Report

Current Biology

Fine-scale sampling uncovers the complexity of migrations in 5th–6th century Pannonia

Highlights

- Novel genomic data from 53 individuals from 5th–8th century Europe were generated
- The sites exhibit significant genetic diversity comparable to large swaths of Europe
- Associations between genomic ancestry and burial customs vary in the analyzed sites
- Spatial modeling analyses indicate gene flow from northern Europe into Pannonia

Authors

Deven N. Vyas, István Koncz, Alessandra Modi, ..., Tivadar Vida, Patrick J. Geary, Krishna R. Veeramah

Correspondence

david.caramelli@unifi.it (D.C.), vida.tivadar@btk.elte.hu (T.V.), geary@ias.edu (P.J.G.), krishna.veeramah@stonybrook.edu (K.R.V.)

In brief

Vyas, Koncz, et al. conduct a multidisciplinary analysis of 5th–6th century communities from the territory of former Roman Pannonia. Despite archaeological similarities, they find significant differences in genomic ancestry between the sites as well as evidence of gene flow from Northern Europe into the region starting in the late 5th century.

Vyas et al., 2023, Current Biology *33*, 3951–3961
September 25, 2023
https://doi.org/10.1016/j.cub.2023.07.063

CellPress

Current Biology

CellPress
OPEN ACCESS

Report

Fine-scale sampling uncovers the complexity of migrations in 5th–6th century Pannonia

Deven N. Vyas,[1,17] István Koncz,[2,17] Alessandra Modi,[3] Balázs Gusztáv Mende,[4] Yijie Tian,[1] Paolo Francalacci,[5] Martina Lari,[3] Stefania Vai,[3] Péter Straub,[6] Zsolt Gallina,[7] Tamás Szeniczey,[8] Tamás Hajdu,[8] Luisella Pejrani Baricco,[9] Caterina Giostra,[10] Rita Radzevičiūtė,[11] Zuzana Hofmanová,[11,12] Sándor Évinger,[13] Zsolt Bernert,[13] Walter Pohl,[14,15] David Caramelli,[3,*] Tivadar Vida,[2,*] Patrick J. Geary,[16,*] and Krishna R. Veeramah[1,18,*]

[1]Department of Ecology and Evolution, Stony Brook University, 650 Life Sciences Building, Stony Brook, NY 11794, USA
[2]Institute of Archaeological Sciences, ELTE - Eötvös Loránd University, Múzeum krt. 4/B, 1088 Budapest, Hungary
[3]Dipartimento di Biologia, Università degli Studi di Firenze, Via del Proconsolo 12, 50122 Firenze, Italy
[4]Institute of Archaeogenomics, Research Centre for the Humanities, Tóth Kálmán utca 4, 1097 Budapest, Hungary
[5]Dipartimento di Scienze della Vita e dell'Ambiente, Università di Cagliari, Via T. Fiorelli 1, 09126 Cagliari, Italy
[6]8200 Veszprém, Hungary
[7]Ásatárs Ltd, Futár utca 12, 6000 Kecskemét, Hungary
[8]Department of Biological Anthropology, ELTE - Eötvös Loránd University, Pázmány Péter sétány 1/c, 1117 Budapest, Hungary
[9]Soprintendenza Archeologia, Belle Arti e Paesaggio per la Città Metropolitana di Torino, piazza San Giovanni 2, 10122 Torino, Italy
[10]Dipartimento di Storia, Archeologia e Storia dell'Arte, Università Cattolica del Sacro Cuore, Largo A. Gemelli, 1, 20123 Milano, Italy
[11]Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
[12]Department of Archaeology and Museology, Faculty of Arts, Masaryk University, Arna Nováka 1/1, Brno 60200, Czech Republic
[13]Department of Anthropology, Hungarian Natural History Museum, Ludovika tér 2-6, 1083 Budapest, Hungary
[14]Institute for Medieval Research, Austrian Academy of Sciences, Dr-Ignaz-Seipel-Platz 2, 1020 Vienna, Austria
[15]Institute for Austrian Historical Research, University of Vienna, Universitätsring 1, 1010 Vienna, Austria
[16]School of Historical Studies, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA
[17]These authors contributed equally
[18]Lead contact
*Correspondence: david.caramelli@unifi.it (D.C.), vida.tivadar@btk.elte.hu (T.V.), geary@ias.edu (P.J.G.), krishna.veeramah@stonybrook.edu (K.R.V.)
https://doi.org/10.1016/j.cub.2023.07.063

SUMMARY

As the collapse of the Western Roman Empire accelerated during the 4th and 5th centuries, arriving "barbarian" groups began to establish new communities in the border provinces of the declining (and eventually former) empire. This was a time of significant cultural and political change throughout not only these border regions but Europe as a whole.[1,2] To better understand post-Roman community formation in one of these key frontier zones after the collapse of the Hunnic movement, we generated new paleogenomic data for a set of 38 burials from a time series of three 5th century cemeteries[3–5] at Lake Balaton, Hungary. We utilized a comprehensive sampling approach to characterize these cemeteries along with data from 38 additional burials from a previously published mid-6th century site[6] and analyzed them alongside data from over 550 penecontemporaneous individuals.[7–19] The range of genetic diversity in all four of these local burial communities is extensive and wider ranging than penecontemporaneous Europeans sequenced to date. Despite many commonalities in burial customs and demography, we find that there were substantial differences in genetic ancestry between the sites. We detect evidence of northern European gene flow into the Lake Balaton region. Additionally, we observe a statistically significant association between dress artifacts and genetic ancestry among 5th century genetically female burials. Our analysis shows that the formation of early Medieval communities was a multifarious process even at a local level, consisting of genetically heterogeneous groups.

RESULTS AND DISCUSSION

With the dissolution of the Western Roman Empire, the 5th century was a period of great political, cultural, and demographic changes in Europe.[1,2] This was particularly true in the Middle Danube Region, which had long served as a frontier zone of the Roman Empire and also became a border zone after its division into the Western and Eastern Empires after 395 CE. With the abandonment of the Pannonian provinces by the Roman civil and military administration, most likely in 433 CE, it had already lost its former political and military importance. The region's subsequent development was first determined by the period of Hunnic rule in the middle of the 5th century, after which it came under the influence of various "barbarian" groups

(Goths, Heruls, Langobards, etc.). While the political history of the region is well documented in written sources, these come from contemporary or later external authors writing in Greek or Latin (such as Marcellinus Comes, Priscus of Panium, Sidonius Apollinaris, Jordanes, Procopius of Caesarea, Menander Protector, etc.) from the Western or Eastern Roman Empire[1,20] who put much emphasis on the movement of various ethnic groups but did not describe how these impacted the life of communities. The archaeological record shows the transformation of settlement structures and patterns, the appearance of new archaeological phenomena, and the continuous emergence of new communities that founded small burial sites in the region, quite unlike the large late Roman cemeteries of the 4th century, which could sometimes contain thousands of graves.[1,2]

Amorim et al.[6] performed a comprehensive paleogenomic characterization of Szólád, a Langobard-period cemetery found on the southern shore of Lake Balaton, in present-day Hungary, dating to the middle of the 6th century. We demonstrated that the community was organized primarily around a large, three-generation male biological kindred (i.e., sets of closely biologically related individuals) enriched for genomic ancestry found in modern northern Europeans, and that it was possible to distinguish at least two groups of burials with different genomic ancestry that varied with regard to the design of the graves, the dress accessories, and other grave goods.

In this study we present the analysis of paleogenomic data derived from a dense spatial and temporal sampling approach[13,21,22] for 38 individuals from three cemeteries (Fonyód, Hács, and Balatonszemes) near Szólád that date to the middle and second half of the 5th century CE. They were chosen for analysis based on their geographical proximity to Szólád and are considered fairly typical of the region and this period. They chronologically represent a time transect of the second part of the 5th century, with Fonyód dated to the middle third, Hács to the second half, and Balatonszemes to the end of the 5th century (Figure 1). This allowed us to investigate whether the genetic heterogeneity observed at Szólád represented a long-standing post-Roman structure in the region or was the result of 6th century population movements as described in the written sources. In particular, we were interested in (1) to what extent the continuous emergence of these short-lived communities was the result of extensive human mobility in the region and the appearance of new population groups, as attested by the written sources, and (2) what role biological relatedness played in the construction of social kinship and organization of these post-Roman communities.[6,24] In other words, did genetic variation correspond to significant social and cultural differences in burial customs and spatial organization of the sites?

5th century Pannonia: Fonyód, Hács, and Balatonszemes

Fonyód, Hács, and Balatonszemes (STAR Methods) are located within 18 km of each other on elevated loess ridges close to the southern shore of Lake Balaton, Hungary. Archaeological dating of these sites is based on certain jewelry types (brooches, pins, earrings, etc.) and tools ("nomadic mirrors" of the Čmi-Brigetio type, double-sided combs).[3–5,25,26] All three sites likely represent short-lived rural communities. 19, 29, and 19 graves were excavated at the three sites, respectively; however, due to empty graves and taphonomic issues, only 14 graves had human remains at Fonyód, only 15 from Hács had remains (including Hacs_5, which is a mixture of two—a male and female—burials due to modern disturbances) that were not lost or destroyed and had adequate preservation, and only 13 had human remains at Balatonszemes.

Including information from some of the lost, destroyed, or badly preserved burials from Hács (and, when available, genetic sex in cases where adult osteological sex assessments are not present), there was a substantial bias toward adult female burials (i.e., 21 adult females, 5 adult males, 17 nonadults); this pattern remains at each site (Fonyód, 7 adult females, 2 adult males, 5 nonadults; Hács, 9, 2, 7; Balatonszemes, 7, 2, 5). Hács and Balatonszemes are also very similar in terms of burial customs, as artifacts are mostly found in female burials, in contrast to male burials, which are generally poorly furnished.[4,5,26] Artificial cranial deformation (ACD) is also observable in all sites, but it is only prevalent in Fonyód (found in 7/11 preserved crania), while the other two sites only contain single examples (STAR Methods). The sites also show differences in terms of spatial organization of the burials, with Fonyód being unique[26] due to its six grave clusters lying at roughly equal distances of 50–60 m from each other (Figure 2).

