.

Irwin J, Egyed B, Saunier J, Szamosi G, O'Callaghan J, et al. (2007) Hungarian mtDNA population databases from Budapest and the Baranya county Roma. Int J Legal Med 121(5):377-83.

Irwin J, Saunier J, Strouss K, Paintner C, Diegoli T, et al. (2008) Mitochondrial control region sequences from northern Greece and Greek Cypriots. Int J Legal Med 122(1):87-9.

Irwin JA, Ikramov A, Saunier J, Bodner M, Amory S, et al. (2010) The mtDNA composition of Uzbekistan: a microcosm of Central Asian patterns. Int J Legal Med 124(3):195-204.

Irwin JA, Saunier JL, Beh P, Strouss KM, Paintner CD, Parsons TJ. (2009) Mitochondrial DNA control region variation in a population sample from Hong Kong, China. Forensic Sci Int Genet. 2009; 3(4):e119-25.

Irwin JA, Saunier JL, Strouss KM, Diegoli TM, Sturk KA, et al. (2008) Mitochondrial control region sequences from a Vietnamese population sample. Int J Legal Med 122(3):257-9.

Kaessmann H, Zöllner S, Gustafsson AC, Wiebe V, Laan M, et al. (2002) Extensive linkage disequilibrium in small human populations in Eurasia. Am J Hum Genet 70(3):673-85.

Kaldma K, Kivisild T, Parik J, Villems R. Mitochondrial DNA lineages in Ossetes.

Karachanak S, Carossa V, Nesheva D, Olivieri A, Pala M, et al. (2011) Bulgarians vs the other European populations: a mitochondrial DNA perspective. Int J Legal Med. 2012; 126(4):497-503.

Karafet T, Xu L, Du R, Wang W, Feng S, et al. (2001) Paternal Population History of East Asia: Sources, Patterns, and Microevolutionary Processes Am J Hum Genet 69:615–628.

Karlsson AO, Wallerström T, Götherström A, Holmlund G (2006) Y-chromosome diversity in Sweden - a long-time perspective. Eur J Hum Genet 14: 963-970.

Kasperaviciūte D, Kucinskas V, Stoneking M (2004) Y chromosome and mitochondrial DNA variation in Lithuanians. Ann Hum Genet 68(Pt 5):438-52.

Keyser C, Bouakaze C, Crubézy E, Nikolaev VG, Montagnon D, Reis T, Ludes B. Ancient DNA provides new insights into the history of south Siberian Kurgan people. Hum Genet. 2009; 126(3):395-410.

Kharkov VN, Stepanov VA, Medvedeva OF, Spiridonova MG, Voevoda MI, et al. (2007) Gene pool differences between northern and southern Altai inferred from the data on Y-chromosomal haplogroups. Russ J Genet 43:675–687.

King RJ, Ozcan SS, Carter T, Kalfoğlu E, Atasoy S, et al. (2008) Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. Ann Hum Genet 72: 205-214.

Kittles RA, Bergen AW, Urbanek M, Virkkunen M, Linnoila M, et al. (1999) Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. Am J Phys Anthropol 108(4):381-99.

Kivisild T, Reidla M, Metspalu E, Rosa A, Brehm A, et al. (2004) Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. Am J Hum Genet 75(5):752-70.

Kivisild T, Tolk HV, Parik J, Wang Y, Papiha SS, et al. (2002) The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol 19(10):1737-51.

Kolman CJ, Sambuughin N, Bermingham E (1996) Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 142(4):1321-34.

Komso D (2006) The Mesolithic in Croatia. Opusc.archaeol. 30, 55-92.

Kong QP, Yao YG, Liu M, Shen SP, Chen C, et al. (2003) Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. Hum Genet 113(5):391-405.

Kornienko IV, Ivanov PL. Hypervariable region II data from the Russian Federation.

Kouvatsi A, Karaiskou N, Apostolidis A, Kirmizidis G (2001) Mitochondrial DNA sequence variation in Greeks. Hum Biol 73(6):855-69.

Koyama H, Iwasa M, Maeno Y, Tsuchimochi T, Isobe I, et al. (2002) Mitochondrial sequence haplotype in the Japanese population. Forensic Sci Int 125(1):93-6.

Krings M, Salem AE, Bauer K, Geisert H, Malek AK, et al. (1999) mtDNA analysis of Nile River Valley populations: a genetic corridor or a barrier to migration?. Am J Hum Genet 64(4):1166-76.

Lacan M, Keyser C, Ricaut FX, Brucato N, Duranthon F, Guilaine J, Crubézy E, Ludes B. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. Proc Natl Acad Sci U S A. 2011; 108(24):9788-91

Lacan M, Keyser C, Ricaut FX, Brucato N, Tarrús J, Bosch A, Guilaine J, Crubézy E, Ludes B. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. Proc Natl Acad Sci U S A. 2011; 108(45):18255-9.

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, et al. (1996) The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet 58(6):1309-22.

Lalueza-Fox C, Sampietro ML, Gilbert MT, Castri L, Facchini F, Pettener D, Bertranpetit J. Unravelling migrations in the steppe: mitochondrial DNA sequences from ancient central Asians. Proc Biol Sci. 2004; 271(1542):941-947.

Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, et al. (2006) Regional diff erences among the Finns: a Y-chromosomal perspective. Gene Jul 19;376(2):207-15.

Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, et al. (2008) Migration Waves to the Baltic Sea Region. Ann Hum Genet 72(Pt 3):337-48.

