Ancient DNA reveals diverse community organizations in the 5th millennium BCE Carpathian Basin

Anna Szécsényi-Nagy[1], Cristian Virag[2], Kristóf Jakab[1], Nadin Rohland[3,4], Harald Ringbauer[5], Alexandra Anders[6], Pál Raczky[6], Tamás Hajdu[7], Krisztián Kiss[1,7], Tamás Szeniczey[8], Sándor Évinger[9], Tamás Keszi[10], Zsuzsanna M. Virág[11], Olivia Cheronet[12], Swapan Mallick[3,4,13], Ali Akbari[3,4,13], Ron Pinhasi[12], David Reich[3,4,13,14], Zsuzsanna Siklósi[6,15]

1. Institute of Archaeogenomics, HUN-REN Research Centre for the Humanities, 1097 Budapest, Hungary
2. Satu Mare County Museum, 440031 Satu Mare, Romania
3. Department of Genetics, Harvard Medical School, Boston, MA 02138, USA
4. Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
5. Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany
6. Institute of Archaeological Sciences, ELTE Eötvös Loránd University, 1088 Budapest, Hungary
7. Department of Biological Anthropology, Institute of Biology, Faculty of Science, ELTE Eötvös Loránd University, 1117 Budapest, Hungary
8. Faculty of Health Sciences, University of Miskolc, 3515 Miskolc, Hungary
9. Department of Anthropology, Hungarian Natural History Museum, Hungarian National Museum Public Collection Centre, 1083 Budapest, Hungary
10. Intercisa Museum, Dunaújváros, Hungary
11. Budapest History Museum, Aquincum Museum, 1031 Budapest, Hungary
12. Department of Evolutionary Anthropology, University of Vienna, 1090 Vienna, Austria
13. Howard Hughes Medical Institute, Boston, MA 02138, USA
14. Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
15. MTA-ELTE Lendület "Momentum" Innovation Research Group, 1088 Budapest, Hungary

Corresponding authors

Anna Szécsényi-Nagy (szecsenyi-nagy.anna@abtk.hun-ren.hu), David Reich (reich@genetics.med.harvard.edu) and Zsuzsanna Siklósi (siklosi.zsuzsanna@btk.elte.hu)

Keywords:
Archaeogenomics, Late Neolithic, Copper Age, Carpathian Basin, identity by descent, kinship, inbreeding, 5th Millennium BCE

Abstract

We present a comprehensive genetic investigation of Late Neolithic (LN) and Early Copper Age (ECA) populations living in the Carpathian Basin, leveraging whole genome data from 125 previously unreported individuals. Using population genetics, kinship analyses and the study of networks of identity-by-descent haplotype segment sharing, we elucidate the social and genetic dynamics of these communities between 4800-3900 cal BCE. Despite changes in settlement patterns, burial practices, and material culture, we document a high degree of genetic continuity. While one set of individuals we analyzed from a large community cemetery was genetically diverse, another site was more homogenous and closed, with numerous consanguineous relationships and evidence of patrilineality and patrilocality. These results document radically different kinship systems in contemporaneous ECA communities using similar material culture and living only about 100 km apart.

List of Abbreviations

ANF: Anatolian Neolithic farmers
CA: Copper Age
CB: Carpathian Basin
ECA: Early Copper Age
EHG: Eastern European hunter-gatherer
GHP: Great Hungarian Plain
HG: Hunter-gatherer
IBD: Identity by descent
LD: Linkage Disequilibrium
LN: Late Neolithic
MCA: Middle Copper Age
MN: Middle Neolithic
mtDNA: mitochondrial DNA
N: Neolithic
Ne: Effective population size
PCA: Principal Component Analysis
ROH: Runs of homozygosity
SNP: Single nucleotide polymorphism
WHG: Western hunter-gatherer

Introduction

The geographical position of the Carpathian Basin (CB) makes it a crossroads between South-East and Central Europe. The fertile river valleys of the Danube and its tributaries offered optimal conditions for settlement and facilitated the development of efficient transport and communication networks for early farming communities. Since the advent of the field of aDNA research, the prehistoric populations of the CB have been the subject of intensive study, primarily due to the critical role they played in the Neolithisation of Europe[1,2]. The population genetic history of the Neolithic (N) (6000-4500 cal BCE) of the CB is well-established[3–6]. However, the subsequent Copper Age (CA) (4500-2800 cal BCE) is underrepresented in the available genetic data[4,7–10].

The Copper Age is characterized by the spread of significant technological innovations, including metallurgy[11,12] and the wheel and wagon[13,14], which profoundly impacted human culture. This period is regarded to coincide with the emergence of salient social ranking and craft specialization[15,16].

During the LN (4900-4500 cal BCE), tells and large horizontal settlements were established in the CB[17], with some covering an area of up to 60-80 ha. These sites were distinguished by a concentration of populations[18–20]. The number of inhabitants of the LN settlements has been estimated at several thousand based on the number of excavated buildings and burials[19,21]. The use of these extensive settlements terminated by the end of the LN, between 4500 and 4450 cal BCE[22]. Profound transformations were observed on the GHP, affecting all segments of life. In contrast to the large horizontal LN settlements, a dense network of small, farm-like settlements emerged across the GHP[23,24]. During the LN, the deceased were buried within the settlement boundaries or through rites that left no archaeologically visible traces[25]. However, in the ECA, from 4400-4350 cal BCE, formal cemeteries – the first in the CB – appeared that were spatially separated from the settlements[22]. Meanwhile, the practice of settlement burials persisted[26,27], with even LN tells being repurposed for burial grounds[28,29]. These changes coincided with shifts in the material culture, particularly in the pottery style. Subsequently, on the GHP, the Tiszapolgár[30] and then the Bodrogkeresztúr style emerged[31,32]. In parallel, in Transdanubia (Western Hungary, west of the Danube River), the pottery style of the Lengyel complex underwent transformation[33,34]. The analysis of grave goods reveals disparities in wealth, while the use of prestige and status goods provides evidence of social differentiation[25,35]. It has been proposed that the ECA formal cemeteries functioned as a kind of common burial place, a central place for the inhabitants of several nearby smaller settlements[22,36].

The interpretations of these transformations are diverse, encompassing a range of hypotheses. These include migration[37], economic changes[38], climate change[39], and internal social reorganizations[40]. However, some elements of cultural continuity between the LN and ECA of the GHP rather support an internal, peaceful transformation[27]. This interpretation is further supported by anthropological and strontium isotope analyses, which indicate no significant increase in individual mobility or evidence of a new immigrant population between the LN and the ECA[41,42].

We report genome-wide data from 125 LN and ECA burials to address two interrelated questions (Table S1). First, we test if a population shift occurred between the LN and ECA, and if so, such a shift correlates with the observed cultural and economic transformations. Secondly, we aim to understand the organization of the LN and ECA

communities, specifically whether we can discern evidence of internal reorganization related to the changes between the two periods. To investigate these questions, we conducted a comparative analysis between two significant LN archaeological sites, Aszód-Papi földek and Polgár-Csőszhalom, and two large, almost completely excavated ECA cemeteries, Tiszapolgár-Basatanya (referred to as Basatanya in this paper) and Urziceni-Vamă, which exhibit a high degree of cultural similarities to each other. The Polgár microregion, situated on the Upper Tisza River, provides an ideal setting for investigating the question of local continuity. Extensive excavations in this 91 km² area have uncovered several archaeological sites spanning from the Middle Neolithic (MN, Polgár-Ferenci-hát ((5470-5070 cal BCE)) through the LN (Polgár-Csőszhalom) to the ECA (Polgár-Nagy-Kasziba and Basatanya)[43]. To provide a broader regional perspective, we incorporated human remains from additional contemporaneous CA sites, including those from Budapest-Albertfalva-Hunyadi János út, Ivánccsa-Lapos, and Rákóczifalva-Bagi föld site 8 (SI Chapter 1, Table S2).

Studies on genetic relatedness are almost completely absent from prehistoric periods in Hungary[6,10], although previous analyses of other areas of Europe documented a variety of kinship and residence practices[44–48]. To obtain insight into internal transformations affecting the organization of communities that underlie archaeologically visible cultural changes, we combined population genetic methods with biological relatedness analyses[49,50] and identity by descent (IBD) analysis[51].

Results

We generated genome-wide data by capturing more than 1.2 million single nucleotide polymorphisms (SNPs) in 125 individuals who lived during the LN and ECA of the CB, and report enhanced capture or shotgun genomes of a further 6+18 individuals for which aDNA data were published previously[4,52] (**Fig. 1**). We combined the newly reported data with previously published data from the CB[4,6,52–54] and other genomic data from relevant ancient Europeans (see Methods).

As different regions were reached by new technological innovations at different times, there is a variation in the chronological nomenclature across the regions surrounding the CB. We use the terminology specific to the given region and provide corresponding radiocarbon dates of the major reference groups in **Figs. 1-2**. Archaeological cultures used for classification should not be interpreted as coherent social groups that necessarily existed in the past[55–57]. Behind the diversity of material culture lie a range of social identities, and the extent to which material cultures correspond to coherent social groups must be tested and evaluated on a case-by-case basis and not assumed from the outset[58,59]. To avoid the pitfall of conflating archaeological cultures with specific populations, we use individuals and communities as the basic units of analysis, and test whether there is any correlation between the use of various pottery styles and biological relatedness.

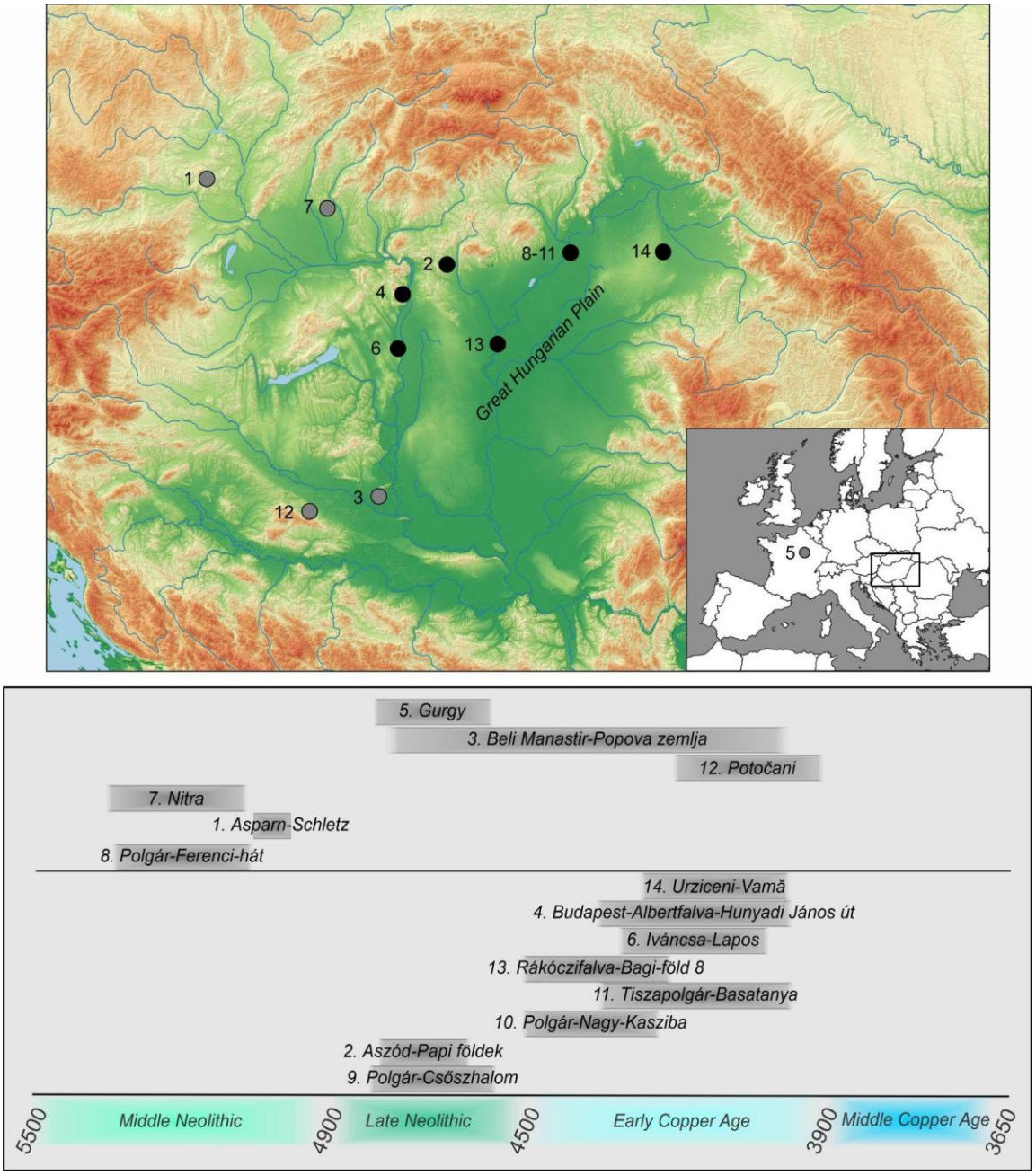

Figure 1: Map of the Carpathian Basin with the sites mentioned in the text and their chronologies. Date range is given in gray bands at the bottom in cal BCE. Black markers indicate sites for which we generated new aDNA data, gray markers indicate separately published comparative data. Numbers 8-11 cover the sites from the Polgár microregion. Basemap was made with Natural Earth (free vector and raster map data @ naturalearthdata.com) and SRTM provided by NASA (10.5067/MEASURES/SRTM/SRTMGL1.003).

Genomic composition of the studied communities

We carried out a principal component analysis (PCA) on genome-wide data using *smartpca*[60], projecting ancient individuals onto modern Western Eurasian genetic variation to visualize key trends in both the study period and area (**Fig. 2A**). The new data, combined with previously published N genomes[4,61,62], indicate that LN populations

of the CB exhibit moderate hunter-gatherer (HG) ancestry, with a position between Neolithic Anatolia, Western Europe, and northeastern (Bohemian) LN groups. Within the CB, GHP communities show higher HG ancestry compared to their Transdanubian counterparts[4]. In the GHP, the LN period showcases a more homogeneous ancestry pattern than the MN (see **Table S3** for Levene's and Bartlett's tests for homogeneity). The ECA populations on the GHP display greater homogeneity and tighter clustering ($p<0.05$, Bartlett's test, **Table S3**), where only two PCA outliers were identified. In contrast, contemporaneous ECA individuals from Urziceni-Vamă (Romania) and others from Romania and Bulgaria[63] exhibit signs of increased ancestry diversity, reflecting new Eastern European genetic contacts. Using Mahalanobis distance and testing based on a chi-square distribution, we detected in Urziceni-Vamă five female and three male outliers on different PC axes ($p<0.05$, **Table S3**), with some individuals shifting toward the Eastern European steppe, and others toward Northwestern or Anatolian populations.