5th–8th century Italy: Bardonecchia and Torino Lavazza

The two penecontemporary Italian sites of Bardonecchia and Torino Lavazza were included in this study, so as to increase the number of available high-quality Medieval Italian sequenced individuals. These data supplement the often lower-coverage Italian data from Rome from Antonio et al.[7] (Table S1).

Genome sequencing

From Fonyód, we obtained genomic data from 13 of 14 graves with human remains, from Hács from 14/15 preserved burials (the 15th, Hacs_23, was not accessible for sampling), and from Balatonszemes 11 of 13 graves with human remains. Altogether 38 samples were newly sequenced and analyzed (STAR Methods). Ten samples were used in whole-genome sequencing (WGS) (coverage mean 8.36×, range 5–11×) and the remaining 28 in genome-wide SNP capture sequencing for ~1.2 million SNPs (1240K)[28–30] (coverage mean 1.47×, range 0.02–3.46×) (Table S1). Among WGS samples, virtually all of the 1,126,140 autosomal 1240K SNPs were covered (1,098,348 to 1,120,721 SNPs), while for captures the number of SNPs covered ranged from 24,558 to 859,675 (Table S1). Mitochondrial contamination rate medians[31] were 1%–3% for all individuals, while nuclear contamination point estimates[32] from the 14 males were between 0.4% and 2%. In addition, 14 new samples from Bardonecchia and Torino Lavazza (5th–8th century Italy) underwent 1240K capture with coverages from 0.39 to 3.85× (covering 306,226 to 845,032 SNPs), while one from Bardonecchia underwent WGS (Bard_T1, coverage 6.54×, 1,121,223 SNPs).

Temporal variation in population genetic structure in the 5th century

We conducted a principal component analysis (PCA) of 561 ancient individuals (with coverage $\geq$ 0.1×); ancient individuals were transformed onto a background of modern POPRES[33] reference individuals using a Procrustes-transformation technique[6,34]

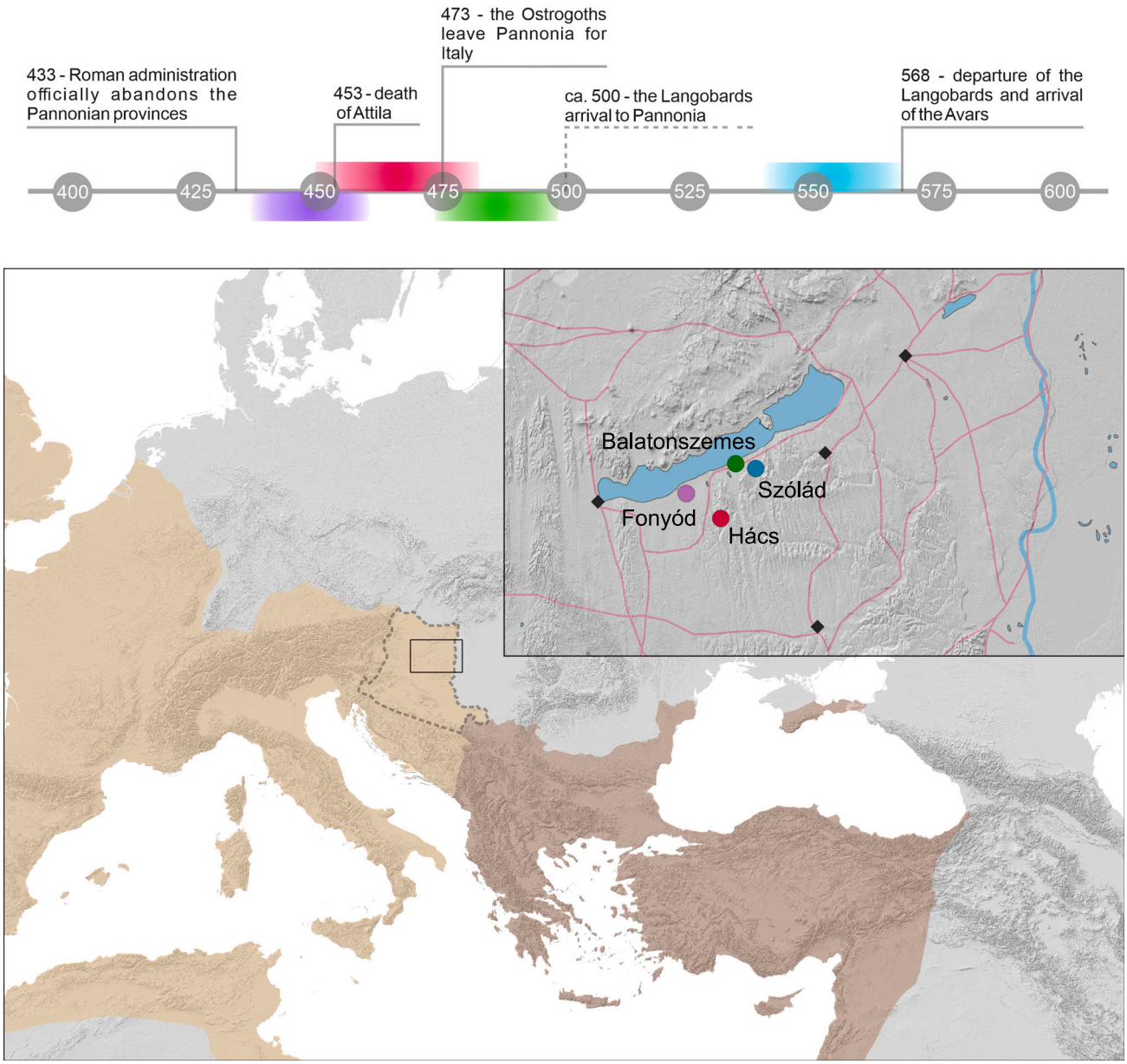

Figure 1. Location and chronological position of the four investigated sites within the (Western) Roman Empire

Roman roads are indicated with red lines, while late Roman inner fortresses of the Lake Balaton area are marked with black. The Pannonian provinces (Prima, Secunda, Valeria, and Savia) are outlined based on their combined borders c. 400 CE. Map created with QGiS v3.22.1 and resources used from Ancient World Mapping Center. "Coastline, River, Inland water, Open water, Roman road." http://awmc.unc.edu/wordpress/map-files/.[23]

using smartPCA.[35,36] We found that all 5th/6th century Lake Balaton individuals (n = 69 due to the coverage cut-off) reflected primarily modern European genetic diversity (Figures 3A and S1). In addition, we analyzed a penecontemporary (c. 4th–8th century CE) set of 492 comparative individuals (477 previously published alongside 15 newly produced from Bardonecchia and Torino Lavazza) sampled from across Europe (Table S1; STAR Methods).[6–12,37] Confidence ellipses of PC1 and PC2 coordinates demonstrate that while the penecontemporary regions cluster primarily with modern individuals from the same geographic regions (i.e., they are localized) (Figure 3B), our 5th–6th century Lake Balaton individuals have a much more diverse genomic ancestry, encompassing a range reflecting the entire north to south axis of modern European genetic variation in the POPRES dataset[33] (though not western or eastern Europe) (Figure 3A). While all Lake Balaton individuals were sampled from a single region of 200 km^2 over ~100 years, the penecontemporary populations generally encompass much larger geographic regions and wider time frames (and thus we would expect more genetic diversity): this suggests that this part of central Europe experienced particularly high rates of gene flow during the 5th–6th centuries. We explicitly modeled gene flow in a spatial context across Europe using our Lake

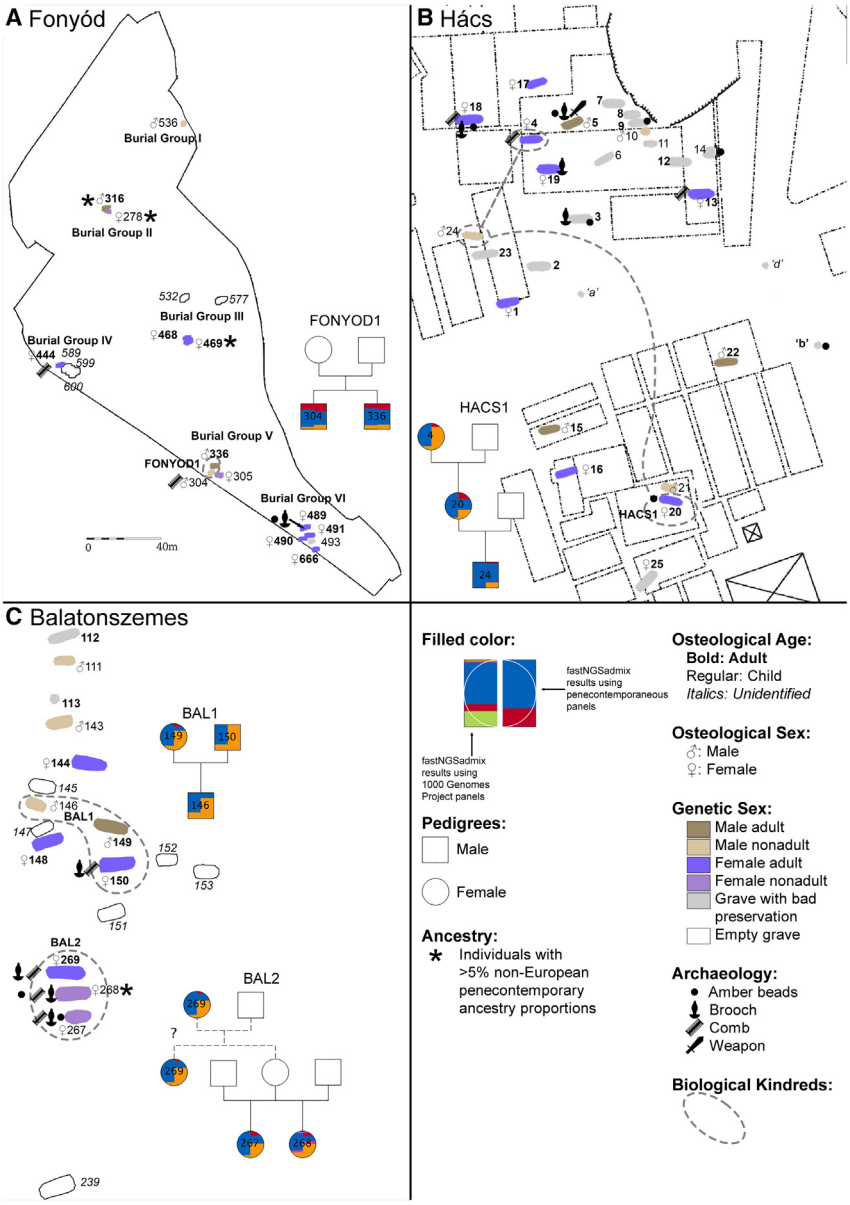

Figure 2. Cemetery maps of Fonyód, Hács, and Balatonszemes

Burials are identified based on age, sex (osteological and genetic), and sampling status as well as the presence of amber beads, brooches, double- and single-sided combs, and weapons. Dotted outlines indicate biological relatedness. Pedigrees of the biologically related individuals are overlaid on the cemetery maps with colors corresponding to fastNGSadmix results from Figure 4.