Larruga JM, Díez F, Pinto FM, Flores C, González AM (2001) Mitochondrial DNA characterisation of European isolates: the Maragatos from Spain. Eur J Hum Genet 9(9):708-16.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S. et al. (2014) Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513, 409–413.Lee EJ, Krause-Kyora B, Rinne C, Schütt R, Harder M, et al. (2013) Ancient DNA insights from the Middle Neolithic in Germany. Archaeol Anthropol Sci. doi:DOI 10.1007/s12520-013-0173-0.

Lee EJ, Makarewicz C, Renneberg R, Harder M, Krause-Kyora B, Müller S, Ostritz S, Fehren-Schmitz L, Schreiber S, Müller J, von Wurmb-Schwark N, Nebel A, Nebel A. (2012) Emerging genetic patterns of the european neolithic: Perspectives from a late neolithic bell beaker burial site in Germany. Am J Phys Anthropol. 2012; doi: 10.1002/ajpa.22074.

Lee HY, Yoo JE, Park MJ, Chung U, Shin KJ (2006) Mitochondrial DNA control region sequences in Koreans: identification of useful variable sites and phylogenetic analysis for mtDNA data quality control. Int J Legal Med 120(1):5-14.

Lehocký I, Baldovic M, Kádasi L, Metspalu E (2008) A database of mitochondrial DNA hypervariable regions I and II sequences of individuals from Slovakia. Forensic Sci Int Genet 2(4):e53-9.

Li H, Cai X, Winograd-Cort ER, Wen B, Cheng X, Qin Z, et al. (2007) Mitochondrial DNA diversity and population differentiation in southern East Asia. Am J Phys Anthropol 134(4):481-8.

Liu C, Wang SY, Zhao M, Xu ZY, Hu YH, et al. (2011) Mitochondrial DNA polymorphisms in Gelao ethnic group residing in Southwest China. Forensic Sci Int Genet 5(1):e4-10.

Loueslati BY, Cherni L, Khodjet-Elkhil H, Ennafaa H, Pereira L, et al. (2006) Islands inside an island: reproductive isolates on Jerba island. Am J Hum Biol 18(1):149-53.

Luis JR, Rowold DJ, Regueiro M, Caeiro B, Cinnioğlu C, et al. (2004) The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. Am J Hum Genet 74: 32–544.

Lutz S, Weisser HJ, Heizmann J, Pollak S (1998) Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Legal Med 111(2):67-77.

Mabuchi T, Susukida R, Kido A, Oya M (2007) Typing the 1.1 kb control region of human mitochondrial DNA in Japanese individuals. J Forensic Sci 52(2):355-63.

Maca-Meyer N, Sánchez-Velasco P, Flores C, Larruga JM, González AM, et al. (2003) Y chromosome and mitochondrial DNA characterization of Pasiegos, a human isolate from Cantabria (Spain). Ann Hum Genet 67(Pt 4):329-39.

Maca-Meyer N., Sanchez-Velasco P., Flores C., Larruga J.M., Gonzalez A.M., et al. (2003) Y chromosome and mitochondrial DNA characterization of Pasiegos, a human isolate from Cantabria (Spain). Ann Hum Genet Jul;67(Pt 4):329-39.

Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, et al. (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am J Hum Genet 64(1):232-49.

Malmström H, Gilbert MT, Thomas MG, Brandström M, Storå J, Molnar P, Andersen PK, Bendixen C, Holmlund G, Götherström A, Willerslev E. (2009) Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. Curr Biol. 2009; 19(20):1758-62.

Malyarchuk B, Derenko M, Grzybowski T, Lunkina A, Czarny J, et al. (2004) Differentiation of mitochondrial DNA and Y chromosomes in Russian populations. Hum Biol 76(6):877-900.

Malyarchuk BA, Derenko MV (2001) Mitochondrial DNA variability in Russians and Ukrainians: implication to the origin of the Eastern Slavs. Ann Hum Genet 65(Pt 1):63-78.

Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Drobnic K, et al. (2003) Mitochondrial DNA variability in Bosnians and Slovenians. Ann Hum Genet 67(Pt 5):412-25.

Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Woźniak M, et al. (2002) Mitochondrial DNA variability in Poles and Russians. Ann Hum Genet 66(Pt 4):261-83.

Malyarchuk BA, Perkova MA, Derenko MV, Vanecek T, Lazur J, et al. (2008) Mitochondrial DNA Variability in Slovaks, with Application to the Roma Origin. Ann Hum Genet 72(Pt 2):228-40.

Maliarchuk BA, Derenko MV (2008) Gene Pool Structure of Russian Populations from the European Part of Russia Inferred from the Data on Y Chromosome Haplogroups Distribution. Genetika Feb;44(2):226-31.

Malyarchuk BA, Vanecek T, Perkova MA, Derenko MV, Sip M (2006) Mitochondrial DNA variability in the Czech population, with application to the ethnic history of Slavs. Hum Biol 78(6):681-96.

Mannino MA, Catalano G, Talamo S, Mannino G, Di Salvo R, et al. (2012) Origin and Diet of the Prehistoric Hunter-Gatherers on the Mediterranean Island of Favignana (Ègadi Islands, Sicily). PLoS ONE 7(11): e49802. doi:10.1371/journal.pone.0049802

Marjanovic D, Fornarino S, Montagna S, Primorac D, Hadziselimovic R, et al. (2005) The peopling of modern Bosnia-Herzegovina: Y-chromosome haplogroups in the three main ethnic groups. Ann Hum Genet 69: 757– 763.