We used the qpAdm software to model with Anatolian N farmer (ANF), and Western and Eastern European HG (WHG and EHG) sources, the potential admixture proportions in the newly generated LN-ECA genomes (see Methods, **Table S5A-F**); the inclusion of multiple hunter-gatherer groups reflects an attempt to cover the diversity of HG groups contributing to people of this region[4,7]. In our analyses, ANF was the dominant component in all N-CA populations of the GHP (88% on average, **Fig. 2B, Fig. S1-S2**). The distribution of EHG and WHG components shows limited EHG in the LN GHP communities on one end, and more in ECA Urziceni-Vamă on the other (a two-sample t-test of their pairwise comparison resulted in $p<0.002$ for the EHG component variance **Fig. 2B, Table S5A-B**).

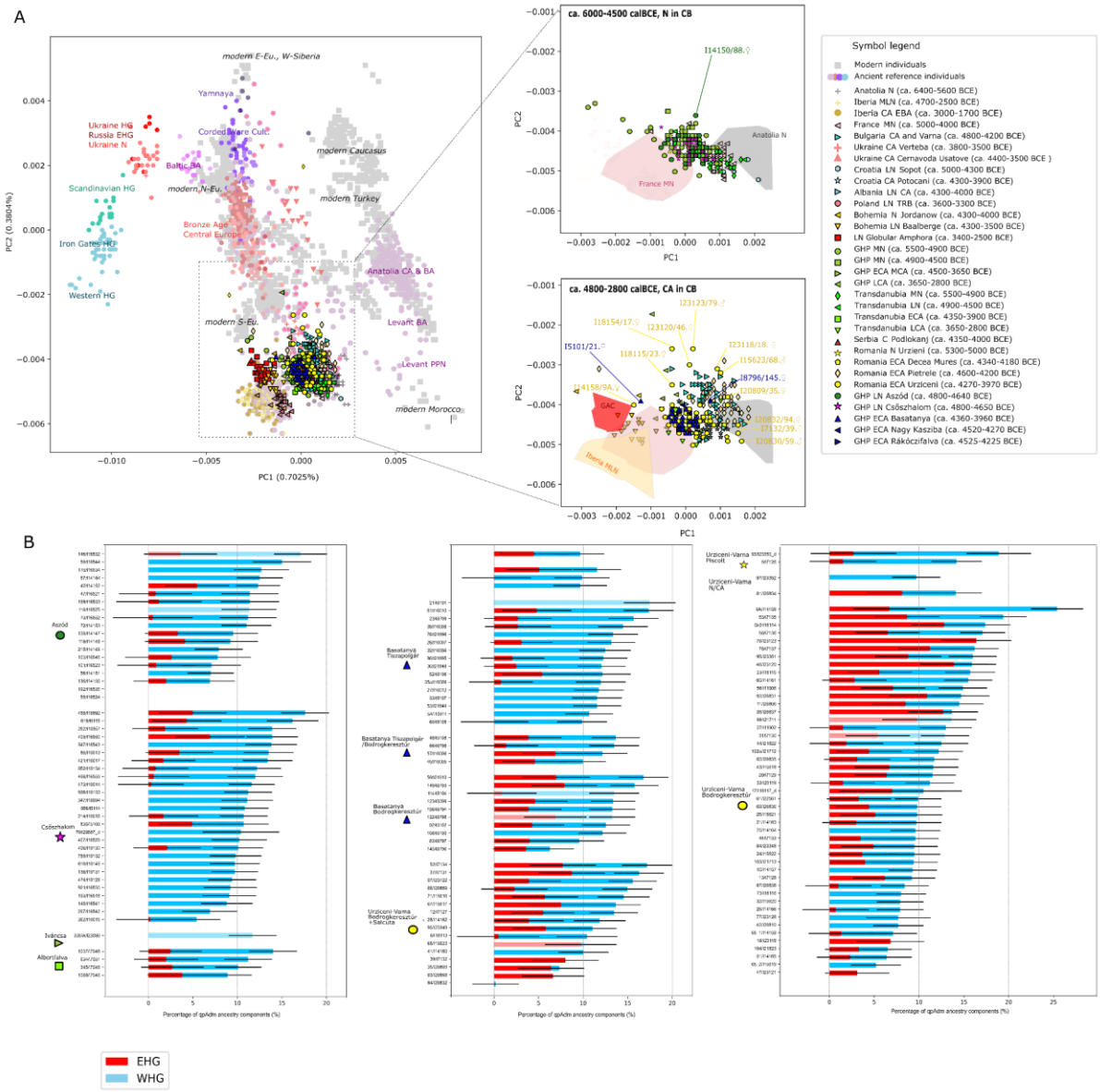

Figure 2: A: Scatterplot showing PC1-2 of a principal component analysis, B: Individual ancestry inferences based on a 3-way qpAdm model. Among the newly generated individuals, outliers (all >50k SNPs detected, tested by Mahalanobis distance calculation combined with chi-square distribution, at the significance level of 0.05, see Methods and **Table S3**) are marked on the zoomed-in parts of panel A. The two zoom-in panels focus on the N and CA samples from CB with comparative data, whose chronology is indicated in the figure legend. WHG: Western Hunter-Gatherer, EHG: Eastern Hunter-Gatherer, the Anatolian N ancestry component is not shown in panel B. Grave numbers and laboratory IDs are indicated on the individual barplots. Faint bars are qpAdm tests with p<0.05 indicating a lack of fit. For further acceptable models see **Table S5**.

Both the LN Aszód-Papi földek and LN Polgár-Csőszhalom populations can be modeled as descended directly without additional admixture from the earlier local MN Polgár-Ferenci-hát population using qpAdm (p=0.26 and 0.40 respectively), consistent with their similar position in PCA. We tested ancestry sources of the subsequent ECA GHP samples on an individual level, and could model 24/30 individuals at Basatanya using the Polgár-Csőszhalom community as a single source (p>0.05, **Table S5E**). At Urziceni-Vamă, 28 out of 50 ECA individuals could be modeled using a local MN source alone; this lower proportion of fitting models suggests additional admixture into this community from exogenous sources (**Table S5E**). Similarly, only half of the ECA individuals fit as

derived without admixture from a population genetically identical or closely related to LN Polgár-Csőszhalom (p>0.05). Consistent with the ancestry variability observed in f_4-statistics (**Fig. S3**), sampled ECA Urziceni-Vamă individuals have evidence of several streams of post-N influx and show diverse qpAdm component proportions, which, however, do not correlate with pottery styles (Bodrogkeresztúr- and Salcuța-styles, see **Tables S4-5**). Sporadic evidence of CHG ancestry is evident graves 17, 18 and 31, which could be early signs of steppe influence into the area, since CHG+EHG appears jointly, like in later Yamnaya-related groups[52] (in nine cases, CHG could be fit without an EHG component (p>0.05, with component Z-score>2 **Table S5D**)). In a group-based qpAdm analysis, the genetic composition of the Urziceni-Vamă group is most similar to Basatanya, whereas the comparative Romanian CA Pietrele and Bulgarian CA Varna site groups show significant differences in their WHG composition[63] (**Table S5F**). This reflects the fact that the Varna group has genetic components from all WHG-EHG and CHG in similar proportion (4-6%), alongside the predominant ANF ancestry. Although the population from which the Varna people came are not a suitable group source for Urziceni-Vamă overall, five Urziceni individuals who do not fit the LN GHP source can be modeled using Varna and CA Pietrele as a single source (p>0.05, **Table S5E**).

The Tiszapolgár-style graves at the Basatanya site show a non-significantly lower EHG component compared to those that have Bodrogkeresztúr-style pottery (ca. 2.6% on average compared to 4.6%, two-sample t-test p=0.0765, see **Tables S5B**). There are also relatively few genetic outliers at Basatanya, with only three individuals failing the ANF-WHG-EHG 3-way qpAdm model test (**Table S5A**).

Parental relatedness and effective population size

To obtain insights into the population composition and effective population size (N_e) of the studied groups, we first inferred and analyzed runs of homozygosity[64] (ROH). We compared the genome fraction in ROH with previously published N-CA datasets[6,44]. An excess of larger ROH segments (>20 cM) at the MN Polgár-Ferenci-hát site provide evidence of sporadic close-kin unions, which are absent in other Early Neolithic groups of Central Europe[6,65] (Fisher's exact test with Polgár-Ferenci-hát vs Combined LBK sites yield a p-value = 0.035, **Fig. 3A, Table S6A**). In the subsequent LN Aszód-Papi földek and Polgár-Csőszhalom communities, contrasting on a microregional level to the MN phenomenon, there is evidence of only one close-kin union on the order of 2nd-cousin level (5th degree) parental relatedness. In the following ECA, two individuals from Urziceni-Vamă and six out of 30 studied individuals at Basatanya had significantly increased levels of long ROH (sum >50 cM in segment >20 cM) indicative for consanguineous relationships between 3rd-degree relatives, such as first cousins (see details on **Fig. S4**), which is notably high in ancient datasets (Basatanya vs. sum of ten reference groups: p-value = 0.0001, **Table S6A)**.

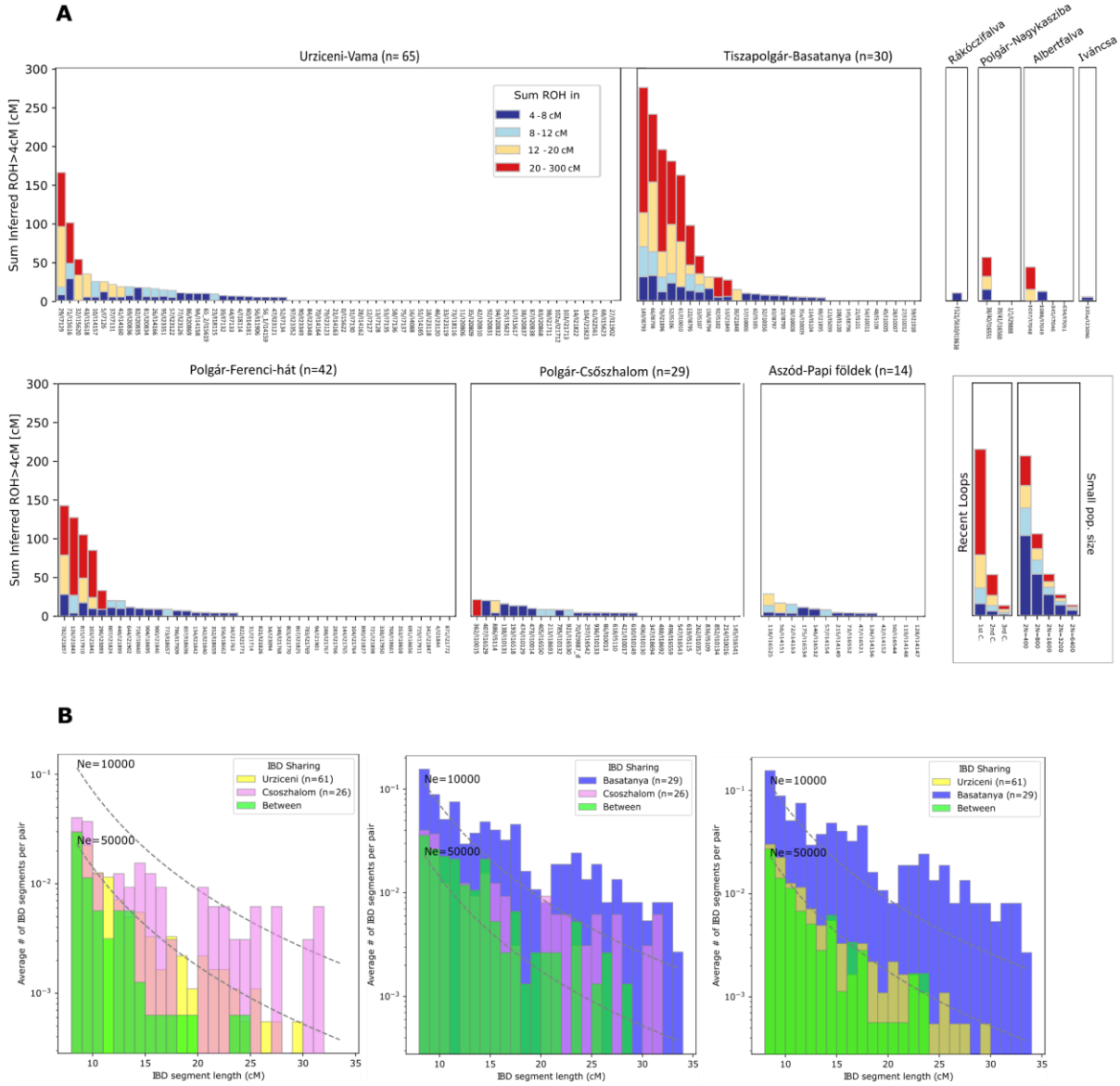

Figure 3: Inferred Runs of Homozygosity (A) and pairwise IBD patterns (B) of the studied sites' populations. A: Plot of ROH for each individual, summing them in the form of stacked, colored bars, depicting different ROH lengths. Individuals from the same sites are grouped into boxes, indicating the number of individuals suitable for ROH analysis (>400k 1240k SNPs covered) in the title of each box. In a separate panel, the legend is depicted with expected ROH for offspring of close kin in an outbred population ("Recent Loops") and panmictic populations of different sizes, using the calibrations from Ringbauer et al.[66]. We observe signals of consanguinity between close kin in two cases at Urziceni, six cases at Basatanya and four cases at MN Polgár-Ferenci-hát[6]. Signs of small effective populations are detectable in shorter ROHs. Further in-depth analyses are seen in **Figure S4** and **Table S6**. B: Inter and intra site IBD sharing patterns and IBD-based effective population size estimates. The plots demonstrate the larger effective population sizes of the Urziceni and Csőszhalom communities compared to Basatanya, and show that inter-site IBD connections between Basatanya–Urziceni occur at the same rate as the within-cite connections at Urziceni, showing that the Urziceni population is so larger that individuals from it are no more closely related to each other than they are to Basatanya individuals.

Both at Basatanya and Urziceni-Vamă, pottery styles do not show a correlation with signs of inbreeding, leading us to conclude that the practice of consanguineous marriages was not associated with distinct pottery stylistic traditions of the sampled communities. It is, however, noteworthy that the outlier individuals observed on PCA bear no or only limited (4-8 cM) ROH signals.