(A) Cemetery map of Fonyód. Due to the dimensions of the cemetery, graves sizes are not to scale with the size of the cemetery. This figure was modified from Gallina and Straub.[26]

(B) Cemetery map of Hács. This figure was modified from Kiss.[5]

(C) Cemetery map of Balatonszemes. Bal_26 is not included in this figure, as this individual was buried 200 m west of the main site. This figure was modified from Miháczi-Pálfi.[27]

See also Tables S1, S3, and S4.

Balaton and penecontemporary individuals using FEEMS[38] and indeed found a strong rate of gene flow between the Lake Balaton region and populations across Northern Europe ~1,000 km away (Figures 3D and S2).

To further explore the structure within our Lake Balaton cemeteries, we also used a model-based clustering method (fastNGSadmix)[39] to characterize the genomic ancestry of each individual using 1000 Genomes Project (1000G)[40] populations as possible sources. In addition, given the localized geographic structure observed in the PCA for the penecontemporary Europeans, we constructed a panel of ancient c. 4th–8th century penecontemporary reference individuals with a similar geographic distribution to the 1000G populations to form eight ancient panels (EASIA, MEDEU, NAFRICA, NGBI, SASIA, SCAND, and SUBSAHARAN) (STAR Methods).[7,9–12,14–19] The modern and ancient reference sets yielded similar patterns (Figures 4A, 4B, and S3), with almost all individuals possessing some combination of primarily Tuscan (TSI), Central European and Great Britain (CEU+GBR), and Finnish (FIN) modern ancestry or Mediterranean (MEDEU), northern Germany and British (NGBI), and Scandinavian/Estonian (SCAND) penecontemporary ancestry. We conducted linear regressions comparing ancestry proportions of MEDEU versus TSI, NGBI versus CEU+GBR, and SCAND versus FIN (for 441 individuals not in a reference panel) and found R^2 values of 0.79, 0.54, and 0.36, respectively (all comparisons $p < 0.001$). However, all three 5th century cemeteries have distinct profiles with regard to the relative proportions of these components, with the penecontemporary analysis showing these differences most clearly. Most notably, there is a clear progressive significant increase in SCAND ancestry and corresponding decrease in NGBI ancestry over time for the 5th century sites (Fonyód, Hács, Balatonszemes; SCAND 0.07 < 0.26 < 0.38; NGBI 0.40 > 0.31 > 0.14; Figure 4C).

Fonyód, which dates to the peak of the Hunnic power in the region, differs markedly from the later two 5th century sites from both an archaeological and genomic perspective (STAR Methods). Its unique spatial structure suggests that it is a burial site of a short-lived coexistence of Hunnic period groups, part of them practicing ACD, a tradition that is considered a foreign phenomenon in Pannonia and less prevalent in the later sites.[41] Genomically, Fonyód also has significantly (as assessed by non-overlapping 95% CIs) more Mediterranean ancestry than the other Lake Balaton sites and is somewhat southern shifted on the PCA (Figures 3 and 4). Overall, the site shows more non-European diversity, with two individuals possessing ~6% South and/or East Asian ancestry (Fonyod_278 and Fonyod_316) and

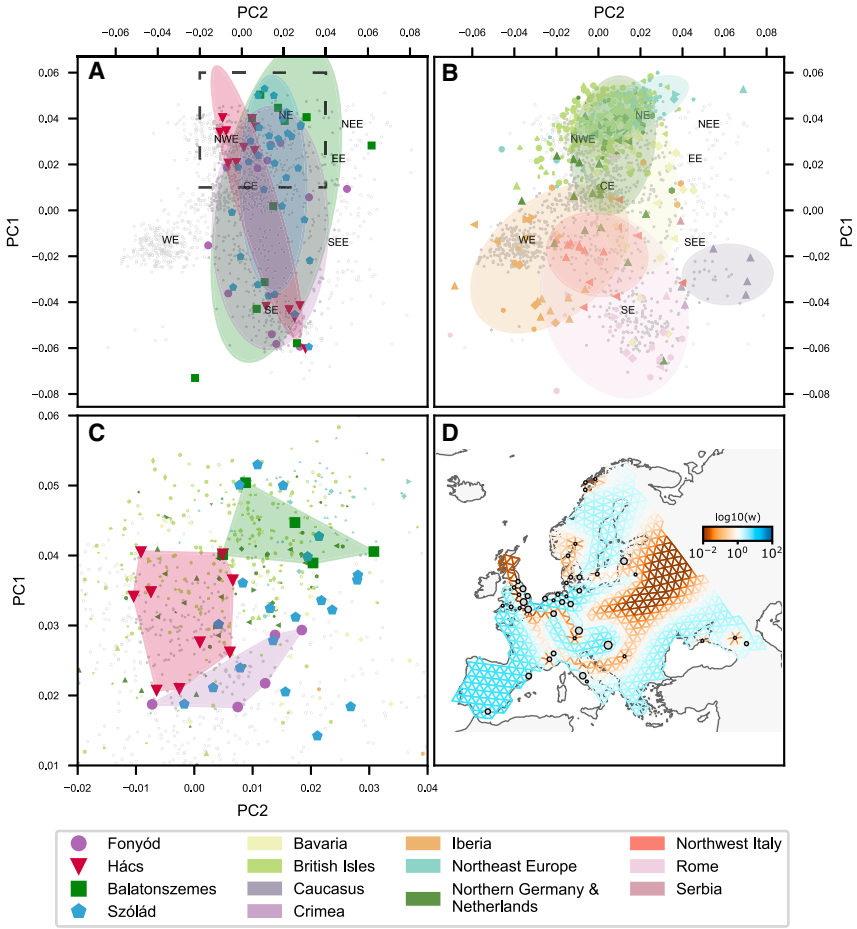

Figure 3. Principal component analysis and spatial gene flow analysis

Procrustes PCA of 561 individuals[6–12,37] transformed onto a PCA with 1,385 modern European individuals from the POPRES dataset[33] using pseudohaploid genotype calls from 328,670 SNPs with individuals. Penecontemporaneous reference individuals are colored in pastel colors based on their region, while modern individuals from POPRES are colored in gray.

(A) PCA results from the Lake Balaton communities are plotted with covariance confidence ellipses with radii corresponding to 1.5 standard deviations overlaying the PCA. The dashed box identifies the northern clusters from the Lake Balaton communities.

(B) PCA results from the penecontemporary populations are plotted with covariance confidence ellipses with radii corresponding to 1.5 standard deviations.

(C) A zoom-in of the area within the dashed box from (A). This plot focuses on the northern cluster of the PCA with polygons identifying the lack of overlap between the 5th century Lake Balaton communities.

(D) A FEEMS plot estimating gene flow between the 4th and 8th century communities analyzed in the PCA with all four Lake Balaton communities analyzed together. Bluer colors indicate more migration, while brown indicates less/no migration.

See also Figures S1 and S2 and Table S1.

one individual 12% African ancestry (Fonyod_469) (Figure 4). Thus, the genomic profile of Fonyód may reflect heterogeneity of the late Roman period and/or recent influx from the East.

Both Hács and Balatonszemes are enriched for individuals with very high proportions of northern European ancestry (NGBI+SCAND) that are not as prominent at Fonyód (Figure 3C), potentially indicating the arrival of new groups to the region consistent with written records describing the emergence of new "barbarian" powers after the fall of the Hunnic empire, in the second half of the 5th century. The high level of gene flow between Lake Balaton individuals and northern Europe is only seen in our FEEMS analysis when examining these later two communities, while the use of the earlier Fonyód site results in a barrier to gene flow to this region and greater gene flow with southern Europe (Figure S2). However, because of variation in the relative SCAND/NGBI components, individuals with higher PC1 values (i.e., more northern European ancestry) show no overlap in the PCA between the two sites (Figure 3C), with Hács angling toward the modern northwestern Europe and Balatonszemes northeastern Europe. These differences could indicate that groups from similar, yet distinguishable, sources from northern Europe arrived in the region in multiple waves during the second half of the 5th century. On the other hand, the presence of individuals with high amounts of MEDEU ancestry is a constant in all four sites, and they show overlap in the PCA, suggesting this may represent a more stable local genomic signature during this entire period. To test our findings of northern gene flow into the Balaton region subsequent to Fonyód, we also conducted qpAdm[29] analyses; these findings are consistent with our described findings from FEEMS and fastNGSadmix, though we caution that the qpAdm method is not ideal for analyzing closely related Late Antique/Early Medieval source populations (STAR Methods; Figure S4; Table S2).