Martinez L, Mirabal S, Luis JR, Herrera RJ (2008) Middle Eastern and European mtDNA lineages characterize populations from eastern Crete. Am J Phys Anthropol 137(2):213-23.

Martinez L, Underhill PA, Zhivotovsky LA, Gayden T, Moschonas NK, et al. (2007) Paleolithic Y-haplogroup heritage predominates in a Cretan highland plateau. Eur J Hum Genet 15, 485-493.

Maruyama S, Minaguchi K, Saitou N (2003) Sequence polymorphisms of the mitochondrial DNA control region and phylogenetic analysis of mtDNA lineages in the Japanese population. Int J Legal Med 117(4):218-25.

McEvoy B, Richards M, Forster P, Bradley DG (2004) The Longue Duree of genetic ancestry: multiple genetic marker systems and Celtic origins on the Atlantic facade of Europe. Am J Hum Genet 75(4):693-702.

Meinilä M, Finnilä S, Majamaa K (2001) Evidence for mtDNA admixture between the Finns and the Saami. Hum Hered 52(3):160-70.

Melchior L, Lynnerup N, Siegismund HR, Kivisild T, Dissing J. Genetic diversity among ancient Nordic populations. PLoS One. 2010; 5(7):e11898.

Melton T, Clifford S, Martinson J, Batzer M, Stoneking M (1998) Genetic evidence for the proto-Austronesian homeland in Asia: mtDNA and nuclear DNA variation in Taiwanese aboriginal tribes. Am J Hum Genet 63(6):1807-23.

Mergen H, Oner R, Oner C (2004) Mitochondrial DNA sequence variation in the Anatolian Peninsula (Turkey). J Genet 83(1):39-47.

Messina F, Scorrano G, Labarga CM, Rolfo MF, Rickards O (2010) Mitochondrial DNA variation in an isolated area of Central Italy. Ann Hum Biol 37(3):385-402.

Metspalu E, Kivisild T, Kaldma K, Reidla M, Villems R. Mitochondrial DNA lineages and the history of the Roms (Gypsies).

Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, et al. (2004) Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. BMC Genet 5:26.

Mikkelsen M, Sørensen E, Rasmussen EM, Morling N (2010) Mitochondrial DNA HV1 and HV2 variation in Danes. Forensic Sci Int Genet 4(4):e87-8.

Mirabal S, Regueiro M, Cadenas AM, Cavalli-Sforza LL, Underhill PA, et al. (2009) Y-Chromosome distribution within the geolinguistic landscape of northwestern Russia. Eur J Hum Genet 17:1260–1273.

Mogentale-Profizi N, Chollet L, Stévanovitch A, Dubut V, Poggi C, et al. (2001) Mitochondrial DNA sequence diversity in two groups of Italian Veneto speakers from Veneto. Ann Hum Genet 65(Pt 2):153-66.

Moore LT, McEvoy B, Cape E, Simms K, Bradley DG (2006) A Y-chromosome signature of hegemony in Gaelic Ireland. Am J Hum Genet Feb;78(2):334-8.

Morozova I, Evsyukov A, Kon'kov A, Grosheva A, Zhukova O, et al. (2012) Russian ethnic history inferred from mitochondrial DNA diversity. Am J Phys Anthropol 147: 341–351. doi: 10.1002/ajpa.21649

Nasidze I, Quinque D, Rahmani M, Alemohamad SA, Stoneking M (2008) Close genetic relationship between Semitic-speaking and Indo-European-speaking groups in Iran. Ann Hum Genet 72(Pt 2):241-52.

Nasidze I, Quinque D, Rahmani M,Alemohamad SA, Asadova P, et al. (2009) mtDNA and Y-Chromosome Variation in the Talysh of Iran and Azerbaijan. Am J Phys Anthropol Jan;138(1):82-9.

Nikitin AG, Newton JR, Potekhina ID. Mitochondrial haplogroup C in ancient mitochondrial DNA from Ukraine extends the presence of East Eurasian genetic lineages in Neolithic Central and Eastern Europe. J Hum Genet. 2012 Jun 7. doi: 10.1038/jhg.2012.69.

Nishimaki Y, Sato K, Fang L, Ma M, Hasekura H, et al. (1999) Sequence polymorphism in the mtDNA HV1 region in Japanese and Chinese. Leg Med (Tokyo) 1(4):238-49.

Non AL, Al-Meeri A, Raaum RL, Sanchez LF, Mulligan CJ (2011) Mitochondrial DNA Reveals Distinct Evolutionary Histories for Jewish Populations in Yemen and Ethiopia. Am J Phys Anthropol 144(1):1-10.

Oota H, Kitano T, Jin F, Yuasa I, Wang L, et al. (2002) Extreme mtDNA homogeneity in continental Asian populations. Am J Phys Anthropol 118(2):146-53.

Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. (2001) Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29(1):20-1.

Opdal SH, Rognum TO, Vege A, Stave AK, Dupuy BM, et al. (1998) Increased number of substitutions in the D-loop of mitochondrial DNA in the sudden infant death syndrome. Acta Paediatr 87(10):1039-44.

Orekhov V, Poltoraus A, Zhivotovsky LA, Spitsyn V, Ivanov P, et al. (1999) Mitochondrial DNA sequence diversity in Russians. FEBS Lett. 445(1):197-201.

Ottoni C, Martínez-Labarga C, Loogväli EL, Pennarun E, Achilli A, et al. (2009) First genetic insight into Libyan Tuaregs: a maternal perspective. Ann Hum Genet 73(Pt 4):438-48.