These shorter observed homozygous segments (4-8 cM) attest to background relatedness, typical of populations with low recent population size[66] (N_e). We calculated the *Ne* based on the length distribution of ROH 4-20 cM (likely arising from co-ancestry mostly within the last ~50 generations) across comparative N and CA populations. We also calculated diploid N_e based on IBD segment sharing of high quality genomes (see Methods) (**Fig. 3B, Table S6**). During the MN-CA periods, we tracked fluctuations in hapROH *Ne* over time in the Polgár microregion (**Table S6B**), and over 30% larger Ne in Urziceni compared to Polgár-Csőszhalom and Basatanya. However, we observe a decrease in effective population size of the Polgár area ECA communities in the IBD-based *Ne* calculations (by approximately a quarter at Basatanya compared to the LN). This decreasing trend of IBD-*Ne* signals appears to be restricted to the Polgár microregion, as compared to Basatanya it was 6-7x larger at Urziceni-Vamă and 2-3x higher in the mass grave from Potočani (Croatia), both belonging to contemporaneous ECA communities (**Table S6C**).

As a comparison, we added the well-studied N Gurgy (France) dataset to the *Ne* analyses, which is understood as a patrilocal population with controlled female exogamy, where most pairs of parents were related to each other via co-ancestors within the preceding 5-30 generations[44]. The population stability recorded there, with the relatively even distribution of the ROH segments and on extensive pedigrees[44], stands in contrast to what we observe on the GHP. When comparing the ROH patterns of Gurgy with the two well-recorded communities of this study, we find that Basatanya's ROH patterns are more varied, including both individuals without ROH and others with longer homozygous segments. Consequently, community organization seems to have been unstable at the LN-ECA period of the CB, when the ancestors of the Basatanya population experienced a relatively recent effective population decline (apparent in an abundance of 12-20 cM ROHs, and numerous IBD sharing), most plausibly reflecting a closed community organization at Basatanya compared to the more open system at Urziceni-Vamă.

Genetic network of Neolithic-Copper Age individuals and communities

To assess the genetic structure and the genetic network of the LN-ECA populations in relation to other contemporaneous populations, we inferred and analyzed long shared haplotypes, commonly referred to as IBD segments, based on phased and imputed autosomal haplotypes using the ancIBD software[51] (see Methods).

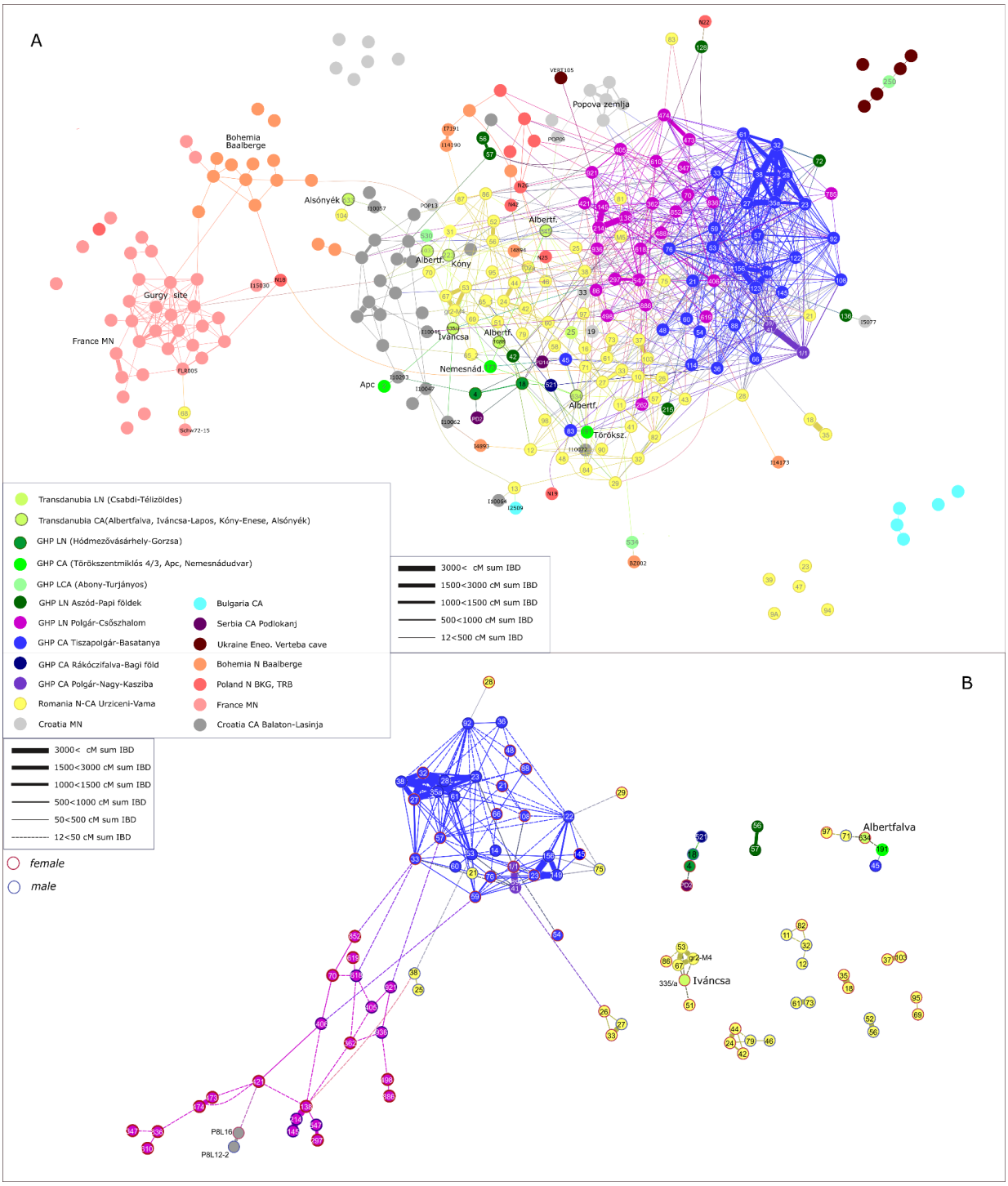

Figure 4. Network of IBD segments shared between the individuals from the Carpathian Basin and other roughly contemporaneous Neolithic and Copper Age individuals of Europe. We used a ForceAtlas2 layout in Gephi software v0.10.1 for both graphs (see Methods), which considers the number of connections (centrality) of each node (genome) and where edges (IBD connections) between nodes act like springs. The total length of IBD sharing was used to weight the edges as indicated in the legend. Thicker edges represent closer relatives, including familial relationships between individuals, while thinner edges depict more distant connections. Numbers in the nodes indicate grave number at the given site in this study, while reference samples are labeled by lab codes as provided in AADR[67] v54.1. **A: Network of sharing 1x12 cM IBD segments between individuals from the Carpathian Basin and other roughly contemporaneous Neolithic and Copper Age individuals of Europe. B: At least 2x12 cM IBD segments connectivity shown between the Late Neolithic and Copper Age individuals in the Carpathian Basin.** Genetic sex is indicated by contours of circles in different colors. These networks

illustrate how interconnected the CB individuals were with each other and their adjacent regions (A), and how this network even fragments communities of the same sites with stricter filtering, zooming into the GHP area (B), while still keeping Basatanya, Polgár-Nagy-Kasziba and Polgár-Csőszhalom connected.

Using all individuals with adequate quality data (see Methods) and applying a filter for at least 12 cM shared segments between two individuals, the ForceAtlas2 algorithm[68] produces a dense graph of the CB LN-CA samples showing scattered IBD sharing with N populations of France, Poland and Bohemia and a more intensive network with MN and CA individuals in Croatia (**Fig. 4A**).

The degree centrality (*k)*, representing the number of links held by each node, calculated from the network shown in **Figure 4A,** averages 0.077 for Basatanya—the highest value among the sites of this study (**Table S8**)—placing it in a core position within the network. Despite only including individuals from 19% of the excavated graves, Basatanya exhibits nearly three times more within-cemetery IBD connections than Urziceni and twice as many as Csőszhalom (*k*w, **Table S8A**). Nevertheless, it is also linked with the contemporaneous nearby Polgár-Nagy-Kasziba site and the LN population of Polgár-Csőszhalom**.** All analyzed individuals from these three sites – except graves 45 and 83 from Basatanya and 488 at Csőszhalom – form a single cluster based on Leiden community detection applied on the same network[69], including 17 individuals from the Urziceni-Vamă cemetery embedded within this cluster (**Fig. S5, Table S7**). Analyzing the sites further as modules of the network, we demonstrate that other LN and CA genomes from present-day Hungary and the majority of the Urziceni-Vamă community exhibit less dense within-module clustering than Basatanya and Csőszhalom (**Table S8A**). Although, 48% of the graves excavated at Urziceni-Vamă were part of the IBD analyses, some Urziceni-Vamă individuals (graves 9A, 23, 18, 35, 39, 68) show no connections to anyone on the graph or appear as outliers within their community (**Fig. 4A**), consistent in four cases with the PCA outlier detection as well (**Table S3**).

Late Neolithic community organization on the Great Hungarian Plain and the LN-CA transition in the Polgár microregion

Twenty-nine individuals (16 females and 13 males) had high enough quality data for IBD analysis from the LN Polgár-Csőszhalom site. They show high within-site (within-module, *k*=0.05) connectivity, but even stronger between-site (between-module) connections (*k*B/*k*=0.732, representing the ratio of between-module connections to total connections, **Table S8**), documented also by the maximum IBD segment length (strength) ratios (**Figs. S13-14**). We measured cliques within the network as complete subgraphs—subsets of at least three nodes where each pair of nodes is adjacent, forming a complete subgraph. Clique analysis reveals that Polgár-Csőszhalom individuals are more likely to participate in between-module cliques (103 cliques) compared to within-module cliques (24 cliques), with females predominantly driving these inter-site connections (p=0.0015 in Chi^2 test, **Table S8-9, Fig. S13**). Three small kindreds/pedigrees identified at Polgár-Csőszhalom are detailed in the Supplementary Information, Chapter 2 (**Fig. S9**). Individuals from LN Polgár-Csőszhalom maintain links in the IBD network with the ECA Basatanya cluster, even under strict filtering (2x12 cM; **Fig. 4B**). This continuity is observed through three males (graves 406, 618, 921) and

one female (grave 852), suggesting local continuity in the LN-ECA transitional period. The three youngest graves (graves 421, 498 and 619) at Polgár-Csőszhalom do not represent direct connections with Basatanya, suggesting that the interaction between the populations occurred earlier than the end of the Polgár-Csőszhalom site. The maternal and paternal haplogroup frequency distribution at the two sites are consistent with maternal continuity (p=0.421), while the paternal composition differs significantly (p=0.022) (**Fig. S10a-b**, **S11, Table S10**).

Early Copper Age community and cemetery organization on the Great Hungarian Plain

Tiszapolgár-Basatanya and Polgár-Nagy-Kasziba sites

We sampled 14 females and 16 males from 154 ECA graves at Basatanya (**Fig. S6C,** see Methods). Among them 29 were suitable for IBD analysis. The Basatanya population stands out with the highest degree centrality, clustering coefficient, and within-module strength among the three major GHP sites, indicating a highly interconnected community (**Table S8**). Many of these individuals are part of extended pedigrees, with close biological relatedness reflected in long stretches of IBD (details in the next sections). Nevertheless, a significant portion of the population also maintains between-module connections (**Figs. S13-14**).

A distinctive feature of the Basatanya community is its resilience in the IBD network, even when stricter filters are applied (requiring at least 2x 12 cM or longer for network inclusion; **Fig. 4B**). The Leiden algorithm also connects individuals with different pottery styles, with only graves 83 and 45 not being included (**Fig. S5, Table S7**). Mapping the weighted within-cemetery 2x12 cM IBD connections, we observe a high level of interconnectedness among most individuals (average degree is 7.24, graph density is 0.259), with the exception of those buried in graves 83 and 45 (**Fig. 5**). Although the genomic and uniparental genetic make-up of these two IBD outlier males is not unusual for the time or region, their IBD connection patterns point outside the site, toward Urziceni-Vamă and ECA populations on the southern GHP and Transdanubia. The male in grave 83 had no detected ancestor or descendant in the community, whereas male 45 has a single 12 cM IBD connection to an individual in grave 36. Their limited ROH signals suggest also that they were likely outsiders, not part of the typical closed community patterns seen in the Basatanya community. Other individuals with high ROH are scattered throughout the cemetery, without any clear association to family ties or pottery styles (**Fig. 5**). Despite these two male IBD outliers, analyzing the 72 cliques observed at Basatanya, we found more female connections between sites, whereas males predominantly participated in cliques within the Basatanya community (p=0.0036, Chi^2 test, **Fig. S15C**).

Through a combination of READ, KIN, ROH, and IBD analyses (see Methods), we identified three biological kindreds at Basatanya, two of which represent larger pedigrees (**Fig. 5**), linking graves in relative proximity in different sections of the cemetery. Family A, located in the western section, is a complex, extended kin group buried with Tiszapolgár-style pottery. Through READ and KIN analyses, we identified the nuclear family consisting of a mother (grave 27), father (grave 28), son (grave 35a), and daughter (grave 32). IBD and KIN results further indicate that the individual in grave

23 is in second-degree ancestral relation to grave 28 and third-degree paternal relative of the male in grave 35a. The Y-chromosome lineage I2a1b1a(2b)-L1316/Z161 (ISOGG v15.73) can be traced across at least three generations in this pedigree, and was detected in precursor Polgár-Csőszhalom community too. This lineage was present in 10 out of the 16 males analyzed at Basatanya site (Fig. S12). Additionally, another paternal lineage, C1a2b-Z44491, appears across two distant generations within this pedigree. In contrast, four distinct mitochondrial lineages were found in family A (**Fig. 5, Fig. S10 Table S1**).

Family B is associated with Bodrogkeresztúr-style burials and includes a mother (grave 123) and her two sons (graves 149 and 156), along with three more distant relatives (graves 59, 76, and 122), with grave 122 representing a descendant through the maternal line. Interestingly, grave 76 contained Tiszapolgár-style inventory, linking the two pottery styles through biological relatedness. We found a marked difference in ROH between the two siblings in graves 149 and 156, with 149 likely being the offspring of close relatives (probably 3rd-degree related, e.g. first cousins, as seen in **Fig. S4**). Taking into account the significant variability in ROH expected from simulations[66], the observed IBD/ROH patterns, and our kinship analyses, 156 could also be an offspring of the same parents (see Supplementary Information Chapter 2).