Interestingly, Szólád demonstrates a profile that encompasses genomic variation observed in all three 5th century sites, albeit it has a much larger sample size. Northern European ancestry is even more prominent than at Hács and Balatonszemes, driven primarily (but not totally) by a large nine-member pedigree identified previously[6] (Figure 4). The individuals in this pedigree along with most other northern-like individuals lack the high SCAND component observed in Balatonszemes (qualitatively resembling those from Fonyód and Hács instead; the three notable exceptions being 2 second-degree relatives, Sz_41 and Sz_42, with 92% and 73% SCAND ancestry, respectively, and Sz_4 with 59%), suggesting no major direct continuity between these two sites. Szólád is notable for strontium isotope data, suggesting most adults at the site were non-local, regardless of genomic ancestry.[6] However our data make clear that the major patterns of genetic ancestry observed at Szólád were already established during the second half of the 5th century in and around the Lake Balaton area. Thus, this community could have formed from the existing diverse regional pool of genetic variation established ~50 years earlier, rather than just being the result of the arrival of a new population group, i.e., the Langobards, as interpreted by both historical and archaeological research.[42–45]

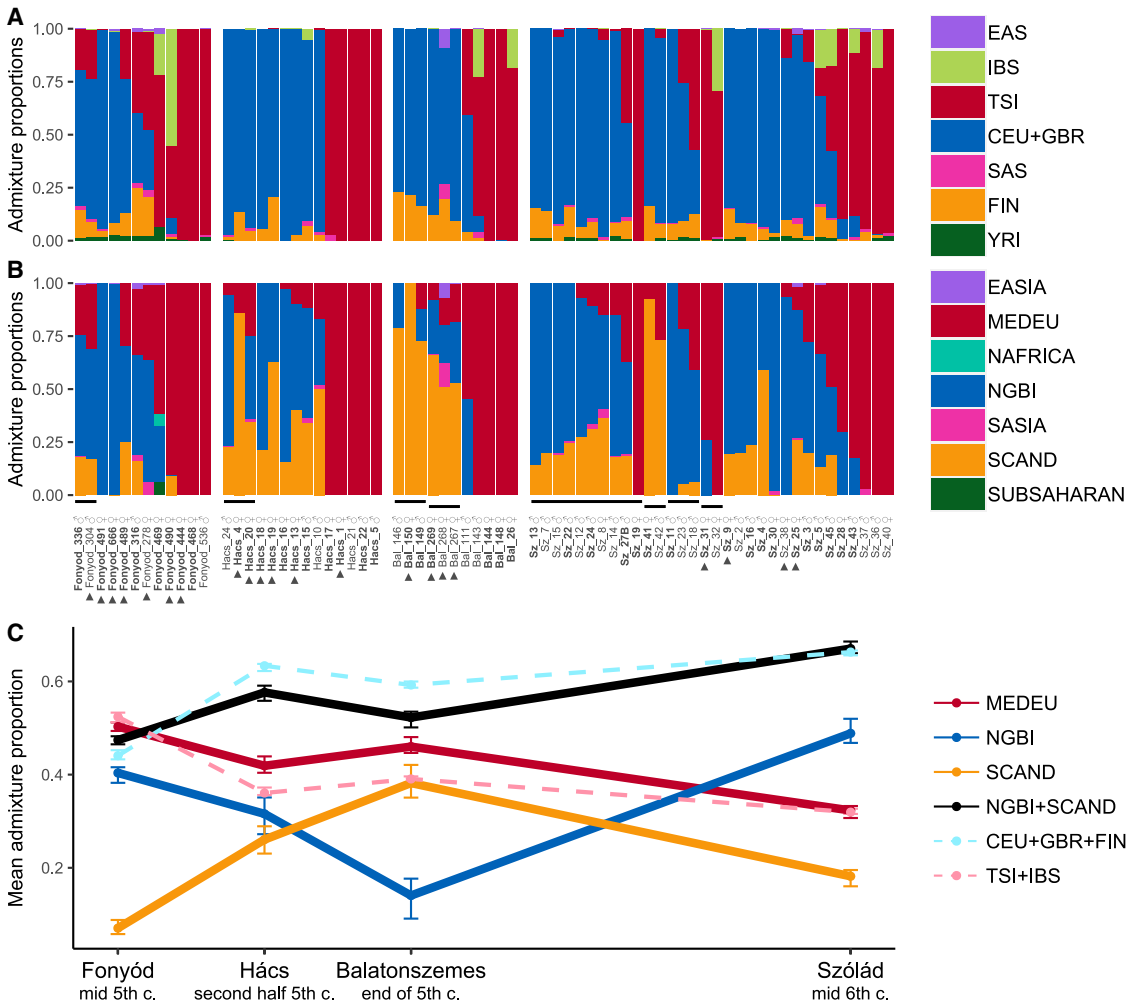

Figure 4. Supervised ancestry proportions from individuals from Fonyód, Hács, Balatonszemes, and Szólád

Individuals with <0.1× coverage were excluded from these analyses. The ♂ and ♀ symbols identify genetic males and genetic females, respectively, while triangles identify individuals buried with dress accessories (in the case of Fonyód, Hács, and Balatonszemes) or brooches specifically (in the case of Szólád).

(A) Proportions were estimated using fastNGSadmix and using 1000 Genomes populations as references[40] (CEU+GBR, Northern Europeans from Utah [CEU] and British in England and Scotland [GBR]; FIN, Finnish in Finland; IBS, Iberian populations in Spain; TSI, Tuscans from Italy; EAS, the East Asian super-population; SAS, the South Asian super-population; YRI, Yoruba in Ibadan, Nigeria).

(B) Proportions were estimated using penecontemporaneous individuals to form reference populations[7,9–12,14–19] (MEDEU, Italy and Iberia [Mediterranean Europe]; NGBI, what are now northern Germany and Britain; SCAND, what are now Scandinavia/Estonia; EASIA, what is now Hanben, Taiwan; NAFRICA, what is now Sudan; SASIA, Roopkund Lake in what is now India; SUBSAHARAN, sites from sub-Saharan Africa). Individuals are sorted based on increasing MEDEU and then by decreasing NGBI.

(C) A line graph showing the change in ancestry proportions through time over the four sites. Error bars correspond to 95% confidence intervals calculated based on bootstrap analyses.

See also Figure S3 and Table S1.

Biological relatedness and social relationships

The small burial groups at Fonyód were previously interpreted as signs of strong social ties,[26] and Hács and Balatonszemes have been described as family cemeteries.[5,25,46] We used lcMLkin[47] to identify close biological relatives in all three 5th century sites. At all three sites, biological kindreds consisted of only a few very close first- and second-degree relatives, mostly on the mother's side (Figure 2; Table S3). These analyses were also validated by a READ analysis that found the same kinship relationships (STAR Methods).[48] This is in stark contrast to Szólád, where the cemetery was organized largely around male biological relatives with a large extended pedigree.[6] Biologically related individuals at Fonyód and Balatonszemes were buried in close proximity to each other, and in the latter case their connection is also clearly reflected by the similarities in the burial customs, suggesting that these connections held meaningful social values (STAR Methods). However, the relatively low number of biologically related individuals, the small size of the kindreds, and the lack of biological relations in most burial groups—probably also influenced by the short occupation period of the sites and the low number of buried individuals—suggest that other factors may have had a major influence on the formation of these 5th century communities.

To examine to what extent the broader biological backgrounds of individuals (as determined by genomic ancestry) beyond close biological kinship were acknowledged as meaningful social ties, we also compared burial customs demonstrating variation with the POPRES PCA PC1 and PC2 coordinates using a logistic regression framework (Table S4). We observed a strong significant association between the presence of various jewelry types and dress accessories (i.e., polyhedral earrings, various brooch types, bracelets, amber beads, etc.) characteristic of the 5th century female burials[49,50] and genetic variation ($p = 0.008$, $n = 20$ for adults; $p = 0.002$, $n = 23$ for adults and nonadults). This appears to be primarily driven by most individuals with dress accessories possessing high amounts of northern European ancestry (13/16 are majority SCAND/NGBI) (Figures 4 and S1), and as such, site-specific regressions are only significant for Hács ($p = 0.01$, $n = 8$) and Balatonszemes ($p = 0.008$, $n = 7$) but not Fonyód ($p = 0.08$, $n = 8$) (despite all three tests having similar power; Table S4). These results point to individuals with northern European genetic backgrounds being treated differently in death by their respective communities, perhaps marking cultural and/or social differences between locals and more recent migrants in the region. We also found a strong association between the PCA results and ACD for the individuals at Fonyód (with preserved crania) ($p = 0.001$, $n = 10$) (STAR Methods; Table S4). Though close biological relatives were rare, it is also noteworthy that all biological kindreds involved individuals with predominantly northern genomic ancestry, maybe reflecting another aspect of this biologically structured social organization.

While Szólád shows a similar genomic profile to the three 5th century sites, it is strikingly different in terms of funerary practices, spatial organization, and demographics. There was no significant association between genetic variation and brooches—an artifact that in various forms and types is present in the 5th century sites as well as the 6th century Szólád—from genetically female burials (including and excluding nonadults) ($p < 0.56$, $n = 11$), suggesting that social and economic differences were no longer formulated along fault lines of genetic background in this community. Whether this weakening of the association between this artifact and genomic ancestry reflects a more general shift in social kinship practices (at least with respect to women) since the 5th century and the establishment of Langobard rule is difficult to conclude with just our data, as even among other 6th century Lake Balaton sites Szólád is considered somewhat unique, with the vast majority of adult and adolescent males (15/19) being buried with weapons.[42,51] The importance of a northern genomic background appears to still be manifested in this community socially, but rather through the dominance of the cemetery by a spatially clustered and primarily male kindred.