Ottoni C, Martinez-Labarga C, Vitelli L, Scano G, Fabrini E, et al. (2009) Human mitochondrial DNA variation in Southern Italy. Ann Hum Biol 36(6):785-811.

Pakendorf B, Novgorodov IN, Osakovskij VL, Danilova AP, Protod'jakonov AP, et al. (2006) Investigating the effects of prehistoric migrations in Siberia: genetic variation and the origins of Yakuts. Hum Genet 120(3):334-53.

Pakendorf B, Novgorodov IN, Osakovskij VL, Stoneking M (2007) Mating patterns amongst Siberian reindeer herders: inferences from mtDNA and Y-chromosomal analyses. Am J Phys Anthropol 133(3):1013-27.

Pakendorf B, Wiebe V, Tarskaia LA, Spitsyn VA, Soodyall H, et al. (2003) Mitochondrial DNA evidence for admixed origins of central Siberian populations. Am J Phys Anthropol 120(3):211-24.

Parson W, Parsons TJ, Scheithauer R, Holland MM (1998) Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: application of mtDNA sequence analysis to a forensic case. Int J Legal Med 111(3):124-32.

Passarino G, Cavalleri GL, Lin AA, Cavalli-Sforza LL, Børresen-Dale AL, et al. (2002) Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms. Eur J Hum Genet 10(9):521-9.

Pereira L, Cerný V, Cerezo M, Silva NM, Hájek M, et al. (2010) Linking the sub-Saharan and West Eurasian gene pools: maternal and paternal heritage of the Tuareg nomads from the African Sahel. Eur J Hum Genet 18(8):915-23.

Pereira L, Cunha C, Amorim A (2004) Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. Int J Legal Med 118(3):132-6.

Pereira V, Gomes V, Amorim A, Gusmão L, João Prata M (2010) Genetic characterization of uniparental lineages in populations from Southwest Iberia with past malaria endemicity. Am J Hum Biol 22(5):588-95.

Pericić M, Lauc LB, Klarić IM, Rootsi S, Janićijevic B, et al. (2005) High-Resolution Phylogenetic Analysis of Southeastern Europe (SEE) Traces Major Episodes of Paternal Gene Flow Among Slavic Populations. Mol Biol Evol Oct;22(10):1964-75.

Pfeiffer H, Brinkmann B, Hühne J, Rolf B, Morris AA, et al. (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Legal Med 112(5):291-8.

Pfeiffer H, Forster P, Ortmann C, Brinkmann B (2001) The results of an mtDNA study of 1,200 inhabitants of a German village in comparison to other Caucasian databases and its relevance for forensic casework. Int J Legal Med 114(3):169-72.

Pfeiffer H, Steighner R, Fisher R, Mörnstad H, Yoon CL, Holland MM (1998) Mitochondrial DNA extraction and typing from isolated dentin-experimental evaluation in a Korean population

Phillips-Krawczak C, Devor E, Zlojutro M, Moffat-Wilson K, Crawford MH (2006) MtDNA variation in the Altai-Kizhi population of southern Siberia: a synthesis of genetic variation. Hum Biol 78(4):477-94.

Pichler I, Mueller JC, Stefanov SA, De Grandi A, Volpato CB, et al. (2006) Genetic structure in contemporary south Tyrolean isolated populations revealed by analysis of Y-chromosome, mtDNA, and Alu polymorphisms. Hum Biol 78(4):441-64.

Picornell A, Gómez-Barbeito L, Tomàs C, Castro JA, Ramon MM (2005) Mitochondrial DNA HVRI variation in Balearic populations. Am J Phys Anthropol 128(1):119-30.

Piercy R, Sullivan KM, Benson N, Gill P (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Int J Legal Med 106(2):85-90.

Pimenoff VN, Comas D, Palo JU, Vershubsky G, Kozlov A, et al. (2008) Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers Eur J Hum Genet Oct;16(10):1254-64.

Pinto F, González AM, Hernández M, Larruga JM, Cabrera VM (1996) Genetic relationship between the Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. Ann Hum Genet 60(Pt 4):321-30.

Plaza S, Calafell F, Helal A, Bouzerna N, Lefranc G, et al. (2003) Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. Ann Hum Genet 67(Pt 4):312-28.

Pliss L, Tambets K, Loogväli EL, Pronina N, Lazdins M, et al. (2006) Mitochondrial DNA portrait of Latvians: toward the understanding of the genetic structure of Baltic-Speaking populations. . Ann Hum Genet 70(Pt 4):439-58.

Poetsch M, Wittig H, Krause D, Lignitz E (2004) Corrigendum to ‘‘Mitochondrial diversity of a northeast German population sample’’. Forensic Sci Int 145(1):73-7.

Prieto L, Zimmermann B, Goios A, Rodriguez-Monge A, Paneto GG, et al. (2011) The GHEP-EMPOP collaboration on mtDNA population data--A new resource for forensic casework. Forensic Sci Int Genet 5(2):146-51.

Pshenichnov A, Balanovsky O, Utevska O, Metspalu E, Zaporozhchenko V (2013), Genetic affinities of Ukrainians from the maternal perspective. Am. J. Phys. Anthropol., 152: 543–550. doi: 10.1002/ajpa.22371

Pult I, Sajantila A, Simanainen J, Georgiev O, Schaffner W, et al. (1994) Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. Biol Chem Hoppe Seyler 375(12):837-40.