Compared with strontium ($^{87}Sr/^{86}Sr$) isotope data[42], we found that all males in families A and B had on a microregional scale local strontium values, while the individuals in graves 114 and 145, who had no close biological relatives, exhibited non-local strontium isotope values in their first molars; thus the genetic and isotopic analyses converge in showing these individuals are outliers. The uniparental, genetic relatedness, IBD clique analyses and the strontium evidence all speak for a patrilocal and patrilineal community organization.

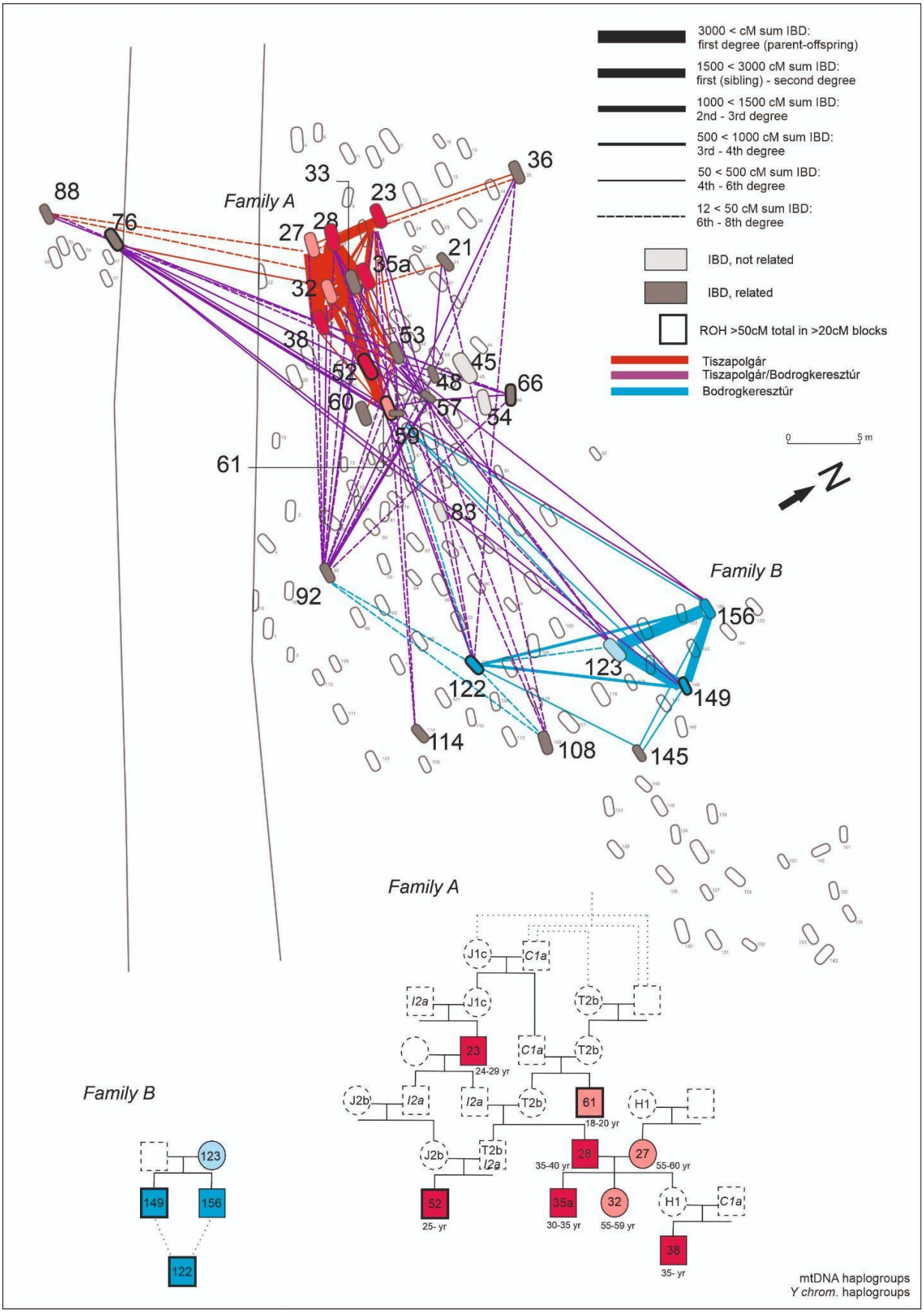

Figure 5. Map of Tiszapolgár-Basatanya cemetery with biological relationship between the buried individuals. All pairs with at least 2x12 cM IBD connections are marked on the cemetery map. Edges

are weighted according to the total length of IBD sharing, as indicated in the legend. Colors of the edges represent different pottery styles: Tiszapolgár (red), Bodrogkeresztúr (blue) and lines connecting the two (violet). Family A and B are represented as pedigrees. The pedigree was reconstructed based on KIN, READ, IBD, ROH analyses and mtDNA and Y haplotype sharing, with the black outline highlighting graves and individuals in the pedigrees bearing at least 50 cM sum IBD or ROH over 20 cM in length. Color shades of the individuals indicate local $^{87}Sr/^{86}Sr$ signal (vivid color) or lack of Sr isotope data (faint symbols). The presented pedigree of Family A is the simplest version explainable by the results. Besides this, several further versions could be conceivable, considering the missing family members. Family A and B have also further distant (3rd+ degree, such as graves 59 and 145 in Family B) members with uncertain positions in the pedigree that are shown by IBD connections on the cemetery map. In addition to the presented families A and B, we discovered third-degree relatives using KIN (in graves 57 and 92) in family C. IBD results suggest that the relatives had an avuncular 2nd relationship or they are third generation relatives (**Figure S16**).

The Polgár-Nagy-Kasziba site is clearly embedded within the Basatanya – Polgár-Csőszhalom IBD cluster, with individuals in graves 1/1 and 39/41, displaying significant inter-site connections. We sampled a nuclear family here, consisting of a mother (grave 38/40), a father (grave 39/40), and their 3-4-year-old daughter (grave 1/1). The daughter is a 6-8th degree relative of a male in grave 23 at Basatanya, as revealed through IBD analysis (**Figures 4B and S16**). Both father and daughter show multiple connections to individuals in Basatanya, making it challenging to determine the directionality of the connections between these chronologically and geographically close sites (Supplementary Information).

Urziceni-Vamă site

From the ECA Urziceni-Vamă cemetery, we sampled 43 females and 34 males (**Fig. S6D**, Methods), with 63 individuals among them suitable for IBD analyses. Despite our coverage of this cemetery being most complete–with 47.3% of the graves from an almost completely excavated site–it forms the loosest IBD network, with the fewest internal connections among the three major GHP sites (see Methods, **Fig. 4A, Table S11**). The average number of connections within Urziceni-Vamă itself is 2.96, with a graph density of 0.125, which is significantly lower than the density of the connections of the contemporaneous Basatanya (~1/4 of that, **Table S8**). In the 2x12 cM IBD network, many connections disintegrate (**Figure 4A** vs. **B**), leaving predominantly 1st-2nd degree relationships dominating the second network, and only a few Urziceni individuals remain connected to the Polgár-Csőszhalom – Basatanya cluster.

Eight close biological kindreds were detected at Urziceni-Vamă in the IBD analyses, with three additional kindreds identified through READ and KIN analyses, which allow study of individuals with lower quality data (see Methods and **Table S1**). These pedigrees consist of 2-3 individuals, with most connections involving adult siblings, and no evidence of biological fathers. Despite the presence of close relatives, kindreds do not form substructures on the IBD graph, mapped on the cemetery, except for three brothers from graves 4, 53, and 67 (**Fig. 6**). This is also reflected in that some individuals have high between-module strength, but their degree centrality remains low (**Figs. S13-S14**).

Eighteen cliques were counted within Urziceni-Vamă, with four including children, which is ~⅕ of the between-module cliques observed at this site (79 total). Importantly, both male and female individuals exhibit high between-site connections, with no significant difference in connectivity between sexes (p=0.23, **Fig. S15**). Interestingly, two females in a mother-child relationship (graves 18 and 46) entered the community from external

groups with elevated Eastern Hunter-Gatherer (EHG) components (9% and 14%, respectively).

After subdividing the ECA Urziceni-Vamă dataset based on pottery styles (Bodrogkeresztúr and Bodrogkeresztúr-Salcuța), we observe significant differences in network connectivity. Individuals buried with Bodrogkeresztúr-style grave goods (n=48) form within-style cliques (7 cliques with 22 nodes), while those (n=14) with Bodrogkeresztúr-Salcuța-style grave goods do not form within-style cliques. The (k_B/k *style)* ratio is significantly larger for the individuals in Bodrogkeresztúr-Salcuța-style graves compared to the values of the individuals in Bodrogkeresztúr-style graves in line with the style specific pattern of the cliques ($p<0.001$ in Fisher's exact test, **Table S8**, **Fig. S13**). Notably, pottery styles do not strictly align with family structures, as demonstrated by mixed-style cliques (5 cases) and the appearance of mixed Bodrogkeresztúr-Salcuța-style pottery in only one grave per families C, E, F, J and K. Whereas the available radiocarbon data do not make it possible to separate chronologically the two styles, the archaeological chronology, the pedigrees and the age at death of certain individuals speaks for the A version of the family trees F and K. This phenomenon might indicate that the Bodrogkeresztúr-Salcuța-style pottery appeared in the cemetery quickly, within one or two generations time range.

The uniparental composition reflects the site's broader and looser network. Altogether 44 different mtDNA subhaplogroups were described (see details in SI, **Fig. S10**), including rare haplogroups pointing toward the southeast, such as N1b (based on its rarity in the area and era and its general appearance in the Near East[70,71], and others (e.g. J2b1, J1c) with parallels to the Varna and Pietrele sites[63]. The paternal genetic diversity is also larger at Urziceni-Vamă than at Basatanya (**Fig. S11**). Y-lineages include C2a, which is associated with Mesolithic people and H2 and G2a elements of Neolithic origin[1], as well as potentially new external influences such as R1b from steppe-outliers (graves 12 and 79).

	Tiszapolgár-Basatanya ECA	**Urziceni-Vamă ECA**	**Significance in difference between the sites**
Sex distribution (in excavated graves)	43.05% male, 35.76% female, 19.87% children, 1.32% symbolic	27.88% male, 43.27% female, 21.15% children, 8% not determined	
Radiocarbon model for the cemetery (68.3%)	4355-4270 to 4150-3990 cal BCE	4285-4130 to 4035-3940 cal BCE	
Pottery styles	Tiszapolgár, Bodrogkeresztúr	Bodrogkeresztúr, Salcuța	
% ind. analyzed in PCA (n analysed /	19.4%	54.5%	

excavated)	(30/154)	(72/132)	
% ind. analyzed in IBD (n analysed / excavated)	18.8% (29/154)	47.3% (61/129)	
% PCA outliers	6.6% (2/30)	11.1% (8/72)	p=0.49
% ind. with EHG Z score>2 in qpAdm	3.3% (1/30)	24.6% (16/65)	**p=0.026**
IBD kw/k	0.433	0.313	p=0.86
IBD cliques W/B	0.97 (423/432)	0.49 (50/102)	**p=0.0001**
IBD cliques Chi^2 W/B module male/female diff.	**p=0.0036**	p=0.2289	**p= 1.79×10^{-5}** (males W/B)
IBD Ne	6000-7500	38,900-50,900	6.65x diff.
n Y-chr diff.lineages/ total	18.8% (3/16)	36.6% (11/30)	p=0.357
n diff. female subhaplogroups/ total	60% (18/30)	58% (44/75)	p=1
n Inbreeding (min 50cM, >20cM blocks)	20% (6/30)	2.8% (2/69)	**p= 0.009**
Cemetery organization	3 kindreds with 2-6-8 members, related inds. buried in proximity of each other	11 kindreds with 2-3 members only, buried also in larger distance, no fathers in pedigrees	
Community structure	patrilineal, patrilocal	non patrilineal, non patrilocal	

Table 1. Comparison of the two ECA communities through the most important metrics and characteristics. P-values were calculated from Z-test, two-tailed test, Fisher's exact, and Chi^2 tests. Individual data of this table is presented in **Tables S1, S3, S5, S6, S8, S10.**

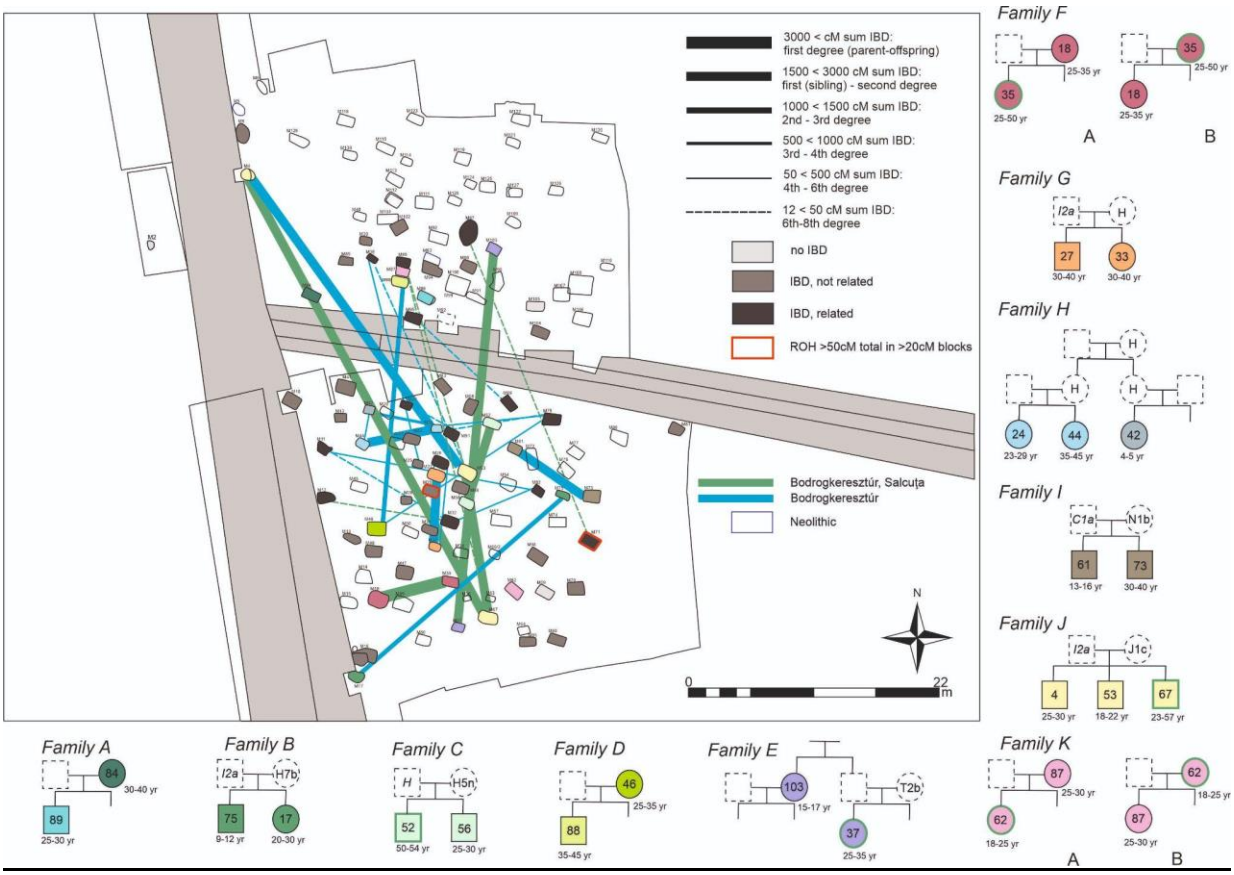

Figure 6. Biological relatedness in the Urziceni-Vamă population. At least 2x12 cM IBD connections are marked on the cemetery map. Edges are weighted according to summed IBD, as indicated in the legend. Colors of the edges represent IBD connections among individuals buried with different pottery styles: Bodrogkeresztúr (blue), Bodrogkeresztúr-Salcuța (green). Pedigrees shaded both on the map and the pedigree consistently. The pedigrees were reconstructed based on KIN, READ, IBD analyses and mtDNA and Y chromosomal lineage sharing (Table S12). Highlighted graves and individuals on the map bear at least 50 cM sum IBD of ROH over 20 cM in length. The presented pedigree of Family E and H are versions consistent with the results, but several alternative versions are possible, considering the missing family members. Green outlines of the symbols in the pedigrees signal Bodrogkeresztúr-Salcuța-style pottery in the grave.