Conclusion

A common assumption of ethnoarchaeology posited a common ancestral, ethnic, and cultural heritage of Early Medieval migratory communities following the dissolution of the Roman Empire in the West.[52,53] Recent historical, archaeological, and anthropological studies have found this to be a vast oversimplification, with the material culture demonstrating significant complexity both on site and on the regional level. This study adds striking evidence of this complexity. Rather than homogeneity, our three post-Roman 5th century sites from Lake Balaton exhibit considerable genomic diversity compared to penecontemporaneous Europe. This region experienced particularly high levels of gene flow during this period, and there are significant shifts in genomic ancestry through our 5th century time transect over a period of only ~50 years that indicate migration into the region from various sources, probably from areas in northern Europe, consistent with the continuously changing political landscape of the period described in the historical record.[54,55] As these post-Roman communities formed from mixtures of locals and migrants, social ties between individuals of similar genomic backgrounds appear to have still been maintained to some extent, but these links could shift rapidly rather than being enduring states of social organization, and the importance of close biological relatedness varied considerably. Given the immense complexity observed in just these four small cemeteries, it is clear that comprehensive fine-grained spatiotemporal genomic sampling will be critical to unpack the processes that underlie the subsequent development of modern Europe.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Archaeological material and ethics permission
  - Description of archaeological sites
  - Artificial cranial deformation
- METHOD DETAILS
  - Ancient DNA lab work and sequence processing
  - Hács, Balatonszemes, Bardonecchia, and Torino Lavazza
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Processing of novel genomic dataset
  - Mitochondrial and Y chromosome analyses
  - Comparative dataset
  - Biological relatedness assessment
  - Principal component analysis
  - Gene flow modeling analysis
  - Model-based clustering analyses
  - Analyses using qpAdm

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2023.07.063.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 856453 ERC-2019-SyG), the Anneliese Maier Research Award of the Alexander von Humboldt Foundation, the Max Planck Society, the German Federal Ministry for Education and Research, the Swedish Riksbankens Jubileumfond, the Czech Grant Agency (GACR 21-17092X),

the Gerard B. Lambert Foundation, the Institute for Advanced Study Director's Office, and the Italian Ministry of Education, University and Research (project "Dipartimenti di Eccellenza 2018–2022" and PRIN2017 grant no. 20177PJ9XF). We thank Szilvia Honti, Péter Németh, and the Rippl-Rónai Museum, Kaposvár, for providing access to the material from Balatonszemes; the Hungarian Natural History Museum, Budapest, for providing access to the anthropological material from Fonyód and Hács; and the late Elena Bedini for access to the material from Bardonecchia and Torino Lavazza. We also thank Johannes Krause for funding and performing all SNP capture and sequencing.

AUTHOR CONTRIBUTIONS

Conceptualization, P.J.G., K.R.V., D.N.V., and I.K.; DNA preparation and sequencing, A.M., R.R., and S.V.; formal analysis, D.N.V. and I.K.; investigation, D.N.V., I.K., B.G.M., T.S., and T.H.; resources, B.G.M., I.K., Y.T., M.L., P.S., Z.G., T.S., T.H., L.P.B., C.G., S.É., and Z.B.; visualization, D.N.V., I.K., and Y.T.; writing, D.N.V., K.R.V., and I.K.; supervision, W.P., Z.H., D.C., T.V., P.J.G., and K.R.V.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 6, 2023
Revised: June 20, 2023
Accepted: July 31, 2023
Published: August 25, 2023

REFERENCES

1. Hardt, M. (2011). Pannonien im Spannungsfeld zwischen Römer- und Völkerwanderungszeit – eine geschichtliche Einführung. In Keszthely-Fenékpuszta im Kontext spätantiker Kontinuitätsforschung zwischen Noricum und Moesia Castellum Pannonicum Pelsonense 2, O. Heinrich-Tamáska, ed. (Leidorf), pp. 15–28.
2. Vida, T. (2011). Die Zeit zwischen dem 4. und dem 6. Jahrhundert im mittleren Donauraum aus archäologischer Sicht. In Römische Legionslager in den Rhein- und Donauprovinzen Abhandlungen, M. Konrad, ed. (Bayerische Akademie der Wissenschaften, phil.-hist. Klasse / NF 138), pp. 615–650.
3. Straub, P. (2008). Adalékok a Balaton környéki 5. századi temetők Felső-Duna vidéki kapcsolatához. Zalai Múzeum *17*, 189–207.
4. Miháczi-Pálfi, A. (2017). Form-und herstellungstechnische Analyse der Bügelfibeln von Balatonszemes aus dem dritten Viertel des 5. Jahrhunderts*35-36* (ANTÆUS: Communicationes Ex Instituto Archaeologico Academiae Scientiarum Hungaricae), pp. 67–74.
5. Kiss, A. (1995). Das germanische Gräberfeld von Hács-Béndekpuszta (Westungarn) aus dem 5.–6. Jahrhundert. Acta Antiquae Academiae Scientiarum Hungaricae *36*, 275–342.
6. Amorim, C.E.G., Vai, S., Posth, C., Modi, A., Koncz, I., Hakenbeck, S., La Rocca, M.C., Mende, B., Bobo, D., Pohl, W., et al. (2018). Understanding 6th-century barbarian social organization and migration through paleogenomics. Nat. Commun. *9*, 3547.
7. Antonio, M.L., Gao, Z., Moots, H.M., Lucci, M., Candilio, F., Sawyer, S., Oberreiter, V., Calderon, D., Devitofranceschi, K., Aikens, R.C., et al. (2019). Ancient Rome: a genetic crossroads of Europe and the Mediterranean. Science *366*, 708–714.
8. Damgaard, P.d.B., Marchi, N., Rasmussen, S., Peyrot, M., Renaud, G., Korneliussen, T., Moreno-Mayar, J.V., Pedersen, M.W., Goldberg, A., Usmanova, E., et al. (2018). 137 ancient human genomes from across the Eurasian steppes. Nature *557*, 369–374.
9. Gretzinger, J., Sayer, D., Justeau, P., Altena, E., Pala, M., Dulias, K., Edwards, C.J., Jodoin, S., Lacher, L., Sabin, S., et al. (2022). The Anglo-Saxon migration and the formation of the early English gene pool. Nature *610*, 112–119.
10. Margaryan, A., Lawson, D.J., Sikora, M., Racimo, F., Rasmussen, S., Moltke, I., Cassidy, L.M., Jørsboe, E., Ingason, A., Pedersen, M.W., et al. (2020). Population genomics of the Viking world. Nature *585*, 390–396.
11. Olalde, I., Mallick, S., Patterson, N., Rohland, N., Villalba-Mouco, V., Silva, M., Dulias, K., Edwards, C.J., Gandini, F., Pala, M., et al. (2019). The genomic history of the Iberian Peninsula over the past 8000 years. Science *363*, 1230–1234.
12. Schiffels, S., Haak, W., Paajanen, P., Llamas, B., Popescu, E., Loe, L., Clarke, R., Lyons, A., Mortimer, R., Sayer, D., et al. (2016). Iron Age and Anglo-Saxon genomes from East England reveal British migration history. Nat. Commun. *7*, 10408–10409.
13. Veeramah, K.R. (2018). The importance of fine-scale studies for integrating paleogenomics and archaeology. Curr. Opin. Genet. Dev. *53*, 83–89.
14. Harney, É., Nayak, A., Patterson, N., Joglekar, P., Mushrif-Tripathy, V., Mallick, S., Rohland, N., Sedig, J., Adamski, N., Bernardos, R., et al. (2019). Ancient DNA from the skeletons of Roopkund Lake reveals Mediterranean migrants in India. Nat. Commun. *10*, 3670.
15. Prendergast, M.E., Lipson, M., Sawchuk, E.A., Olalde, I., Ogola, C.A., Rohland, N., Sirak, K.A., Adamski, N., Bernardos, R., Broomandkhoshbacht, N., et al. (2019). Ancient DNA reveals a multistep spread of the first herders into sub-Saharan Africa. Science *365*, eaaw6275.
16. Sirak, K.A., Fernandes, D.M., Lipson, M., Mallick, S., Mah, M., Olalde, I., Ringbauer, H., Rohland, N., Hadden, C.S., Harney, É., et al. (2021). Social stratification without genetic differentiation at the site of Kulubnarti in Christian Period Nubia. Nat. Commun. *12*, 7283.
17. Skoglund, P., Thompson, J.C., Prendergast, M.E., Mittnik, A., Sirak, K., Hajdinjak, M., Salie, T., Rohland, N., Mallick, S., Peltzer, A., et al. (2017). Reconstructing prehistoric African population structure. Cell *171*, 59–71.e21.
18. Wang, K., Goldstein, S., Bleasdale, M., Clist, B., Bostoen, K., Bakwa-Lufu, P., Buck, L.T., Crowther, A., Dème, A., McIntosh, R.J., et al. (2020). Ancient genomes reveal complex patterns of population movement, interaction, and replacement in sub-Saharan Africa. Sci. Adv. *6*, eaaz0183.
19. Wang, C.-C., Yeh, H.-Y., Popov, A.N., Zhang, H.-Q., Matsumura, H., Sirak, K., Cheronet, O., Kovalev, A., Rohland, N., Kim, A.M., et al. (2021). Genomic insights into the formation of human populations in East Asia. Nature *591*, 413–419.
20. Bratož, R. (2003). Pannonien. In Reallexikon der germanischen Altertumskunde, H. Jankuhn, ed. (De Gruyter), pp. 469–483.
21. Pohl, W., Krause, J., Vida, T., and Geary, P. (2021). Integrating genetic, archaeological, and historical perspectives on Eastern Central Europe, 400–900 AD: brief description of the ERC Synergy Grant – HistoGenes 856453. Historical Studies on Central Europe *1*, 213–338.
22. Geary, P., and Veeramah, K. (2016). Mapping European population movement through genomic research. medievalworlds *4*, 65–78. https://doi.org/10.1553/medievalworlds_no4_2016s65.
23. QGIS Development Team (2022). QGIS Geographic Information System (QGIS Association).
24. O'Sullivan, N., Posth, C., Coia, V., Schuenemann, V.J., Price, T.D., Wahl, J., Pinhasi, R., Zink, A., Krause, J., and Maixner, F. (2018). Ancient genome-wide analyses infer kinship structure in an Early Medieval Alemannic graveyard. Sci. Adv. *4*, eaao1262.
25. Bondár, M., Honti, S., Márkus, G., and Németh, P.G. (2007). Balatonszemes-Szemesi berek. In Gördülo ido: régészeti feltárások az M7-es autópálya Somogy megyei szakaszán Zamárdi és Ordacsehi között, K. Belényesy, S. Honti, and V. Kiss, eds. (Somogy Megyei Múzeumok Igazgatósága), pp. 123–135.
26. Gallina, Z., and Straub, P. (2014). Fonyód–Vasúti-dűlő 2 (Mérnöki-telep) kora népvándorlás kori sírjai. In AVAROK PUSZTÁI Régészeti tanulmányok Lőrinczy Gábor 60. születésnapjára, A. Alexandra, B. Csilla, and T. Attila, eds. (MTA BTK MŐT), pp. 213–233.