Puzyrev VP, Stepanov VA, Golubenko MV, Puzyrev KV, Maksimova NR, et al. (2003) MtDNA and Y-chromosome lineages in the Yakut population. Genetika 39(7):975-81.

Qian YP, Chu ZT, Dai Q, Wei CD, Chu JY, et al. (2001) Mitochondrial DNA polymorphisms in Yunnan nationalities in China. J Hum Genet 46(4):211-220.

Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, et al. (2004) Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. Am J Hum Genet 74(5):827-45.

Rando JC, Pinto F, González AM, Hernández M, Larruga JM, et al. (1998) Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European, near-eastern, and sub-Saharan populations. Ann Hum Genet 62(Pt 6):531-50.

Regueiro M, Cadenas AM, Gayden T, Underhill PA, Herrera RJ (2006) Iran: tricontinental nexus for Y-chromosome driven migration. Hum Hered 61: 132-143.

Regueiro M, Rivera L, Damnjanovic T, Lukovic L, Milasin J,et al. (2012) High levels of Paleolithic Y-chromosome lineages characterize Serbia Gene Apr 25;498(1):59-67.

Reidla M, Mitochondrial DNA lineages in Georgia.

Rhouda T, Martínez-Redondo D, Gómez-Durán A, Elmtili N, Idaomar M, et al. (2009) Moroccan mitochondrial genetic background suggests prehistoric human migrations across the Gibraltar Strait. Mitochondrion 9(6):402-7.

Richard C, Pennarun E, Kivisild T, Tambets K, Tolk HV, et al. (2007) An mtDNA perspective of French genetic variation. Ann Hum Biol 34(1):68-79.

Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, et al. (1996) Paleolithic and neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59(1):185-203.

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, et al. (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 67(5):1251-76.

Robino C, Crobu F, Di Gaetano C, Bekada A, Benhamamouch S, et al. (2008) Analysis of Y-chromosomal SNP haplogroups and STR haplotypes in an Algerian population sample. Int J Legal Med 122:251–255.

Roewer L, Willuweit S, Krüger C, Nagy M, Rychkov S, et al. (2008) Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. Int J Legal Med May;122(3):219-23.

Rousselet F, Mangin P (1998) Mitochondrial DNA polymorphisms: a study of 50 French Caucasian individuals and application to forensic casework. Int J Legal Med 111(6):292-8.

Rowold DJ, Luis JR, Terreros MC, Herrera RJ (2007) Mitochondrial DNA geneflow indicates preferred usage of the Levant Corridor over the Horn of Africa passageway. J Hum Genet 52(5):436-47.

Rubicz R Evolutionary consequences of recently founded Aleut communities in the Commander and Pribilof Islands.

Rubicz R, Melton PE, Spitsyn V, Sun G, Deka R, et al. (2010) Genetic structure of native circumpolar populations based on autosomal, mitochondrial, and Y chromosome DNA markers. Am J Phys Anthropol 143(1):62-74.

Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, et al. (1995) Genes and languages in Europe: an analysis of mitochondrial lineages. Genome Res 5(1):42-52.

Sajantila A, Salem AH, Savolainen P, Bauer K, Gierig C, et al. (1996) Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc Natl Acad Sci U S A 93(21):12035-9.

Salas A, Comas D, Lareu MV, Bertranpetit J, Carracedo A (1998) mtDNA analysis of the Galician population: a genetic edge of European variation. Eur J Hum Genet 6(4):365-75.

Sampietro ML, Lao O, Caramelli D, Lari M, Pou R, Martí M, Bertranpetit J, Lalueza-Fox C. Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. Proc Biol Sci. 2007; 274(1622):2161-7.

Sanchez JJ, Hallenberg C, Borsting C, Hernandez A, Morling N (2005) High frequencies of Y chromosome lineages characterized by E3b1, DYS19-11, DYS392-12 in Somali males. Eur J Hum Genet 13:856–866.

Sánchez-Quinto F, Schroeder H, Ramirez O, Avila-Arcos MC, Pybus M, Olalde I, Velazquez AM, Marcos ME, Encinas JM, Bertranpetit J, Orlando L, Gilbert MT, Lalueza-Fox C. Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. Curr Biol. http://dx.doi.org/10.1016/j.cub.2012.06.005

Saunier JL, Irwin JA, Strouss KM, Ragab H, Sturk KA, et al. (2009) Mitochondrial control region sequences from an Egyptian population sample. Forensic Sci Int Genet 3(3):e97-103.

Scheible M, Alenizi M, Sturk-Andreaggi K, Coble MD, Ismael S, et al. (2011) Mitochondrial DNA control region variation in a Kuwaiti population sample. Forensic Sci Int Genet 5(4):e112-3.

Schönberg A, Theunert C, Li M, Stoneking M, Nasidze I (2011) High-throughput sequencing of complete human mtDNA genomes from the Caucasus and West Asia: high diversity and demographic inferences. Eur J Hum Genet 19(9):988-94.

Schurr TG, Sukernik RI, Starikovskaya YB, Wallace DC (1999) Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. Am J Phys Anthropol 108(1):1-39.

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, et al. (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science Nov 10;290(5494):1155-9.

Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, et al. (2006) Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. Am J Hum Genet 78: 202-221.

Shen P, Lavi T, Kivisild T, Chou V, Sengun D, et al. (2004) Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. Hum Mutat 24: 248-260.

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, et al. (1993) mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am J Hum Genet 53(3):549-62.