Discussion

We present genome-wide data from 125 individuals from 4800-3900 BCE, corresponding to the LN-ECA periods of the Carpathian Basin, focusing on agro-pastoral communities of the Great Hungarian Plain.

Our analyses show that the genetic ancestry of LN GHP communities derives from the preceding Middle Neolithic (MN) populations, with a noticeable decline in genetic diversity over time. These populations had minimal Eastern Hunter-Gatherer (EHG) ancestry, aligning with the genetic trends described by Lipson et al.[4]. During the ECA, however, we observed differences in genetic composition among the GHP sites. In the Polgár microregion of Hungary, f_4-statistics, qpAdm and IBD analyses (**Figure 2 and 4**) consistently indicate that the cultural changes between LN and ECA occurred without the introduction of larger immigrant groups. The ECA population is predominantly descended from local LN communities, sharing mostly maternal lineages and exhibiting

even greater genetic homogeneity, which likely indicates increased isolation. Only two adult female individuals (graves 21 and 145) from the Basatanya cemetery were identified as outliers in their genetic ancestry composition. These unrelated individuals likely joined the local community during adulthood, and connected to it through their descendants (see IBD analyses). In contrast, contemporaneous ECA sites like Urziceni-Vamă show greater genetic diversity (PCA, f_4, qpAdm) and stronger connections to Eastern European regions. Similar patterns are observed at Copper Age sites such as Pietrele and Varna in southeastern Europe[63]. Interestingly, most of the outliers at Urziceni-Vamă are females (n=8 out of 11), and their varied ancestry does not correlate with pottery styles, burial practices, or grave goods.

Evidence suggests that individuals from the Volga Basin region may have sporadically appeared in the ECA of the GHP, as shown by the burial at Csongrád-Kettőshalom in Hungary dated at 4330-4070 cal BCE[72,73]. However, these cases appear isolated in the currently available dataset. For instance, the Decea Mureşului cemetery from Romania, which culturally reflects an early infiltration of Eastern steppe groups as well, showed genetic results in one analyzed male (grave 10) that are similar to early farming populations of the GHP[74].

The IBD networks reflect the diverse internal organizations of the studied communities. We observed that closely related individuals were buried near one another at both LN Aszód-Papi földek and Polgár-Csőszhalom (**Figs. S7 and S9**). While similar grave goods were noted (**Table S2**), the low number of detected biological connections makes it challenging to assess whether these similarities are due to kinship or mortuary customs tied to local, age, or gender-specific practices. Females at Polgár-Csőszhalom played a significant role in connecting different sites, suggesting extensive external relationships, as indicated by the high number of inter-site IBD connections. Polgár-Csőszhalom is most closely connected to Basatanya from the same microregion but separated by a 300-year gap within the dataset, sharing long IBD segments with Basatanya, and more connections with other sites than within (**Fig. 3B**). We detected 6-8th degree or more distant relations on maternal lineages of three males and one female at Polgár-Csőszhalom to all families at Basatanya (**Fig. 4**). As these individuals at Polgár-Csőszhalom were not among the chronologically youngest graves, we interpret it as some ancestors of the Basatanya populations diverged earlier from the Csőszhalom community than the Polgár-Csőszhalom site was abandoned.

In the LN period, the effective population size (N_e) was uniformly large across the studied sites (**Table S6**). In contrast, the ECA presents a more complex scenario. While N_e remained large or increased at Urziceni-Vamă and Potočani, the Polgár microregion experienced a significant N_e decline, as shown by IBD N_e analysis. A relatively large N_e suggests the presence of extensive marriage networks at Urziceni, while a 6.6x smaller N_e can imply more isolated, closed communities at Basatanya. Whether this contraction was also accompanied with population decline on the GHP, needs to be verified on further sampled ECA groups.

The closure of ECA communities in the Polgár microregion is also supported by the numerous third-fourth-degree kinship connections within the Basatanya cemetery (**Table S11-12, Fig. S16**). This pattern likely extended beyond just one site and reflects a broader microregional trend, exemplified by strong IBD ties between Basatanya and the nearby Polgár-Nagy-Kasziba site (located approximately 10 km away, **Figures 1, 4-5**), indicating a highly interrelated and connected community system. At Basatanya,

despite the use of various pottery styles and the identification of at least three pedigrees, the IBD network reveals a strongly interrelated cemetery community. In terms of social organization, females were not generally exogamic; however, IBD analyses show that they play a key role in linking sites across the region, while males were primarily responsible for intrasite connections at Basatanya **(Table S8, Fig. S15)**. This, along with the uneven distribution of the uniparental lineages, the 8-member family structure and the stable isotope evidence, points to a patrilineal and patrilocal society at Basatanya.

A long-term comparative analysis of human populations over the past 15,000 years indicates that inbreeding has been relatively rare. Two distinct forms of inbreeding can be identified: one resulting from genetic drift in small populations and the other due to consanguinity[75]. Inbreeding due to small population size was more prevalent in Paleolithic and Mesolithic hunter-gatherer groups and in isolated island communities[76], though studies suggest that deliberate avoidance of consanguinity was common[77]. As Neolithic population growth reduced the likelihood of inbreeding, its occurrence became even rarer. When it did occur, it was typically within farming communities under conditions of strict social stratification, elite isolation, and inheritance of private property[46,78–80].

The prevalence of first-cousin marriages exhibits considerable variation across cultures, spanning a spectrum from complete prohibition to socially preferred practices in modern societies. A longitudinal analysis indicates that the preference for first-cousin marriages is a relatively recent phenomenon, influenced by various social factors, including the inheritance of private property[75,81–83].

In the case of the ECA Basatanya cemetery, the elevated rates of runs of homozygosity (ROH) are indicators of first-cousin marriages. Similar trends can be seen at Malta and the Aegean islands, where the islands' isolation likely contributed to inbreeding patterns[46,76]. Contrary to these examples, the Basatanya community does not share the geographic isolation but archaeological evidence points to a densely populated network of small settlements, although the contemporaneity and the precise relationship between these settlements and cemeteries are still unclear. This suggests that the isolation of the Basatanya community was more likely sociocultural in nature rather than geographical.

The prevalence of consanguineous marriages in Family A at Basatanya could indicate a deliberate strategy to preserve and transfer wealth within the family lineage. However, we did not observe clear correlation between biological relatedness and the wealth of burials, based on the analysis of individuals sharing long IBD segments and those who do not. When considering the grave goods of Family A across four generations and the members of Family B, it is possible that access to prestige goods from distant areas gradually diminished, or the family lost interest in them, as the community became more isolated. This is evidenced by the absence of such items in the graves of members of Family B (Graves 123, 149 and 156) and in the younger generations of Family A (Graves 32, 35a and 38). Another possible interpretation is that only one child inherited the property or rank, as evidenced by some richly furnished burials at Basatanya (**Figure 5, Table S2, SI**). However, since only a limited portion of the cemetery has been analyzed, these interpretations remain speculative until further analysis can be conducted on a larger sample. In summary, the Basatanya cemetery appears to reflect a highly cohesive, patrilineal community, where first-cousin marriages were common over generations. The burial tradition of placing spouses (graves 27-28) and close relatives

near one another echoes similar patterns observed in the MN[6] and LN periods of the region.

The ECA Urziceni-Vamă cemetery presents a strikingly different social structure compared to the Basatanya cemetery. Despite partly identical and partly similar burial traditions, it was used by a more open and probably larger community. One key observation is the lack of direct lineal relationships spanning more than two generations, while sibling relationships are common. Fathers are notably absent from parent-child burial relationships, and even closely related individuals are often buried far apart from each other. In five cases, Bodrogkeresztúr-Salcuța-style pottery appeared alongside Bodrogkeresztúr-style pottery within small pedigrees, but no pedigrees were found where all members used Bodrogkeresztúr-Salcuța-style exclusively. This phenomenon might indicate that the Bodrogkeresztúr-Salcuța-style pottery appeared in the cemetery quickly, within two generations.

Consanguinity at Urziceni was rare contrasting with the frequent first-cousin unions seen in Basatanya, indicating that cousin marriages were not a common practice here. We found no clear correlation between biological relatedness and the grave goods. The small kindreds and absence of strong biological ties between individuals buried near each other suggest a community not organized around strict patrilineal rules. Although it is a still possible explanation that the Urziceni-Vamă cemetery was used for a considerably shorter time than the Basatanya cemetery, the currently available AMS dates suggest otherwise. The cemetery likely remained in use for 140-300 years at 68.3% probability, a timeframe that aligns closely with the Basatanya cemetery's estimated duration of 160-325 years at 68.3% probability, based on Bayesian-modeled radiocarbon dates **(Figures S20, S22)**.

The differences in N_e observed between the LN-ECA GHP and among different sites of the ECA hint on changes in the community organizations and potential population contraction or isolation in the Polgár microregion. While severe internal or external social conflict might have caused such phenomena, no traces of violent action was found at the sites of the Polgár microregion. Furthermore, internal social conflict and the growth of social inequality might be interrelated. The CA is often characterized by the appearance of institutionalized or hereditary social inequality[15,16]. One of the main arguments is usually the wealth of certain CA burials, e.g. the Varna cemetery[84,85], but the stable isotope analyses could not detect any significant difference among the ECA individuals on the GHP[86]. The trend of abandoning the LN tell and large flat settlements on the GHP can be witnessed on several sites. According to a radiocarbon dating project of the ECA-MCA on the GHP, the establishment of ECA formal cemeteries can be dated after 4400-4350 cal BCE[22] (**Figures S19-20, Table S2)**. Despite the hundreds of available AMS data for the LN and ECA periods on the GHP, only a few can fill the gap between 4450 and 4350 cal BCE. There are few burials and tiny farmsteads consisting of two or three buildings from this transitional period (e.g. Polgár-Nagy-Kasziba, Rákóczifalva-Bagi föld, site 8, Vésztő-Bikeri)[22,87]. These radiocarbon data corroborate drastic population changes at the turn of the LN and ECA in the Polgár microregion. However, further site-specific investigations are needed to reveal underlying causes of the demographic and social changes, which are crucial for understanding the broader transformations that occurred during this transitional era of the Carpathian Basin.

Conclusion

Our study demonstrated population continuity at the transition between the LN and the ECA on the GHP, despite cultural changes. Accompanying this phenomenon, ECA communities became more isolated, and first-cousin marriages became a common practice, at least on the Northern GHP. However, this cannot be generalized on the level of archaeological culture, as at least one culturally similar, contemporaneous community on the Eastern margin of the GHP did not follow the same trend. The differing internal organization and genetic connection patterns of the Basatanya and Urziceni-Vamă cemeteries suggest that the later might have been a common burial place for a larger community with more distant connections, while the Basatanya cemetery served a relatively closed, tightly interrelated and probably smaller community.

These case studies highlight the risks of overgeneralizations in both archaeological and genetic research, where broad conclusions about an entire archaeological culture, region or period are drawn from the analysis of a single site. Behind the diversity of material culture and cultural traditions, there might have been a range of cultural and social relationships, different dynamics, only one of which were biological or social. The example of the ECA cemeteries we have presented illustrates that archaeologically similar phenomena can also be underpinned by local communities organized in fundamentally different ways.

Material and Methods

Materials

We sampled eight sites from present-day Hungary and Romania (see Supplementary Information for full archaeological details).

Aszód-Papi földek is a LN site in the Gödöllő hillside, Northern Hungary, where among the settlement features 224 graves were found. The burial activity can be dated from 4760-4700 cal BCE (68.3%) to 4710-4640 cal BCE (68.3%) based on Bayesian-modeled AMS radiocarbon data. The material culture reflects intensive interaction both with the GHP and Transdanubia. Culturally, the Aszód-Papi földek site belongs to the East Transdanubian group of the Lengyel complex but both Lengyel- and Tisza-style pottery appear at the site[25,88–91].

The *Budapest-Albertfalva-Hunyadi János út* site is situated on the western bank of the Danube River in southern Budapest. Four ECA features contained human remains with Ludanice-style sherds, stone tools and a grinding stone[92,93] (**Table S2**). Based on Bayesian-modeled AMS dating, the burials can be dated from 4395-4180 cal BCE (68.3%) to 4160-3930 cal BCE (68.3%) (**Figures S17-18, Table S2**).

The *Iváncsa-Lapos* site is located by the Danube River on a 560–570 m hill range. A single grave (Feature 335/A) dated to the ECA was found at the site. It can be dated to 4325-4070 cal BCE (68.3%) (**Table S2**).