27. Miháczi-Pálfi, A. (2018). A balatonszemesi 5. századi temető kisleletei. Anyagközlés és elemzés [Kleinfunde des Gräberfeldes von Balatonszemesaus dem 5. Jahrhundert. Katalog und Bewertung der Funde]. In "Hadak útján." A népvándorláskor fiatal kutatóinak XXVI. konferenciája. Gazdaság – kereskedelem – kézművesség (26th Conference of Young Scholars on the Migration Period. Economy – Trade – Craftsmanship) Dissertationes Archaeologicae Supplementum, Z. Rácz, I. Koncz, and B. Gulyás, eds. (Institute of Archaeological Sciences, Eötvös Loránd University), pp. 129–161.

28. Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H.A., Kelso, J., and Pääbo, S. (2013). DNA analysis of an early modern human from Tianyuan Cave, China. Proc. Natl. Acad. Sci. USA *110*, 2223–2227.

29. Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature *522*, 207–211.

30. Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503.

31. Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. *16*, 224.

32. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. BMC Bioinf. *15*, 356.

33. Nelson, M.R., Bryc, K., King, K.S., Indap, A., Boyko, A.R., Novembre, J., Briley, L.P., Maruyama, Y., Waterworth, D.M., Waeber, G., et al. (2008). The Population Reference Sample, POPRES: a resource for population, disease, and pharmacological genetics research. Am. J. Hum. Genet. *83*, 347–358.

34. Veeramah, K.R., and Hammer, M.F. (2014). The impact of whole-genome sequencing on the reconstruction of human population history. Nat. Rev. Genet. *15*, 149–162.

35. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.

36. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190.

37. Veeramah, K.R., Rott, A., Groß, M., van Dorp, L., López, S., Kirsanow, K., Sell, C., Blöcher, J., Wegmann, D., Link, V., et al. (2018). Population genomic analysis of elongated skulls reveals extensive female-biased immigration in Early Medieval Bavaria. Proc. Natl. Acad. Sci. USA *115*, 3494–3499.

38. Marcus, J., Ha, W., Barber, R.F., and Novembre, J. (2021). Fast and flexible estimation of effective migration surfaces. eLife *10*, e61927.

39. Jørsboe, E., Hanghøj, K., and Albrechtsen, A. (2017). fastNGSadmix: admixture proportions and principal component analysis of a single NGS sample. Bioinformatics *33*, 3148–3150.

40. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74.

41. Hakenbeck, S. (2018). Infant head shaping in Eurasia in the first millennium AD. In The Oxford Handbook of the Archaeology of Childhood, S. Crawford, D.M. Hadley, and G. Shepherd, eds. (Oxford University Press), pp. 483–504.

42. Alt, K.W., Knipper, C., Peters, D., Müller, W., Maurer, A.-F., Kollig, I., Nicklisch, N., Müller, C., Karimnia, S., Brandt, G., et al. (2014). Lombards on the move – an integrative study of the migration period cemetery at Szólád, Hungary. PLoS One *9*, e110793.

43. Jarnut, J. (1982). Geschichte der Langobarden (W. Kohlhammer).

44. Pohl, W., and Erhart, P. (2005). Die Langobarden: Herrschaft und Identität (Verlag der Österreichischen Akademie der Wissenschaften).

45. von Freeden, U., Vida, T., and Skriba, P. (2007). Ausgrabung des langobardenzeitlichen Gräberfeldes von Szólád, Komitat Somogy, Ungarn: Vorbericht und Überblick über langobardenzeitliche Besiedlung am Plattensee. Germania *85*, 359–384.

46. Hakenbeck, S.E., Evans, J., Chapman, H., and Fóthi, E. (2017). Practising pastoralism in an agricultural environment: an isotopic analysis of the impact of the Hunnic incursions on Pannonian populations. PLoS One *12*, e0173079.

47. Lipatov, M., Sanjeev, K., Patro, R., and Veeramah, K.R. (2015). Maximum likelihood estimation of biological relatedness from low coverage sequencing data. Preprint at bioRxiv. https://doi.org/10.1101/023374.

48. Monroy Kuhn, J.M., Jakobsson, M., and Günther, T. (2018). Estimating genetic kin relationships in prehistoric populations. PLoS One *13*, e0195491.

49. Rácz, Z. (2016). Zwischen Hunnen- und Gepidenzeit. Frauengräber aus dem 5. Jahrhundert im Karpatenbecken. Acta Archaeol. Acad. Sci. Hung. *67*, 301–359.

50. Rácz, Z. (2020). Who were the Gepids and Ostrogoths on the Middle Danube in the 5th Century? An archaeological perspective. In The Migration Period between the Oder and the Vistula, (Brill), pp. 771–789.

51. Barbiera, I. (2005). Changing Lands in Changing Memories: Migration and Identity during the Lombard Invasions (All'Insegna del giglio).

52. Geary, P.J. (2002). The Myth of Nations: The Medieval Origins of Europe (Princeton University Press).

53. Fehr, H. (2010). Germanen und Romanen im Merowingerreich: Frühgeschichtliche Archäologie zwischen Wissenschaft und Zeitgeschehen: Ergänzungsbände zum Reallexikon der Germanischen Altertumskunde (Gruyter, Walter de, & Co.).

54. Heather, P.J. (2005). The Fall of the Roman Empire: A New History of Rome and the Barbarians (Macmillan).

55. Meier, M. (2021). Geschichte der Völkerwanderung: Europa, Asien und Afrika vom 3. bis zum 8. Jahrhundert n.Chr (C.H. Beck).

56. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

57. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008, https://doi.org/10.1093/gigascience/giab008.

58. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760.

59. Kircher, M. (2012). Analysis of high-throughput ancient DNA sequencing data. Methods Mol. Biol. *840*, 197–228.

60. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

61. Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis (Springer).

62. Rubinacci, S., Ribeiro, D.M., Hofmeister, R.J., and Delaneau, O. (2021). Efficient phasing and imputation of low-coverage sequencing data using large reference panels. Nat. Genet. *53*, 120–126.

63. Ginolhac, A., Rasmussen, M., Gilbert, M.T.P., Willerslev, E., and Orlando, L. (2011). mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics *27*, 2153–2155.

64. Briggs, A.W., Good, J.M., Green, R.E., Krause, J., Maricic, T., Stenzel, U., Lalueza-Fox, C., Rudan, P., Brajković, D., Kućan, Ž., et al. (2009). Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science *325*, 318–321.

65. Hunter, J.D. (2007). Matplotlib: a 2D graphics environment. Comput. Sci. Eng. *9*, 90–95.

66. Fan, L., and Yao, Y.-G. (2013). An update to MitoTool: using a new scoring system for faster mtDNA haplogroup determination. Mitochondrion *13*, 360–363.

67. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

68. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.

69. R Core Team (2022). R: A language and environment for statistical computing (R Foundation for Statistical Computing).

70. Lamnidis, T.C., Majander, K., Jeong, C., Salmela, E., Wessman, A., Moiseyev, V., Khartanovich, V., Balanovsky, O., Ongyerth, M., Weihmann, A., et al. (2018). Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. Nat. Commun. *9*, 5018.

71. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

72. Zhang, Z., and Yuan, K.-H. (2018). Practical Statistical Power Analysis Using Webpower and R (Isdsa Press).

73. Pejrani Baricco, L., and Ratto, S. (2014). L'inattesa scoperta di una chiesa paleocristiana. Rivista MuseoTorino *7*, 10–13.

74. Giostra, C., Bedini, E., Caramelli, D., Mallegni, F., and Pejrani Baricco, L. (2012). Per una conoscenza dei Longobardi in Italia: primi risultati delle analisi genetiche su individui provenienti da necropoli del Piemonte. In VI Congresso Nazionale di Archeologia Medievale, L'Aquila, 12-15 settembre 2012 (ITA), pp. 448–453.

75. Pejrani Baricco, L. (2017). Bardonecchia (Torino), necropoli di ambito merovingio. In Longobardi. Un popolo che cambia la storia, catalogo della mostra (Pavia – Napoli – San Pietroburgo, 2017-2018), G.P. Brogiolo, F. Marazzi, and C. Giostra, eds. (Catalogo di mostra internazionale allestita presso il Castello Visconteo di Pavia, il Museo Archeologico Nazionale di Napoli e l'Ermitage di San Pietroburgo), pp. 76–77.

76. Bierbrauer, V. (2004). Die Keszthely-Kultur und die romanische Kontinuität in Westungarn (5.–8. Jh.). Neue Überlegungen zu einem alten Problem. In Von Sachsen bis Jerusalem. Menschen und Institutionen im Wandel der Zeit. Festschrift für Wolfgang Giese zum 65. Geburtstag, H. Seibert, and G. Thoma, eds. (Herbert Utz Verlag), pp. 51–72.

77. Heinrich-Tamáska, O. (2007). Bemerkungen zur Transformation spätantiker Strukturen in Pannonien am Beispiel von Keszthely-Fenékpuszta. Acta Archaeol. Carpathica *42*, 199–229.

78. Schilling, L. (2011). Bestattungen und Gräberfelder von der Spätantike bis zum Frühmittelalter in und um die spätrömische Befestigung von Tác/Gorsium (4.–8. Jh.). In Keszthely-Fenékpuszta im Kontext spätantiker Kontinuitätforschung zwischen Noricum und Moesia. Castellum Pannonicum Pelsonense 2, O. Heinrich-Tamáska, ed. (Leidorf), pp. 381–396.