Shimada MK, Kim CG, Takahashi A, Spitsyn VA, Ikeo K, et al. (2002) Mitochondrial DNA control region sequences for a Buryats population in Russia. In: DNA Polymorphism Edited by: Ishidu H. Tokyo: TOYO Shoten CO., LTD; 10: 151-155.

Shlush LI, Behar DM, Yudkovsky G, Templeton A, Hadid Y, et al. (2008) The Druze: a population genetic refugium of the Near East. PLoS One. 3(5):e2105.

Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E, Gilbert MTP, Götherström A, Jakobsson M. Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe. Science. 2012; 336(6080):466-9

Starikovskaya EB, Sukernik RI, Derbeneva OA, Volodko NV, Ruiz-Pesini E, et al. (2005) Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. Ann Hum Genet 69(Pt 1):67-89.

Stenico M, Nigro L, Bertorelle G, Calafell F, Capitanio M, et al. (1996) High mitochondrial sequence diversity in linguistic isolates of the Alps. Am J Hum Genet 59(6):1363-75.

Stevanovitch A, Gilles A, Bouzaid E, Kefi R, Paris F, et al. (2004) Mitochondrial DNA Sequence diversity in a sedentary population from Egypt. Ann Hum Genet 68(Pt 1):23-39.

Sykes B Blood of the Isles. Bantam Press, London.

Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M (1995) The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. Am J Hum Genet 57(6):1463-75.

Szécsényi-Nagy A, Brandt G, Keerl V, Jakucs J, Haak W, et al. (2014) Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. In Press. Available on bioRxiv. http://dx.doi.org/10.1101/008664

Tagliabracci A, Turchi C, Buscemi L, Sassaroli C (2001) Polymorphism of the mitochondrial DNA control region in Italians. Int J Legal Med 114(4-5):224-8.

Tajima A, Hayami M, Tokunaga K, Juji T, Matsuo M, et al. (2004) Genetic origins of the Ainu inferred from combined DNA analyses of maternal and paternal lineages. J Hum Genet 49(4):187-93.

Tajima A, Sun CS, Pan IH, Ishida T, Saitou N, et al. (2003) Mitochondrial DNA polymorphisms in nine aboriginal groups of Taiwan: implications for the population history of aboriginal Taiwanese. Hum Genet 113(1):24-33.

Tambets K, Rootsi S, Kivisild T, Help H, Serk P, et al. (2004) The Western and Eastern Roots of the Saami--the Story of Genetic "Outliers" Told by Mitochondrial DNA and Y Chromosomes. Am J Hum Genet 74: 661-682.

Tanaka M, Cabrera VM, González AM, Larruga JM, Takeyasu T, et al. (2004) Mitochondrial genome variation in eastern Asia and the peopling of Japan. Genome Res 14(10A):1832-50.

Tetzlaff S, Brandstätter A, Wegener R, Parson W, Weirich V (2007) Mitochondrial DNA population data of HVS-I and HVS-II sequences from a northeast German sample. . Forensic Sci Int 172(2-3):218-24.

Thomas MG, Barnes I, Weale ME, Jones AL, Forster P, et al. (2008) New genetic evidence supports isolation and drift in the Ladin communities of the South Tyrolean Alps but not an ancient origin in the Middle East. Eur J Hum Genet 16(1):124-34.

Tillmar AO, Coble MD, Wallerström T, Holmlund G (2010) Homogeneity in mitochondrial DNA control region sequences in Swedish subpopulations. Int J Legal Med 124(2):91-8.

Tömöry G, Csányi B, Bogácsi-Szabó E, Kalmár T, Czibula A, et al. (2007) Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations. Am J Phys Anthropol 134(3):354-68.

Tonks S, Winney BJ Comparison of sex-linked and autosomal markers in Orkney and other North European populations.

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J,et al. (1994) Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93(2):189-99.

Turchi C, Buscemi L, Giacchino E, Onofri V, Fendt L, et al. (2009) Polymorphisms of mtDNA control region in Tunisian and Moroccan populations: an enrichment of forensic mtDNA databases with Northern Africa data. Forensic Sci Int Genet 3(3):166-72.

Turchi C, Buscemi L, Previderè C, Grignani P, Brandstätter A, et al. (2008) Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing. Int J Legal Med 122(3):199-204.

Vanecek T, Vorel F, Sip M (2004) Mitochondrial DNA D-loop hypervariable regions: Czech population data. Int J Legal Med 118(1):14-8.

Verginelli F, Donati F, Coia V, Boschi I, Palmirotta R, et al. (2003) Variation of the hypervariable region-1 of mitochondrial DNA in central-eastern Italy. J Forensic Sci 48(2):443-4.

Vernesi C, Di Benedetto G, Caramelli D, Secchieri E, Simoni L, et al. (2001) Genetic characterization of the body attributed to the evangelist Luke. Proc Natl Acad Sci U S A 98(23):13460-3.

Vernesi C, Fuselli S, Castrì L, Bertorelle G, Barbujani G (2002) Mitochondrial diversity in linguistic isolates of the Alps: a reappraisal. Hum Biol 74(5):725-30.

Voevoda MI, Shkapenko AL Mitochondrial sequences in the hypervariable region I of West Siberian Finno-Ugrians.

Völgyi A, Zalán A, Szvetnik E, Pamjav H (2009) Hungarian population data for 11 Y-STR and 49 Y-SNP markers. Forensic Sci Int Genet 3:e27–e28.