The complex site of *Polgár-Csőszhalom* is situated in Northeastern Hungary, in the middle of the so-called Polgár Island in the Upper Tisza Region. It consists of a tell surrounded by a multiple rondel, a smaller double rondel and a large flat settlement[94]. The LN site is characterized by the co-occurrence of various pottery styles characteristic of the Tisza-Herpály-Csőszhalom and Lengyel complexes. The graves, located in various areas of the flat settlement, formed small groups consisting of two to five burials. The majority of the graves were dug in association with houses. Based on Bayesian-modeled AMS radiocarbon measurements, the graves of the single layer settlement can be dated from 4799-4769 cal BCE to 4673-4652 cal BCE (68.3%)[95].

The *Polgár-Nagy-Kasziba* site is situated in Northeastern Hungary, in the so-called Polgár Island in the Upper Tisza Region. Four ECA graves were found in the northern edge of the excavated area. The graves contained a large number of grave goods including Tiszapolgár-style pottery[96] (Solnay 2018). Based on Bayesian-modeled AMS radiocarbon measurements, the graves can be dated from 4520-4360 cal BCE to 4335-4270 cal BCE (68.3%)[22].

The *Rákóczifalva-Bagi föld, Site 8* is situated on the eastern bank of the Tisza river in the Middle Tisza region. Traces of an ECA settlement (a building, two wells and several pits surrounded by a ditch) and nine contemporary burials were found. Stylistically, the pottery of the graves bears the characteristics of Tiszapolgár-Kisrétpart-style[22,97]. Based on Bayesian-modeled AMS radiocarbon measurements, the site can be dated from 4525-4380 cal BCE to 4335-4225 cal BCE (68.3%)[22].

The *Tiszapolgár-Basatanya* site is situated in Northeastern Hungary, in the so-called Polgár Island in the Upper Tisza Region. 154 ECA and one LN graves were excavated

here in the 1950s, which are considered to present the complete CA cemetery. Bognár-Kutzián distinguished period I (Tiszapolgár, ECA) and period II (Bodrogkeresztúr, Middle Copper Age) and a few graves with transitional character (Bognár-Kutzián 1963), which chronological periodization have been later reevaluated as in time overlapping pottery styles[22,98]. Based on Bayesian modeled AMS radiocarbon measurements, the cemetery can be dated from 4355-4270 cal BC to 4150-3990 cal BCE (68.3%) with a 135-305 years (68.3%) span of use (**Figures S19-20, Table S2**).

The *Urziceni–Vamă* ECA cemetery, with 3 N and 129 CA graves, is located in the free zone of the Romanian-Hungarian border. Based on Bayesian-modeled AMS radiocarbon measurements, the CA occupation of the site can be dated from 4285-4130 cal BCE to 4035-3940 cal BCE (68.3%) with a 140-300 years (68.3%) span of use (**Figures S21-22, Table S2**). The graves contained Bodrogkeresztúr and Bodrogkeresztúr-Salcuţa-style potteries[99–104].

Methods

Radiocarbon dating

AMS data were calibrated using the IntCal20 curve[105] and the OxCal 4.4.4. software[106]. Site-based Bayesian models were built taking into consideration vertical stratigraphy (**Figures S17-22**). Bayesian modeled AMS data (68.3% probability) were used for dating certain graves and sites.

Sampling

In the case of Budapest-Albertfalva-Hunyadi János út, Iváncsa-Lapos, Polgár-Nagy-Kasziba and Urziceni-Vamă sites, we collected samples from all available graves. In the case of Aszód-Papi földek, Polgár-Csőszhalom and Tiszapolgár-Basatanya, we aimed to sample 30 graves from each site. We selected such graves for sampling to represent both female and male, and burials with wealthy and modest grave-goods and different parts of the burial grounds.

Ancient DNA data generation

A total of 125 samples were prepared at the University of Vienna. Petrous bones were prepared following the method of Pinhasi et al.[107] by extracting the cochlea with a sandblaster followed by milling. Teeth were cleaned using a dental sandblaster and selected fragments powdered using a mixer mill. The powders were then shipped to the Harvard laboratory or extracted in Vienna directly and extracts shipped to Harvard.

At HMS (Harvard Medical School) and in Vienna DNA was extracted from the powders either with the manual procedure using spin columns[108,109] or at HMS with the automated ('robotic') procedure using silica magnetic beads and Dabney Binding Buffer on the Agilent Bravo NGS workstation[110]. DNA (from extracts generated in Vienna or Boston) was then converted into barcoded double-stranded partial Uracil-treated libraries[111] or USER-treated single-stranded libraries[112] using automation and then enriched in solution for sequences overlapping 1.24 million SNPs [1240k[113], Twist[114]] as well as the mitochondrial genome (either spiked into the 1240k or Twist reagent, or in an independent capture). For each library, we sequenced approximately 20 million reads of the enriched library (average 19.63M) using Illumina instruments [NextSeq500, HiSeq X]; we also sequenced several hundred thousand sequences of the unenriched library.

We generated additional data for six already published samples (**Table S1** I-ID_enhanced). For an additional 18 samples we generated deeper shotgun data (from the same libraries used for capture), with an average genome coverage of 3.8x using Illumina HiSeq X instruments (**Table S1**).

Bioinformatic analysis

Samples were sequenced to generate raw paired-end reads; these were prepared for analysis by performing the following steps: preprocessing, alignment, and post-alignment filtering to enable variant calling. Raw reads were demultiplexed by using identifying barcodes and indices to assign each read to a particular sample, prior to stripping these identifying tags from the sequences. Paired-end reads were merged into

a single molecule using base overlap as a guide, Single-ended reads were aligned to the hg19 human reference genome (https://www.internationalgenome.org/category/grch37/) and the basal Reconstructed Sapiens Reference Sequence[115] (RSRS) mitochondrial genome using the samse aligner of bwa[116]. Duplicate molecules were marked based on barcoding bin, start/stop positions and orientation. The computational pipelines with specific parameters are publicly available on github at: https://github.com/dReichLab/ADNA-Tools and https://github.com/dReichLab/adna-workflow. For calling variants, a pseudo-haploid approach is used at targeted SNPs, where a single base is randomly selected from a pool of possible bases at that position filtering by a minimum mapping quality of 10 and base quality 20, after trimming reads by 2 base pairs at both 5' and 3' ends to remove damage artifacts.

Eighteen samples were shotgun sequenced (deep sequenced) from Csőszhalom, Polgár-Nagy-Kasziba and Urziceni-Vamă with an average coverage of 0.74x.

Genetic sex determination

We determined molecular sex based on the ratio of Y chromosome to the sum of X and Y chromosomes sequences. Ratios <0.03 were interpreted as confidently female, and ratios >0.3 were interpreted as confidently male (**Table S1**).

X chromosome contamination in males

X-chromosomal contamination of male individuals was measured with two software packages: ANGSD (version 0.939-10-g21ed01c, htslib 1.14-9-ge769401) and hapCon (hapROH package version 0.60). The doCounts for the former method was run on X-chromosomal regions from positions 5500000 to 154000000, with base quality 30 and mapping quality 25. The toolkit's contamination executable was used to process the resulting file. Parameters were set to -b 5500000 -c 154000000 -d 3 -e 100 -p 1, and the file used by the -h parameter was included with the software by the developers.

For the second method, the required mpileup file was created with samtools[117] mpileup (version 1.10, htslib 1.10.2-3ubuntu0.1), base quality was set to 30, and mapping quality was set to 25. Parameters were used according to the official documentation. All the necessary reference files used were included by the developers in the software package.

Mitochondrial DNA analyses

Mitochondrial haplogroups were determined using Haplogrep[118] v2.4.0, which utilizes Phylotree mtDNA tree[119] Build 17. We estimated mitochondrial contamination using contamMix[120].

We aligned sequences in each group with ClustalO within Ugene[121]. The alignments were checked and corrected manually where necessary. Comparing to the rCRS sequence, we deleted the following ambiguous positions: np 42, 57, 291–317, 447–458, 511–524, 568–573, 594–597, 1718, 2217–2226, 3106–3110, 3159–3167, 5890–5894, 8272–8281, 16184–16193. Median Joining network (nexus) were created with

DnaSP[122] and the figures were drawn with the PopArt[123] (Population Analysis with Reticulate Trees) program.

Compilation of data

Genotypes of 125+6 individuals were merged with previously published genotypes of ancient and modern individuals obtained with 1240K capture or shotgun sequencing data, downloaded from the Allen Ancient DNA Resource[124] . Especially important N-CA data were obtained from[4–7,63,125].

Principal component analysis (PCA)

We used the Eigensoft 8.0.0 package to perform principal component analysis with the Smartpca software[60]. Specifically, we projected the genotypes of the ancient individuals onto a West-Eurasian basemap of genetic variation using shrinkmode, which we calculated using the HO panel of the AADR. PCA plots were constructed in Python using matplotlib package[126]. To detect outliers in the main analysed groups based on PC1 and PC2 values, we used the Mahalanobis Distance approach[127,128], which calculates the distance of each datapoint from the group mean while accounting for the correlation between variables. First, we computed the mean vector for the group across PC1 and PC2, followed by the calculation of the covariance matrix for PC1-2. The Mahalanobis Distance for each individual was then calculated, considering both the deviation from the group mean and the covariance matrix. To assess statistical significance, we calculated the p-value for each individual's Mahalanobis Distance using the chi-squared distribution with two degrees of freedom. If the p-value was below the specified threshold (p=0.05), the individual was flagged as a potential outlier (**Table S3**). We used Levene's test and Bartlett's test in order to assess the homogeneity of variance between groups or sites, based on PC1-2. We used both, to cover normal and not normally distributed datasets adequately, and to check for equality of variances.

f_4-statistics

f_4-statistics was used in the form of f_4(Israel_Natufian, Turkey_Epipalaeolithic, Test, Mbuti) and f_4(Russia_EHG, Loschbour, Test, Mbuti), and f_4(CHG, Russia_EHG, Test, Mbuti), in order to test allele sharing differences between the test individuals and the given references. f_4-scores are expected to follow a normal distribution (**Figure S1**), outliers were detected with Z-score (**Table S4)**.

qpAdm modeling

Several 1-4 ways qpAdm models were constructed to characterize the study groups and individuals, to test for local continuity on the GHP, and to test for potential streams of migration into the region leading to ancestry outliers. Outgroups and source populations are listed in **Table S5**. We used as HG sources WHG (including as proxies Mesolithic genomes from England, and sites such as Loschbour, Bad Dürrenberg) and EHG (including Mesolithic genomes from Ukraine and Russian Karelia) along with CHG (Caucasus hunter-gatherer). Variable HG proportions have previously been shown to provide a powerful tool to distinguish farmer communities. The intermediate WHG-EHG genetic composition of the Serbian Iron Gates Mesolithic population has been described previously, anticipating the intermediate or dual composition of the local HG sources in the CB[5,129]. We excluded one of each 1st degree relative pair and low coverage

individuals (<50k SNP on the 1.15 million autosomal targeted SNPs) from these analyses. A two-sample t-test was used to test the null hypothesis that there is no significant difference (at the significance level of 0.05) in a given component between pairs of selected populations. We computed the Z-score using the Pandas library in Python, to test how far an individual's EHG or WHG component value deviates from the group mean for groups with n>3, assuming that the WHG data are normally distributed and the EHG components follow half-normal or truncated normal distribution (**Fig. S1**). Z-scores beyond ±2 indicate that the individual's value is significantly different from the group (in terms of standard deviations) (**Table S5**).

Determination of Y-chromosome haplogroups

Sequencing data (bam files) were screened with yleaf[130] for SNPs based on ISOGG Y-chromosomal tree v17 (https://isogg.org/tree/).

Kinship analysis - KIN, READ

To assess biological relatedness, three methods were used: READ, KIN, and ancIBD.

The READ analysis[49] can identify biologically related individuals up to the second degree based on the proportion of non-matching alleles, thus providing valuable information about close biological relatedness.

The KIN analysis[50] estimates the relatedness of a pair of individuals by analyzing the identical-by-descent segments they share and can accurately classify up to third-degree relatives. It includes additional models that adjust for contamination and detect inbreeding, and can differentiate between siblings and parent-child pairs.

Identity by descent chromosome segment sharing

ancIBD identifies IBD segments longer than eight centimorgan for aDNA data. This method can distinguish up to sixth degree relatives which are expected to share multiple long IBD segments and which provide strong signals for identifying close and distant biological relatives[51]. Samples were filtered based on coverage (at least 600k SNPs hit on autosomal targets). After filtering, a total of 26 samples from Polgár-Csőszhalom, 7 from Aszód-Papi földek, 29 from Tiszapolgár-Basatanya, 63 from Urziceni-Vamă, 4 from Budapest-Albertfalva-Hunyadi János út, and one sample each from Rákóczifalva-Bagi föld site 8, Polgár-Nagy-Kasziba and Iváncsa-Lapos each were used in the ancIBD analysis. We performed an ancIBD run with 269 individuals, adding samples from previously published articles, in order to observe the genetic connectivity with other European N and CA populations[4,7,53,54,62,113,125,131–142].

IBD segment sharing was visualized with Gephi 0.10.1 using different layout algorithms: ForceAtlas2 with 0.04 gravity and scaling=10[68] and manually constructed layout, which corresponds to the map of the studied sites. The Leiden algorithm was applied with the constant Potts model and 0.04 resolution with 10000 iterations[143] in Gephi 0.10.1.
After filtering the graph for edges > 12 cM, 269 samples remained in the LN-CA graph. We calculated the degree of centrality (*k*) of a node counting links connected to all other nodes in the network. This degree of centrality was normalized: each node was divided by the maximum possible number of connections it could have (which, in an undirected graph, is **N-1**, where **N** is the number of nodes in the graph).

Maximum IBD lengths were used as weights of the links attached to each node (w). These weights were analysed in two categories: "within strength" for the node represents the sum of edge weights connecting it to other nodes within the same module, whereas „between strength" for the node represents the sum of edge weights connecting it to other nodes to other modules by considering the archaeological site of the burial as a module.
The clustering coefficient (C) for each node in a graph is a measure of how interconnected a node's neighbors are. We used the degree_centrality and nx.clustering function of NetworkX and Gephi for defining C and k, and plotted with seaborn[144]. Statistical tests for the differences and similarities of these values were Welch's t-test and Fisher's exact test for *k* and the two-sample Kolmogorov-Smirnov tests for the distribution of the individual data (*k* and C) by sex[127].
We counted cliques (complete subgraphs) with minimum node size 3 within the whole graph (created in NetworkX), differentiating between and within-module cliques and nodes into both sexes. A within-site clique is a clique where all nodes belong to the same site. A between-site clique is a clique where nodes belong to more than one site. We performed the Chi-square test for each site and sex group separately, using the expected frequencies **(Table S8B).**

IBD segment sharing within and among the populations was also used to estimate differences in diploid effective population sizes of the communities, using the TTNe python package v0.0.1a0[145], filtering for non-related (12<100 cM) segments shared. 95% CI of N_e was calculated via maximum likelihood approach (as included in the TTNe algorithm).