79. Bollók, Á. (2016). A fifth-century scriptural amulet from Hács-Béndekpuszta in its Mediterranean context. In Between Byzantium and the Steppe: Archaeological and Historical Studies in Honour of Csanád Bálint on the Occasion of His 70th Birthday, Á. Bollók, G. Csiky, and T. Vida, eds. (Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences), pp. 31–61.

80. Gilbert, M.T.P., Bandelt, H.-J., Hofreiter, M., and Barnes, I. (2005). Assessing ancient DNA studies. Trends Ecol. Evol. *20*, 541–544.

81. Willerslev, E., and Cooper, A. (2005). Ancient DNA. Proc. Biol. Sci. *272*, 3–16.

82. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.

83. Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., et al. (2015). Optimal ancient DNA yields from the inner ear part of the human petrous bone. PLoS One *10*, e0129102.

84. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20130624.

85. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*. db.prot5448.

86. Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. *40*, e3.

87. Gansauge, M.-T., and Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat. Protoc. *8*, 737–748.

88. Gansauge, M.-T., and Meyer, M. (2019). A method for single-stranded ancient DNA library preparation. Methods Mol. Biol. *1963*, 75–83.

89. Rohland, N., Glocke, I., Aximu-Petri, A., and Meyer, M. (2018). Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. Nat. Protoc. *13*, 2447–2461.

90. Gansauge, M.-T., Gerber, T., Glocke, I., Korlevic, P., Lippik, L., Nagel, S., Riehl, L.M., Schmidt, A., and Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. Nucleic Acids Res. *45*, e79.

91. Gansauge, M.-T., Aximu-Petri, A., Nagel, S., and Meyer, M. (2020). 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.

92. DeAngelis, M.M., Wang, D.G., and Hawkins, T.L. (1995). Solid-phase reversible immobilization for the isolation of PCR products. Nucleic Acids Res. *23*, 4742–4743.

93. Gnecchi-Ruscone, G.A., Szécsényi-Nagy, A., Koncz, I., Csiky, G., Rácz, Z., Rohrlach, A.B., Brandt, G., Rohland, N., Csáky, V., Cheronet, O., et al. (2022). Ancient genomes reveal origin and rapid trans-Eurasian migration of 7th century Avar elites. Cell *185*, 1402–1413.e21.

94. Francalacci, P., Morelli, L., Angius, A., Berutti, R., Reinier, F., Atzeni, R., Pilu, R., Busonero, F., Maschio, A., Zara, I., et al. (2013). Low-pass DNA sequencing of 1200 Sardinians reconstructs European Y-chromosome phylogeny. Science *341*, 565–569.

95. Karmin, M., Saag, L., Vicente, M., Wilson Sayres, M.A., Järve, M., Talas, U.G., Rootsi, S., Ilumäe, A.M., Mägi, R., Mitt, M., et al. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. Genome Res. *25*, 459–466.

96. Poznik, G.D., Henn, B.M., Yee, M.-C., Sliwerska, E., Euskirchen, G.M., Lin, A.A., Snyder, M., Quintana-Murci, L., Kidd, J.M., Underhill, P.A., and Bustamante, C.D. (2013). Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. Science *341*, 562–565.

97. Ensor, B.E. (2013). The Archaeology of Kinship: Advancing Interpretation and Contributions to Theory (University of Arizona Press).

98. Hummer, H. (2018). Visions of Kinship in Medieval Europe (Oxford University Press).

99. Brück, J. (2021). Ancient DNA, kinship and relational identities in Bronze Age Britain. Antiquity *95*, 228–237.

100. Allentoft, M.E., Sikora, M., Sjögren, K.G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.

101. Broushaki, F., Thomas, M.G., Link, V., López, S., van Dorp, L., Kirsanow, K., Hofmanová, Z., Diekmann, Y., Cassidy, L.M., Díez-Del-Molino, D.,

et al. (2016). Early Neolithic genomes from the eastern Fertile Crescent. Science *353*, 499–503.

102. Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., et al. (2014). Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. *5*, 5257–5259.

103. Hofmanová, Z., Kreutzer, S., Hellenthal, G., Sell, C., Diekmann, Y., Díez-Del-Molino, D., van Dorp, L., López, S., Kousathanas, A., Link, V., et al. (2016). Early farmers from across Europe directly descended from Neolithic Aegeans. Proc. Natl. Acad. Sci. USA *113*, 6886–6891.

104. Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., et al. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.

105. Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D.C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., et al. (2016). Genomic insights into the origin of farming in the ancient Near East. Nature *536*, 419–424.

106. Mathieson, I., Alpaslan-Roodenberg, S., Szécsényi-Nagy, A., Szécsényi-Nagy, A., Rohland, N., Mallick, S., Olalde, I., Broomandkhoshbacht, N., Candilio, F., Cheronet, O., et al. (2018). The genomic history of south-eastern Europe. Nature *555*, 197–203.

107. Olalde, I., Allentoft, M.E., Sánchez-Quinto, F., Santpere, G., Chiang, C.W.K., DeGiorgio, M., Prado-Martinez, J., Rodríguez, J.A., Rasmussen, S., Quilez, J., et al. (2014). Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature *507*, 225–228.

108. van de Loosdrecht, M., Bouzouggar, A., Humphrey, L., Posth, C., Barton, N., Aximu-Petri, A., Nickel, B., Nagel, S., Talbi, E.H., El Hajraoui, M.A., et al. (2018). Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. Science *360*, 548–552.

109. Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., et al. (2016). The genetic history of Ice Age Europe. Nature *534*, 200–205.

110. Harney, É., Patterson, N., Reich, D., and Wakeley, J. (2021). Assessing the performance of qpAdm: a statistical tool for studying population admixture. Genetics *217*, iyaa045.

111. Peter, B.M. (2022). A geometric relationship of F2, F3 and F4-statistics with principal component analysis. Philos. Trans. R. Soc. Lond. B Biol. Sci. *377*, 20200413.

112. Zimmerman, K.D., Schurr, T.G., Chen, W.-M., Nayak, U., Mychaleckyj, J.C., Quet, Q., Moultrie, L.H., Divers, J., Keene, K.L., Kamen, D.L., et al. (2021). Genetic landscape of Gullah African Americans. Am. J. Phys. Anthropol. *175*, 905–919.

113. Petkova, D., Novembre, J., and Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. Nat. Genet. *48*, 94–100.

114. Al-Asadi, H., Petkova, D., Stephens, M., and Novembre, J. (2019). Estimating recent migration and population-size surfaces. PLoS Genet. *15*, e1007908.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Ancient skeletal element	This study	Bal_26
Ancient skeletal element	This study	Bal_111
Ancient skeletal element	This study	Bal_143
Ancient skeletal element	This study	Bal_144
Ancient skeletal element	This study	Bal_146
Ancient skeletal element	This study	Bal_148
Ancient skeletal element	This study	Bal_149
Ancient skeletal element	This study	Bal_150
Ancient skeletal element	This study	Bal_267
Ancient skeletal element	This study	Bal_268
Ancient skeletal element	This study	Bal_269
Ancient skeletal element	This study	Fonyod_278
Ancient skeletal element	This study	Fonyod_304
Ancient skeletal element	This study	Fonyod_305
Ancient skeletal element	This study	Fonyod_316
Ancient skeletal element	This study	Fonyod_336
Ancient skeletal element	This study	Fonyod_444
Ancient skeletal element	This study	Fonyod_468
Ancient skeletal element	This study	Fonyod_469
Ancient skeletal element	This study	Fonyod_489
Ancient skeletal element	This study	Fonyod_490
Ancient skeletal element	This study	Fonyod_491
Ancient skeletal element	This study	Fonyod_536
Ancient skeletal element	This study	Fonyod_666
Ancient skeletal element	This study	Hacs_1
Ancient skeletal element	This study	Hacs_4
Ancient skeletal element	This study	Hacs_5
Ancient skeletal element	This study	Hacs_10
Ancient skeletal element	This study	Hacs_13
Ancient skeletal element	This study	Hacs_15
Ancient skeletal element	This study	Hacs_16
Ancient skeletal element	This study	Hacs_17
Ancient skeletal element	This study	Hacs_18
Ancient skeletal element	This study	Hacs_19
Ancient skeletal element	This study	Hacs_20
Ancient skeletal element	This study	Hacs_21
Ancient skeletal element	This study	Hacs_22
Ancient skeletal element	This study	Hacs_24
Ancient skeletal element	This study	Bard_T1
Ancient skeletal element	This study	Bard_T2
Ancient skeletal element	This study	Bard_T6
Ancient skeletal element	This study	Bard_T7A
Ancient skeletal element	This study	Bard_T8_rid
Ancient skeletal element	This study	Bard_T10