Volodko NV, Starikovskaya EB, Mazunin IO, Eltsov NP, Naidenko PV, et al. (2008) Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. Am J Hum Genet 82(5):1084-100.

Vona G, Ghiani ME, Calò CM, Vacca L, Memmì M, et al. (2001) Mitochondrial DNA sequence analysis in Sicily. Am J Hum Biol 13(5):576-89.

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, et al. (2001) The Eurasian Heartland: A continental perspective on Y-chromosome diversity Proc Natl Acad Sci U S A Aug 28;98(18):10244-9.

Wen B, Xie X, Gao S, Li H, Shi H,et al. (2004) Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans. Am J Hum Genet 74(5):856-65.

Wilde S, Timpson A, Kirsanow K, Kaiser E, Kayser M, Unterländer M, Hollfelder N, Potekhina ID, Schier W, Thomas MG, Burger J. (2014) Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. Proc Natl Acad Sci U S A. 2014 Apr 1;111(13):4832-7. doi: 10.1073/pnas.1316513111.

Wong HY, Tang JS, Budowle B, Allard MW, Syn CK, et al. (2007) Sequence polymorphism of the mitochondrial DNA hypervariable regions I and II in 205 Singapore Malays. Leg Med (Tokyo) 9(1):33-7.

Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, et al. (2006) Male demography in East Asia: a north-south contrast in human population expansion times. Genetics 172: 2431–2439.

Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP (2002) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70(3):635-51.

Yao YG, Kong QP, Man XY, Bandelt HJ, Zhang YP (2003) Reconstructing the evolutionary history of China: a caveat about inferences drawn from ancient DNA. Mol Biol Evol 20(2):214-9.

Yao YG, Kong QP, Wang CY, Zhu CL, Zhang YP (2004) Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in china. Mol Biol Evol 21(12):2265-80.

Yao YG, Nie L, Harpending H, Fu YX, Yuan ZG, et al. (2002) Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. Am J Phys Anthropol 118(1):63-76.

Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, et al. (2012) The Caucasus as an asymmetric semipermeable barrier to ancient human migrations Mol Biol Evol Jan;29(1):359-65.

Zalloua PA, Platt DE, El Sibai M, Khalife J, Makhoul N, et al. (2008) Genographic Consortium Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. Am J Hum Genet 83:633–642.

Zalloua PA, Xue Y, Khalife J, Makhoul N, Debiane L (2008) Y-chromosomal diversity in Lebanon is structured by recent historical events. Am J Hum Genet 82:873–882.

Zerjal T, Wells RS, Yuldasheva N, Ruzibakiev R, Tyler-Smith C (2002) A genetic landscape reshaped by recent events: Y-chromosomal insights into Central Asia. Am J Hum Genet 71: 466–482.

Zgonjanin D, Veselinović I, Kubat M, Furac I, Antov M, et al. (2010) Sequence polymorphism of the mitochondrial DNA control region in the population of Vojvodina Province, Serbia. Leg Med (Tokyo) 12(2):104-7.

Zimmermann B, Bodner M, Amory S, Fendt L, Röck A, et al. (2009)Forensic and phylogeographic characterization of mtDNA lineages from northern Thailand (Chiang Mai). Int J Legal Med 123(6):495-501.

Zimmermann B, Brandstätter A, Duftner N, Niederwieser D, Spiroski M, et al. (2007) Mitochondrial DNA control region population data from Macedonia. Forensic Sci Int Genet 1(3-4):e4-9.

Zupanic Pajnic I, Balazic J, Komel R (2004) Sequence polymorphism of the mitochondrial DNA control region in the Slovenian population. Int J Legal Med 118(1):1-4.

15.5 Curriculum Vitae

Anna Szécsényi-Nagy, M.Sc.

Professional experiences

01.2014- Research fellow at the Institute of Archaeology Research Centre for the Humanities, Hungarian Academy of Sciences.

01. 2010- 01.2013 Research associate at the Institute of Anthropology, Johannes Gutenberg University of Mainz.

10.2008-07.2009 10-month DAAD research grant at the Johannes Gutenberg University of Mainz. Archaeogenetic research in the team of Prof Dr Kurt W. Alt, on the project entitled "Balatonszárszó".

09.2006-03.2007 6-month Erasmus scholarship at the Humboldt University of Berlin and 5- month biochemical research traineeship in the Max Delbrück Centrum (MDC) Berlin-Buch, in a project of Prof Dr Udo Heinemann's research team.

2005-2007 Scientific Students' Association work at the Semmelweis Medical University in Budapest, Department of Medical Chemistry, Molecular Biology and Pathobiochemistry, in the research team of Dr Mária Sasvári DSc.

Education

02. 2010- PhD student at Johannes Gutenberg University of Mainz. Supervisor is Prof Dr Kurt W. Alt.

2008-2009 Master thesis at Eötvös Loránd University in Budapest, Institute of Archaeology. Title: "Custom of the cranial deformation in the V-VI. century Carpathian Basin"

2006-2007 Master thesis at the Eötvös Loránd University in Budapest, Faculty of Science, Department of Genetics. Title: „Copy number polymorphisms of the human genome, at the example of C4A/B gene"

2003-2009 Studies of Archaeology at the Eötvös Loránd University in Budapest, Faculty of Humanities. Major subject: archaeology, specialty branch: prehistory and migration period

2002-2007 Studies of Biology at the Eötvös Loránd University in Budapest, Faculty of Science. Major subject: biology, specialty branch: molecular biology, genetics.