Runs of homozygosity analyses

hapROH software[64] can detect signals of recent inbreeding or indicative of a small effective population size. The program was run with default parameters for all pseudo-haploid genotypes with at least 400k SNP covered. The N_e vignette of this program was also used for estimation of effective population sizes with confidence intervals. For the N_e calculations, we removed individuals with more than 50 cM sum ROH above 20 cM stretches. Statistical inferences were made using Fisher's exact test of the scipy.stats module and permutation test (with 10000 permutations) of the SciPy[127] and Numpy[146] libraries in Python.

Data availability statement

All data needed to evaluate the conclusions of the paper are present in the paper and/or the Supplementary Materials. Ancient genome sequences were uploaded to ENA (European Nucleotide Archive), and will be available upon acceptance of the paper. The AADR v54.1 dataset is publicly available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FFIDCW.

Ethics declaration

The individuals studied here were all analyzed with the goal of minimizing damage, with permission from local authorities in each location from which they came. Every sample is represented by stewards such as archaeologists or museum curators, who are

Authors or named in the Acknowledgments. Open science principles require making all data used to support the conclusions of a study maximally available, and we support these principles here by making fully publicly available not only the digital copies of molecules (the uploaded sequences) but also the molecular copies (the ancient DNA libraries themselves, which constitute molecular data storage). Those researchers who wish to carry out deeper sequencing of libraries published in this study should make a request to corresponding author D.R. We commit to granting reasonable requests as long as the libraries remain preserved in our laboratories, with no requirement that we be included as collaborators or co-authors on any resulting publications.

Acknowledgement

We thank the technical support and laboratory assistance of the following people from the Harvard Medical School and the Vienna University: Nicole Adamski, Nasreen Broomandkhoshbacht, Kim Callan, Elizabeth Curtis, Ann Marie Lawson, Megan Michel, Susann Nordenfelt, Jonas Oppenheimer, Lijun Qiu, Kristin Stewardson, Noah Workman, Fatma Zalzala, Kirsten Mandl, Kadir Toykan Ozdogan, Sarah Kellie Duffett Carlson, Beatriz Gamarra Rubio, Anna Wagner, Lea Demetz, Stefanie Hofer, and Guillermo Bravo. We thank Noémi Borbély her support in the uniparental data analyses and Balázs G. Mende his administrative support.

This study was funded by the MTA-ELTE Lendület “Momentum” program of the Hungarian Academy of Sciences. TH was supported by the Bolyai Scholarship of the Hungarian Academy of Sciences.

The ancient DNA data generation and analysis at Harvard was supported by the National Institutes of Health (R01-HG012287), the John Templeton Foundation (grant 61220), by a private gift from Jean-Francois Clin, by the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation, and by the Howard Hughes Medical Institute (DR).

Author contributions

Conceptualization: A.An., P.R., R.P., D.R., Zs.S.,
Data Collection: C.V., A.An., P.R., T.H., K.K., T.Sz., S.É., T.K., Zs.V., O.Ch., R.P., Zs.S.,
Data Curation: K.J., S.M., Zs.S.
Formal Analysis: A.Sz-N., H.R., S.M., A.Ak., Zs.S.
Investigation: A.Sz-N., K.J., N.R., O.Ch., Zs.S.
Writing – Original Draft: A.Sz-N., Zs.S.
Writing – Review & Editing: A.Sz-N., H.R., P.R., D.R., Zs.S.,
Resources: D.R., Zs.S.
Software: K.J., H.R., S.M., A.Ak.

Competing interests

The authors declare no competing interests.

References

1. Szécsényi-Nagy, A. *et al.* Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. *Proceedings of the Royal Society B* **282**, 20150339 (2015).
2. Keerl, V. A River Runs Through It- Ancient DNA Data on the Neolithic Populations of the Great Hungarian Plain. Doctoral thesis. (Johannes Gutenberg University Mainz, 2015).
3. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun* **5**, 5257 (2014).
4. Lipson, M. *et al.* Parallel paleogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
5. Mathieson, I. *et al.* The Genomic History Of Southeastern Europe. *Nature* **555**, 197–203 (2018).
6. Gelabert, P. *et al.* Social and genetic diversity in first farmers of central Europe. *Nature Human Behaviour 2024* 1–12 (2024) doi:10.1038/s41562-024-02034-z.
7. Lazaridis, I. *et al.* The genetic history of the Southern Arc: A bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
8. Novak, M. *et al.* Genome-wide analysis of nearly all the victims of a 6200 year old massacre. *PLoS One* **16**, e0247332- (2021).
9. Szécsényi-Nagy, A. *et al.* An unusual community in death – Reconsidering of the data on the mortuary practices of the Balaton-Lasinja culture in the light of bioarchaeological analyses. in *„Trans Lacum Pelsonem" Prähistorische Forschungen in Südwestungarn (5500–500 v. Chr.) Prehistoric Research in South-Western Hungary (5500–500 BC)* (ed. Eszter Bánffy und Judit P. Barna) 161–186 (Rahden/Westf., Budapest, Leipzig, Keszthely, Frankfurt/M., 2019).
10. Gerber, D. *et al.* Interdisciplinary Analyses of Bronze Age Communities from Western Hungary Reveal Complex Population Histories. *Mol Biol Evol* **40**, (2023).
11. Renfrew, C. The anatomy of innovation. in *Approaches to Social Archaeology* (ed. Renfrew, C.) 390–418 (Edinburgh University Press, Edinburgh, 1984).
12. Scharl, S. *Innovationstransfer in Prähistorischen Gesellschaften.* vol. 10 (Verlag Marie Leidorf GmbH, Rahden/Westf., 2019).
13. Bondár, M. Prehistoric innovations: Wheels and wheeled vehicles. *Acta Archaeologica Academiae Scientiarum Hungaricae* **69**, 271–298 (2018).
14. Hansen, S., Renn, J., Klimscha, F., Büttner -Barbara, J. & Kruse, H.-S. *The Digital Atlas of Innovations: A Research Program on Innovations in Prehistory and Antiquity*. vol. 6 http://journal.topoi.orghttp://creativecommons.org/licenses/by/3.0 (2016).
15. Renfrew, C. Varna and the social context of early metallurgy. *Antiquity* **52**, 199–203 (1978).
16. Renfrew, C. Varna and the emergence of wealth in prehistoric Europe. in *The Social Life of Things: commodities in cultural perspective* (ed. Appadurai, A.) 141–168 (Cambridge University Press, Cambridge, 1986).
17. Raczky, P. Settlements in South-East Europe. in *The Oxford Handbook of Neolithic Europe* (eds. Fowler, C., Harding, J. & Hofmann, D.) 235–253 (Oxford, 2015). doi:10.1093/oxfordhb/9780199545841.013.072.

18. Kalicz, N. Zusammenhänge zwischen dem Siedlungswesen und der Bevölkerungszahl während des Spätneolithikums in Ungarn – Connections between modes of settlement and population size during the Late Neolithic period in Hungary. in *Mensch und Umwelt während des Neolithikums und der Frühbronzezeit in Mitteleuropa. Ergebnisse interdisziplinärer Zusammenarbeit zwischen Archäologie, Klimatologie, Biologie und Medizin* (eds. Lippert, A., Schultz, M., Shennan, S. & Teschler- Nicola, M.) 153–163 (Rahden/Westf., 2001).
19. Osztás, A. *et al.* Coalescent community at Alsónyék: The timings and duration of Lengyel burials and settlement. *Bericht der Römisch-Germanischen Kommission* **94**, 179–182 (2013).
20. Raczky, P. A complex monument in the making at the Late Neolithic site of Polgár-Csőszhalom (Hungary). in *Accross the Mediterranean – Along the Nile. Studies in Egyptology, Nubiology and Late Antiquity Dedicated to László Török on the Occasion of his 75th Birthday* (eds. Bács, T. A., Bollók, Á. & Vida, T.) vol. 1 15–60 (Archaeolingua, Budapest, 2018).
21. Raczky, P., Anders, A., Sebők, K., Csippán, P. & Tóth, Z. The Times of Polgár-Csőszhalom. Chronologies of Human Activities in a Late Neolithic Settlement in Northeastern Hungary. in *Neolithic and Copper Age between the Carpathians and the Aegean Sea. Chronologies and Technologies from the 6th to the 4th Millenium BCE. International Workshop Budapest 2012* (eds. Hansen, S., Raczky, P., Anders, A. & Reingruber, A.) vol. 31 21–48 (Berlin, 2015).
22. Siklósi, Z. & Szilágyi, M. Culture, period or style? Reconsideration of Early and Middle Copper Age Chronology of the Great Hungarian Plain. *Radiocarbon* **63**, 585–646 (2021).
23. Gyucha, A. *Prehistoric Village Social Dynamics. The Early Copper Age in the Körös Region*. (Archaeolingua, Budapest, 2015).
24. Parkinson, W. A. *The Social Organisation of Early Copper Age Tribes on the Great Hungarian Plain*. vol. 1573 (BAR Publishing, Oxford, UK, 2006).
25. Siklósi, Z. *Traces of Social Inequality during the Late Neolithic in the Eastern Carpathian Basin*. vol. 3 (Eötvös Loránd University, Institute of Archaeological Sciences, Budapest, 2013).
26. Siklósi, Z. *et al.* Creating Histories: Different Perspectives, Controversial Narratives at Rákóczifalva, an Early Copper Age Site on the Great Hungarian Plain. *Eur J Archaeol* 1–22 (2022) doi:10.1017/eaa.2022.2.
27. Raczky, P., Anders, A. & Siklósi, Z. Trajectories of Continuity and Change between the Late Neolithic and the Copper Age in Eastern Hungary. in *The Neolithic and Eneolithic in Southeast Europe. New approaches to dating and cultural dynamics in the 6th to 4th Millenium BC* (eds. Schier, W. & Draşovean, F.) 319–346 (Verlag Marie Leidorf GmbH, Rahden/Westf., 2014).
28. Hegedűs, K. & Makkay, J. Vésztő-Mágor. A settlement of the Tisza culture. in *The Late Neolithic of the Tisza region* (eds. Tálas, L. & Raczky, P.) 85–103 (Directorate of the Szolnok County Museums, Budapest - Szolnok, 1987).
29. Draşovean, F. & Schier, W. *Uivar 'Gomilă'. A Prehistoric Tell Settlement in the Romanian Banat. Site, Architecture, Stratigraphy and Dating*. vol. 1 (Leidorf, Rahden/Westf., 2020).
30. Bognár-Kutzián, I. *The Early Copper Age Tiszapolgár Culture in the Carpathian Basin*. vol. 48 (Akadémiai kiadó, Budapest, 1972).

31. Patay, P. Die hochkupferzeitliche Bodrogkeresztur-kultur. *Bericht der Römisch-Germanischen Komission* **55**, 3–71 (1974).
32. Kalicz, N. The new results of the investigation on the Hungarian Copper Age. *Rassegna di Archeologia* **7**, 75–103 (1988).
33. Regenye, J. *et al.* Narratives for Lengyel funerary practice. *Bericht der Römisch-Germanischen Kommission* **97**, 5–80 (2016).
34. Raczky, P. A lengyeli kultúra legkésőbbi szakaszának leletei a Dunántúlon. Funde der spätesten Phase der Lengyel-Kultur in Westungarn. *Archeológiai Értesítő* **101**, 185–210 (1974).
35. Siklósi, Z., Prange, M., Kalicz, N. & Raczky, P. New Data on the Provenance of Early Copper Finds from the Great Hungarian Plain. in *Neolithic and Copper Age between the Carpathians and the Aegean Sea. Chronologies and Technologies from the 6th to the 4th Millennium BCE. International Workshop Budapest 2012* (eds. Hansen, S., Raczky, P., Anders, A. & Reingruber, A.) vol. 31 57–92 (Habelt Verlag, Bonn, 2015).
36. Chapman, J. *Tensions at Funerals*. (Archaeolingua, Budapest, 2000).
37. Gimbutas, M. A. *The Civilization of Goddess: The World of Old Europe*. (Harper, San Fransisco, 1991).
38. Bökönyi, S. *History of Domestic Mammals in Central and Eastern Europe*. (Akadémiai kiadó, Budapest, 1974).
39. Magyari, E. K. *et al.* Review on vegetation, landscape and climate changes in the Carpathian Basin during the Neolithic and Chalcolithic period. *Vegetation History and Archaeobotany* Preprint at https://doi.org/10.1007/s00334-024-00986-w (2024).
40. Parkinson, W. A. & Gyucha, A. A késő neolitikum-kora rézkor átmeneti időszakának társadalomszerkezeti változásai az Alföldön. Rekonstrukciós kísérlet. *Archaeologiai Értesítő* **132**, 37–81 (2007).
41. K. Zoffmann, Z. Prehistoric anthropological finds in the Carpathian Basin and the Penrose connections of the ethnic groups they represent. *Praehistoria* **6**, 103–129 (2005).
42. Giblin, J. I., Knudson, K. J., Bereczki, Z., Pálfi, G. & Pap, I. Strontium isotope analysis and human mobility during the Neolithic and Copper Age: A case study from the Great Hungarian Plain. *J Archaeol Sci* **40**, 227–239 (2013).
43. Mesterházy, G., Serlegi, G., Vágvölgyi, B., Füzesi, A. & Raczky, P. A szociális folyamatok színterei Polgár-Csőszhalom késő neolitikus településének összefüggéseiben. *Archaeologiai Értesítő* **144**, 1–32 (2019).
44. Rivollat, M. *et al.* Extensive pedigrees reveal the social organization of a Neolithic community. *Nature* **620**, 600–606 (2023).
45. Mittnik, A. *et al.* Kinship-based social inequality in Bronze Age Europe. *Science (1979)* **366**, 731–734 (2019).
46. Skourtanioti, E. *et al.* Ancient DNA reveals admixture history and endogamy in the prehistoric Aegean. *Nat Ecol Evol* (2023) doi:10.1038/s41559-022-01952-3.
47. Fowler, C. *et al.* A high-resolution picture of kinship practices in an Early Neolithic tomb. *Nature* **601**, 584–587 (2022).
48. Žegarac, A. *et al.* Ancient genomes provide insights into family structure and the heredity of social status in the early Bronze Age of southeastern Europe. *Sci Rep* **11**, 1–12 (2021).
49. Monroy Kuhn, J. M., Jakobsson, M. & Günther, T. Estimating genetic kin relationships in prehistoric populations. *PLoS One* **13**, e0195491 (2018).