15.6 List of own publications

Papers

Szécsényi-Nagy, A, Brandt, G, Keerl, V, Jakucs, J, Haak, W, Moeller-Rieker, S, Köhler, K, Mende, BG, Fecher, M, Oross, K, Marton, T, Osztás, A, Kiss, V, Pálfi, Gy, Molnár, E, Sebők, K, Czene, A, Paluch, T, Šlaus, M, Novak, M, Pećina-Šlaus, N, Ősz, B, Voicsek, V, Somogyi K, Tóth, G, Kromer, B, Bánffy, E, Alt, KW. (2014) Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. In Press. Available on bioRxiv. http://dx.doi.org/10.1101/008664

Szécsényi-Nagy A, Keerl V, Jakucs J, Brandt G, Bánffy E, Alt KW. (2014) Ancient DNA evidence for a homogeneous maternal gene pool in sixth millennium cal BC Hungary and the central European LBK. In: Whittle, A, Bickle P (eds): Early Farmers: The View from Archaeology and Science. Proceedings of the British Academy 198. Oxford: Oxford University Press for the British Academy. 71-93.

Brandt G, *Szécsényi-Nagy A*, Roth C, Alt KW, Haak W: Palaeogenetics of Europe - the known knowns and the known unknowns. Journal of Human Evolution. http://dx.doi.org/10.1016/j.jhevol.2014.06.017

Brandt G, Haak W, Adler CJ, Roth C, *Szécsényi-Nagy A,* Karimnia S, Möller-Rieker S, Meller H, Ganslmeier R, Friederich S, Dresely V, Nicklisch N, K. Pickrell J, Sirocko F, Reich D, Cooper A, Alt KW, and The Genographic Consortium. (2013) Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity Science 11 October 2013: 342 (6155), 257-261. [DOI:10.1126/science.1241844]

Posters and presentations

Szécsényi-Nagy A, Mende BG: Kultúrák, emberek, gének. Archaeogenetikai vizsgálatok a Dunántúl Kr. e. 6–5. évezredének népességein. (presentation) A Filozófiai és Történettudományok Osztályának tudományos ülése a Magyar Tudomány Ünnepéhez kapcsolódóan, Vándorlás, változás, kontinuitás – évezredek tanulságai a jövőnek. 2014. november 18.

Szécsényi-Nagy A, Brandt G, Jakucs J, Mende BG, Bánffy E, Alt KW: Ancient mitochondrial and Y chromosomal DNA reveals the western Carpathian Basin as a corridor of the Neolithic expansion. (presentation) ISBA6, Basel, Switzerland 27-29th August 2014.

Szécsényi-Nagy, A, Fecher, M Diachronic observations: the signs of kinship and mobility in three Neolithic grave groups at Alsónyék-Bátaszék site. (presentation) Workshop Bioarchaeology and archaeology in Neolithic Europe: the Carpathian Basin, Mainz – RGZM, Germany. 3-5th April 2014.

Szécsényi-Nagy, A: Population genetic evidence of the cultural changes in the Neolithic Transdanubia. (presentation) Workshop Bioarchaeology and archaeology in Neolithic Europe: the Carpathian Basin, Mainz – RGZM, Germany. 3-5th April 2014

Szécsényi-Nagy, A., Jakucs, J., Bánffy, E., W. Alt, K. 6–5th millennium BC cultural changes in Western Hungary tested by ancient DNA. (presentation) 19th Annual meeting of the European Association of Archaeologists. Pilsen, Czech Republic 4-8 September 2013.

Szécsényi-Nagy, A., Fecher, M., Nyerges É.A., Bánffy, E., W. Alt, K. Ancient DNA and isotope analysis of the Starcevo graves at Alsónyék-Bátaszék. (presentation) 19th Annual meeting of the European Association of Archaeologists. Pilsen, Czech Republic 4-8 September 2013.

Szécsényi-Nagy, A. & Keerl V, Jakucs J, Bánffy E, Alt KW "Ancient DNA from middle Neolithic Hungary suggests a common ancestry of the Linear Pottery Culture in Central Europe" (poster) Early Farmers: the view from Archaeology and Science. Cardiff 14-16. Mai, 2012.

Szécsényi-Nagy, A. "Copy number polymorphism of the C4A/B genes" (presentation) BioCoin 2007 (4th annual International Natural Sciences Students Conference). Vilnius, Lithuania 20-24th March 2007.

Szécsényi-Nagy, A. "Gene number polymorphism of the C4A/B complement protein in autism" (presented in Hungarian) Scientific medical student competition of the Semmelweis University. Budapest, Hungary 14-15th February 2007.

15.7 *Erklärung* / Declaration of originality

Hiermit versichere ich, Anna Szécsényi-Nagy, dass ich die vorliegende Arbeit eigenständig angefertigt und keine anderen als die angegebenen Hilfsmittel und Quellen in dieser Arbeit verwendet habe. Die wörtlichen oder dem Inhalt nach aus fremden Arbeiten entnommenen Stellen, Zeichnungen, bildliche Darstellungen und dergleichen als solche genau kenntlich gemacht sind. Die Arbeit wurde in gleicher oder abgewandelter Form noch keiner anderen Stelle als Prüfungsleistung vorgelegt.

I, Anna Szécsényi-Nagy hereby declare that this thesis and the work reported herein was composed by and originated entirely from me, and that the work has not been previously submitted for a degree or diploma at any institution. Information derived from the published and unpublished work (figures, datasets etc.) of others are noted in the text and references are given in the list of sources.

_____ _____

Place, Date Signature