50. Popli, D., Peyrégne, S. & Peter, B. M. KIN: a method to infer relatedness from low-coverage ancient DNA. *Genome Biol* **24**, 10 (2023).
51. Ringbauer, H. *et al.* Accurate detection of identity-by-descent segments in human ancient DNA. *Nat Genet* **56**, 143–151 (2024).
52. Lazaridis, I. *et al.* The genetic history of the Southern Arc: A bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
53. Mathieson, I. *et al.* The Genomic History Of Southeastern Europe. *Nature* **555**, 197–203 (2018).
54. Harney, É. *et al.* A minimally destructive protocol for DNA extraction from ancient teeth. *Genome Res* **31**, 472–483 (2021).
55. Jones, S. *The Archaeology of Ethnicity: Constructing Identities in Past and Present*. (Routledge, London, 1994).
56. Siklósi, Z. A régészeti kultúra fogalmának változása és az etnikai identitás azonosítása az ősrégészeti kutatásokban. *Korall-Társadalomtörténeti folyóirat* **24–25**, 73–88 (2006).
57. *Investigating Archaeological Cultures*. (Springer, New York, NY, 2011). doi:10.1017/CBO9781107415324.004.
58. *The Archaeology of Identity. Approaches to Gender, Age, Status, Ethnicity and Religion.* . (Routledge, London – New York, 2005).
59. Insoll, T. *The Archaeology of Identities. A Reader*. (Routledge, London – New York, 2007).
60. Patterson, N., Price, A. L. & Reich, D. Population Structure and Eigenanalysis. *PLoS Genet* **2**, e190 (2006).
61. Antonio, M. L. *et al.* Stable population structure in Europe since the Iron Age, despite high mobility. *Elife* **13**, e79714 (2024).
62. Maróti, Z. *et al.* The genetic origin of Huns, Avars, and conquering Hungarians. *Current Biology* (2022) doi:10.1016/j.cub.2022.04.093.
63. Penske, S. *et al.* Early contact between late farming and pastoralist societies in southeastern Europe. *Nature* **620**, 358–365 (2023).
64. Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nature Communications 2021 12:1* **12**, 1–11 (2021).
65. Childebayeva, A. *et al.* Population Genetics and Signatures of Selection in Early Neolithic European Farmers. *Mol Biol Evol* **39**, (2022).
66. Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nature Communications 2021 12:1* **12**, 1–11 (2021).
67. Mallick, S. *et al.* The Allen Ancient DNA Resource (AADR) a curated compendium of ancient human genomes. *Sci Data* **11**, 182 (2024).
68. Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* **9**, e98679 (2014).
69. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports 2019 9:1* **9**, 1–12 (2019).
70. Kilinç, G. M. *et al.* The Demographic Development of the First Farmers in Anatolia. *Current Biology* **26**, 2659–2666 (2016).
71. Feldman, M. *et al.* Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat Commun* **10**, 1218 (2019).

72. Lazaridis, I. *et al.* The Genetic Origin of the Indo-Europeans. *bioRxiv* 2024.04.17.589597 (2024) doi:10.1101/2024.04.17.589597.
73. Ecsedy, I. A New Item Relating the Connections with the East in the Hungarian Copper Age (A Marosdécse-type grave in Csongrád). *Móra Ferenc Múzeum Évkönyve* 9–17 (1971).
74. Mattila, T. M. *et al.* Genetic continuity, isolation, and gene flow in Stone Age Central and Eastern Europe. *Commun Biol* **6**, 793 (2023).
75. Ceballos, F. C. *et al.* Human inbreeding has decreased in time through the Holocene. *Current Biology* **31**, 3925-3934.e8 (2021).
76. Ariano, B. *et al.* Ancient Maltese genomes and the genetic geography of Neolithic Europe. *Current Biology* **32**, 2668-2680.e6 (2022).
77. Simões, L. G. *et al.* Genomic ancestry and social dynamics of the last hunter-gatherers of Atlantic France. *Proceedings of the National Academy of Sciences* **121**, 4940-4945. (2024).
78. Broushaki, F. *et al.* Early Neolithic genomes from the eastern Fertile Crescent. *Science (1979)* **353**, 499–503 (2016).
79. Cassidy, L. M. *et al.* A dynastic elite in monumental Neolithic society. *Nature* **582**, 384–388 (2020).
80. Ringbauer, H., Steinrücken, M., Fehren-Schmitz, L. & Reich, D. Increased rate of close-kin unions in the central Andes in the half millennium before European contact. *Current Biology* **30**, R980–R981 (2020).
81. Hamamy, H. *et al.* Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. *Genetics in Medicine* **13**, 841–847 (2011).
82. Jalkh, N. *et al.* Genome-wide inbreeding estimation within Lebanese communities using SNP arrays. *Eur J Hum Genet* **23**, 1364–1369 (2015).
83. Kang, J. T. L., Goldberg, A., Edge, M. D., Behar, D. M. & Rosenberg, N. A. Consanguinity Rates Predict Long Runs of Homozygosity in Jewish Populations. *Hum Hered* **82**, 87–102 (2016).
84. Windler, A., Thiele, R. & Müller, J. Increasing inequality in Chalcolithic Southeast Europe: The case of Durankulak. *J Archaeol Sci* **40**, 204–210 (2013).
85. Klimscha, F. *New Research on the "Kings of Metal": Systems of Social Distinction in the Copper Age of Southeastern Europe*. *Journal of Archaeological Research* (Springer US, 2019). doi:10.1007/s10814-019-09132-x.
86. Giblin, J. I. & Yerkes, R. W. Diet, dispersal and social differentiation during the Copper Age in eastern Hungary. *Antiquity* **90**, 81–94 (2016).
87. Gyucha, A., Yerkes, R. W. & Parkinson, W. A. Settlement Chronology and Layout. in *Bikeri. Two Copper Age Villages on the Great Hungarian Plain* (eds. Parkinson, W. A., Gyucha, A. & Yerkes, R. W.) (The Cotsen Institute of Archaeology Press, Los Angeles, 2021).
88. Kalicz, N. *Kőkori Falu Aszódon*. (Petőfi Múzeum, Aszód, 1985).
89. Kalicz, N. Die Bedeutung der schwarzen Gefäßbemalung der Lengyel-Kultur aus Aszód (Kom. Pest, Ungarn). *Analele Banatului* **14**, 135–157 (2006).
90. Kalicz, N. Aszód: ein gemischter Fundort der Lengyel- und Theiss-Kultur. *Communicationes Archaeologicae Hungariae* 5–54 (2008).
91. Kreiter, A., Kalicz, N., Kovács, K., Siklósi, Z. & Viktorik, O. Entangled traditions: Lengyel and Tisza ceramic technology in a Late Neolithic settlement in northern Hungary. *J Archaeol Sci Rep* **16**, 589–603 (2017).

92. Endrődi, A. Korabronzkori telepfeltárás Budapest-Albertfalván (Budapest, XI. ker., Hunyadi János út, Hrsz.: 43573) The excavation of an Early Bronze Age settlement at Budapest- Albertfalva (Budapest XI, Hunyadi János Road, Lrn.: 43573). *Aquincumi Füzetek* **8**, 81–91 (2002).
93. Reményi, L. & Endrődi, A. Az Albertfalva 6-os számú főút bővítését megelőző régészeti feltárás eredményei - The results of the excavations preceding tha enlargement of rute 6 by Albertfalva. (Budapest XI. ker., Hunyadi János út Hrsz.: 43573, 43575). *Aquincumi füzetek* **9**, 146–150 (2003).
94. Raczky, P., Faragó, N. & Füzesi, A. A special find assemblage from the Late Neolithic tell of Polgár-Csőszhalom: a contextual re-assessment. in *Walking among Ancient Tress* (eds. Grygiel, M. & Obst, P.) 317–348 (Fundacja Badań Archeologicznych Imienia Profesora Konrada Jażdżewskiego, Muzeum Archeologiczne i Etnograficzne w Łodzi, Łódź, 2022).
95. Anders, A. & Siklósi, Z. The times of their death – question of contemporaneity in burials from a Late Neolithic settlement in Polgár-Csőszhalom (NE-Hungary). in *Homines, Funera, Astra 3-4: The Multiple Faces of Death and Burial. Proceedings of the International Symposium on Funerary Anthropology, '1 Decembrie 1918' University (Alba Iulia, Romania)* (eds. Kogălniceanu, R., Gligor, M., Soficaru, A. & Stratton, S.) 35–46 (Archaeopress, Oxford, 2023).
96. Solnay, E. Early Copper Age Graves from Polgár-Nagy-Kasziba. *Dissertationes Archaeologicae* **6**, 179–216 (2018).
97. Szilágyi, M. Kora rézkori településszerkezet a Közép-Tisza-vidéken. (Eötvös Loránd University, 2015).
98. Raczky, P. & Siklósi, Z. Reconsideration of the Copper Age chronology of the eastern Carpathian Basin: A Bayesian approach. *Antiquity* **87**, 555–573 (2013).
99. Virag, C. Cercetări arheologice la Urziceni-Vamă [Archaeological researches at Urziceni-Vamă]. *Acta Musei Porolissensis* **26**, 41–76 (2004).
100. Virag, C. Urziceni, com. Urziceni, jud. Satu Mare Punct: Vamă. *Cronica cercetărilor arheologice din România 2017, Campania 2016* 150–151 (2017).
101. Virag, C. Urziceni, com. Urziceni, jud. Satu Mare Punct: Vamă. *Cronica cercetărilor arheologice din România 2006, Campania 2005* 383–386 (2006).
102. Szücs-Csillik, I. & Virag, C. The orientation of the dead at Urziceni necropolis. An archaeoastronomical approach. in *Archeo-Vest, Nr. IV: In Honorem Adrian Bejan, Interdisciplinaritate în Arheologie și Istorie, Timișoara, 26 novembrie 2016* 591–599 (JATEPress, Szeged, 2016).
103. Oanță-Marghitu, Rodica. Aurul și argintul antic al României : catalog de expoziție. 711 (2014).
104. Drașovean, Florin. *et al. Neolithic Art in Romania : [Published on the Occasion of the Exhibition 'Steinzeitkunst. Frühe Kulturen Aus Rumänien' Historisches Museum Olten, June 3 - October 5, 2008, the First Stage of an International Tour Ending in Bucharest]*. (Arte'm, 2008).
105. Reimer, P. J. *et al.* The IntCal20 Northern Hemishere Radiocarbon Curve (0-55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).
106. Bronk Ramsey, C. Bayesian Analysis of Radiocarbon Dates. *Radiocarbon* **51**, 337–360 (2009).
107. Pinhasi, R., Fernandes, D. M., Sirak, K. & Cheronet, O. Isolating the human cochlea to generate bone powder for ancient DNA analysis. *Nat Protoc* **14**, 1194–1205 (2019).

108. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A.* **110**, 15758–63 (2013).
109. Korlević, P. *et al.* Reducing Microbial and Human Contamination in DNA Extractions from Ancient Bones and Teeth. *Biotechniques* **59**, 87–93 (2015).
110. Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat Protoc* **13**, 2447–2461 (2018).
111. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil – DNA – glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* **370**, 20130624 (2015).
112. Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat Protoc* **15**, 2279–2300 (2020).
113. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
114. Rohland, N. *et al.* Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. *Genome Res* **32**, 2068–2078 (2022).
115. Behar, D. M. *et al.* A 'Copernican' reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet* **90**, 675–84 (2012).
116. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
117. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, 1–4 (2021).
118. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res* **44**, W58–W63 (2016).
119. van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* **30**, E386-94 (2009).
120. Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. *Current Biology* **23**, 553–559 (2013).
121. Okonechnikov, K., Golosova, O. & Fursov, M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).
122. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–2 (2009).
123. Leigh, J. W. & Bryant, D. popart: full-feature software for haplotype network construction. *Methods Ecol Evol* **6**, 1110–1116 (2015).
124. Mallick, S. *et al.* The Allen Ancient DNA Resource (AADR) a curated compendium of ancient human genomes. *Sci Data* **11**, 182 (2024).
125. Papac, L. *et al.* Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Sci Adv* **7**, eabi6941 (2021).
126. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput Sci Eng* **9**, 90–95 (2007).
127. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261–272 (2020).
128. Mahalanobis, P. C. On the Generalised Distance in Statistics. *Proceedings of the National Institute of Sciences of India* **2**, 49–55 (1936).
129. Posth, C. *et al.* Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* **615**, 117–126 (2023).

130. Ralf, A., Montiel González, D., Zhong, K. & Kayser, M. Yleaf: Software for Human Y-Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. *Mol Biol Evol* **35**, 1291–1294 (2018).
131. Childebayeva, A. *et al.* Population Genetics and Signatures of Selection in Early Neolithic European Farmers. *Mol Biol Evol* **39**, (2022).
132. Brunel, S. *et al.* Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proceedings of the National Academy of Sciences* **117**, 12791–12798 (2020).
133. Fernandes, D. M. *et al.* The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nat Ecol Evol* **4**, 334–345 (2020).
134. Freilich, S. *et al.* Reconstructing genetic histories and social organisation in Neolithic and Bronze Age Croatia. *Sci Rep* **11**, 16729 (2021).
135. Gelabert, P. *et al.* Genomes from Verteba cave suggest diversity within the Trypillians in Ukraine. *Sci Rep* **12**, 7242 (2022).
136. Fernandes, D. M. *et al.* A genomic Neolithic time transect of hunter-farmer admixture in central Poland. *Sci Rep* **8**, 14879 (2018).
137. Rivollat, M. *et al.* Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci Adv* **6**, eaaz5344 (2020).
138. Narasimhan, V. M. *et al.* The formation of human populations in South and Central Asia. *Science (1979)* **365**, (2019).
139. Novak, M. *et al.* Genome-wide analysis of nearly all the victims of a 6200 year old massacre. *PLoS One* **16**, e0247332- (2021).
140. Patterson, N. *et al.* Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* **601**, 588–594 (2022).
141. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
142. Sánchez-Quinto, F. *et al.* Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. *Proceedings of the National Academy of Sciences* **116**, 9469–9474 (2019).
143. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* **9**, 5233 (2019).
144. Waskom, M. L. seaborn: statistical data visualization. *J Open Source Softw* **6**, 3021 (2021).
145. Huang, Y., Carmi, S. & Ringbauer, H. Estimating effective population size trajectories from time-series Identity-by-Descent (IBD) segments. *bioRxiv* 2024.05.06.592728 (2024) doi:10.1101/2024.05.06.592728.
146. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).