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ancient skeletal element	This study	Bard_T11
Ancient skeletal element	This study	Bard_T12
Ancient skeletal element	This study	To_Lav_T10US67_ind1
Ancient skeletal element	This study	To_Lav_T10US67_ind2
Ancient skeletal element	This study	To_Lav_T1US6
Ancient skeletal element	This study	To_Lav_T2US16
Ancient skeletal element	This study	To_Lav_T37US60
Ancient skeletal element	This study	To_Lav_T38US344
Ancient skeletal element	This study	To_Lav_T90US339
Chemicals, peptides, and recombinant proteins		
Destilled Water DNA free, UltraPure	Thermo Fisher Scientific	Cat# 10977035
0.5 M EDTA pH 8.0	Thermo Fisher Scientific	Cat# AM9261
Proteinase K	Thermo Fisher Scientific	Cat# AM2548
Isopropanol	Sigma Aldrich	Cat# I9516
Guanidine hydrochloride	Sigma Aldrich	Cat# G4505
Sodium Acetate Solution (3 M), pH 5.2	Thermo Fisher Scientific	Cat# R1181
Tween 20	Sigma Aldrich	Cat# P2287
Buffer PE	Qiagen	Cat# 19065
Buffer PB	Qiagen	Cat# 19066
Tris-EDTA buffer solution	Sigma Aldrich	Cat# 93283
10x Buffer Tango	Thermo Fisher Scientific	Cat# BY5
ATP 100 mM	Thermo Fisher Scientific	Cat# R0441
BSA 20 mg/mL	Roche	Cat# 10711454001
dNTP Mix	Thermo Fisher Scientific	Cat# R1121
USER enzyme	New England Biolabs	Cat# M5505
Uracil Glycosylase inhibitor (UGI)	New England Biolabs	Cat# M0281
T4 Polynucleotide Kinase	New England Biolabs	Cat# M0201
T4 DNA Polymerase	New England Biolabs	Cat# M0203
Bst DNA Polymerase, large fragment	New England Biolabs	Cat# M0275L
Ethanol	Merck	Cat# 1009831000
10x T4 Ligase Buffer	Thermo Fisher Scientific	Cat# EL0011
T4 DNA Ligase	Thermo Fisher Scientific	Cat# EL0011
10x Thermopol Buffer	New England Biolabs	Cat# B9004S
Ampure XP	Bioscience	Cat# BCI-A63881
Agilent D1000 ScreenTapes	Agilent Technologies	Cat# 5067-5582
Agilent D1000 Ladder	Agilent Technologies	Cat# 5067-5586
Agilent D1000 Reagents	Agilent Technologies	Cat# 5067-5583
Agarose	Lonza	Cat# 50004
HyperLadder 25bp (formerly HyperLadder V)	Bioline	Cat# BIO-33057
ECO Safe Nucleic Acid Staining Solution 20,000X	Thermo Fisher Scientific	Cat# 3910001
2X Hi-RPM Hybridization Buffer	Agilent Technologies	Cat# 5190-0403
PfuTurbo Cx Hotstart DNA Polymerase	Agilent Technologies	Cat# 600412
Herculase II Fusion DNA Polymerase	Agilent Technologies	Cat# 600679
Sodiumhydroxide Pellets	Fisher Scientific	Cat# 10306200
Sera-Mag Magnetic Speed-beads. Carboxylate-Modified (1 mm, 3EDAC/PA5)	GE LifeScience	Cat# 65152105050250
Dynabeads MyOne Streptavidin	Thermo Fisher Scientific	Cat# 65602
SSC Buffer (20x)	Thermo Fisher Scientific	Cat# AM9770
GeneAmp 10x PCR Gold Buffer	Thermo Fisher Scientific	Cat# 4379874

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Salmon sperm DNA	Thermo Fisher Scientific	Cat# 15632-011
Human Cot-I DNA	Thermo Fisher Scientific	Cat#15279011
5M NaCl	Sigma Aldrich	Cat# S5150
1M NaOH	Sigma Aldrich	Cat# 71463
1 M Tris-HCl pH 8.0	Sigma Aldrich	Cat# AM9856
50x Denhardt's solution	Thermo Fisher Scientific	Cat# 750018
Methanol, certified ACS	VWR	Cat# EM-MX0485-3
Acetone, certified ACS	VWR	Cat# BDH1101-4LP
Dichloromethane, certified ACS	VWR	Cat# EMD-DX0835-3
Hydrochloric acid, 6N, 0.5N & 0.01N	VWR	Cat# EMD-HX0603-3
5 M Sodium chloride solution	Sigma-Aldrich	Cat# S5150-1L
20% SDS	Serva	Cat# 39575.01
PEG-4000	Thermo Fisher Scientific	Cat# EL001
PEG-8000	Promega	Cat# V3011
Critical commercial assays		
MinElute PCR Purification Kit	QIAGEN	Cat# 28006
TwistAmp Basic Kit	TwistDX	Cat# TABAS03kit
Qubit dsDNA HS Assay Kit, 500 assays	Thermo Fisher Scientific	Cat# Q32854
High Pure Extender Assembly from the Roche High Pure Viral Nucleic Acid Large Volume Kit,40 reactions	Roche	Cat# 5114403001
MiSeq Reagent Kit v3 (150 cycle)	Illumina	Cat# MS-102-3001
NextSeq 500/550 High Output Kit v2 (150 cycles)	Illumina	Cat# FC-404-2002
HiSeq Cluster Kit SR	Illumina	Cat# GD-410-1001
HiSeq 4000 SBS Kit (50/75 cycles)	Illumina	Cat# FC-410-1001/2
NextSeq 500/550 High Output Kit v2 (150 cycles)	Illumina	Cat# FC-404-2002
DyNAmo Flash SYBR Green qPCR Kit	Thermo Fisher Scientific	Cat# F415L
Maxima SYBR Green kit	Thermo Fisher Scientific	Cat# K0251
Oligo aCGH/Chip-on-Chip Hybridization Kit	Agilent Technologies	Cat# 5188-5220
Deposited data		
WGS sequence data from Balaton sites (BAM, fastq files) (NCBI Sequence Read Archive)	This Study	SRA: SRP362235
1240K sequence data from Balaton sites (BAM, fastq files) (NCBI Sequence Read Archive)	This Study	SRA: SRP362166
1240K and WGS sequence data from Italian sites (BAM, fastq files) (NCBI Sequence Read Archive)	This Study	SRA: SRP423598
Software and algorithms		
Admixtools (qpAdm, convertf)	Patterson et al.[56]; Haak et al.[29]	https://github.com/DReichLab/AdmixTools
Angsd v. 0.938	Korneliussen et al.[32]	https://github.com/ANGSD/angsd
bcftools	Danecek et al.[57]	http://www.htslib.org/doc/1.0/bcftools.html
bwa v. 0.7.17-r1188	Li and Durbin[58]	https://github.com/lh3/bwa
fastNGSadmix	Jørsboe et al.[39]	https://github.com/e-jorsboe/fastNGSadmix
fastq processing scripts	Kircher[59]	https://bioinf.eva.mpg.de/fastqProcessing/
FEEMS	Marcus et al.[38]	https://github.com/NovembreLab/feems
gatk v. 3.3	McKenna et al.[60]	https://gatk.broadinstitute.org/hc/en-usq
gatk v. 4.2	McKenna et al.[60]	https://gatk.broadinstitute.org/hc/en-us
ggplot2	Wickham[61]	https://cran.r-project.org/web/packages/ggplot2/index.html
GLIMPSE v. 1.1	Rubinacci et al.[62]	https://github.com/odelaneau/GLIMPSE
Google Maps API v3 Tool	N/A	https://www.birdtheme.org/useful/v3tool.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
lcMLkin	Lipatov et al.[47]	https://github.com/COMBINE-lab/maximum-likelihood-relatedness-estimation
mapDamage v.0.3.3	Ginolhac et al.[63]	https://github.com/ginolhac/mapDamage
Mapping Iterative Assembler	Briggs et al.[64]	https://github.com/mpieva/mapping-iterative-assembler
matplotlib	Hunter[65]	https://matplotlib.org/
MitoTool v. 1.1.2	Fan and Yao[66]	http://mitotool.kiz.ac.cn/
Picard Tools	McKenna et al.[60]	https://broadinstitute.github.io/picard/
Plink v. 1.90b6.9	Purcell et al.[67]; Chang et al.[68]	https://www.cog-genomics.org/plink/
QGiS v3.22.1	QGIS Development Team[23]	https://qgis.org/
R v. 4.1.2	R Core Team[69]	https://cran.r-project.org/
READ	Kuhn et al.[48]	https://bitbucket.org/tguenther/read/
samtools	Danecek et al.[57]	http://www.htslib.org/doc/samtools.html
Schmutzi	Renaud et al.[31]	https://github.com/grenaud/schmutzi
Sex.DetERRmine	Lamnidis et al.[70]	https://github.com/TCLamnidis/Sex.DetERRmine
smartPCA	Price et al.[35]; Patterson et al.[36]	https://github.com/DReichLab/EIG
vcftools	Danecek et al.[71]	https://vcftools.github.io/index.html
Veeramah lab genotype callers	N/A	https://github.com/kveeramah
vegan	N/A	https://cran.r-project.org/web/packages/vegan/index.html
WebPower	Zhang and Yuan[72]	https://cran.r-project.org/web/packages/WebPower/index.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Krishna R. Veeramah (krishna.veeramah@stonybrook.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Our newly generated sequence data from 53 individuals are available from the NCBI Sequence Read Archive (SRA) database under accession SRA: SRP362166 for 1240K data from Fonyod, Hács, and Balatonszemes; under accession SRA: SRP362235 for WGS data from Hács and Balatonszemes; and under SRA: SRP423598 for 1240K and WGS data from Bardonecchia and Torino Lavazza. We accessed the POPRES (Population Reference Sample) dataset collected and published by Nelson et al.[33] from dbGaP (accession dbGaP: phs000145.v4.p2).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Archaeological material and ethics permission

The authors declare that they had requested and got permission from the stakeholders, excavator, and processor anthropologists and archaeologists for the destructive DNA analyses of the anthropological material presented in this study.

We generated new genome-wide data from skeletal remains of 53 ancient individuals from the following sites:

Fonyód - Mérnöki telep (n = 13); Archaeological context reported in Gallina and Straub (2014).[26]
Hács - Béndekpuszta (n = 14); Archaeological context reported in Kiss (1995).[5]
Balatonszemes - Szemesi berek (n = 11); Archaeological context reported in Miháczi-Pálfi (2018).[27]
Torino - Lavazza (n = 7); Archaeological context reported in Pejrani Baricco and Ratto (2014).[73]
Bardonecchia (n = 8); Archaeological context reported in Giostra et al. (2012)[74] and Pejrani Baricco (2017).[75]

Description of archaeological sites

Fonyód, Hács, & Balatonszemes (Hungary)

All three sites are located on elevated loess ridges close to the southern shore of Lake Balaton, Hungary, an area that became increasingly important for the Roman military and civil administration with the foundation of a series of inner fortresses in the 3rd-4th centuries. These inner fortresses remained in use at least until the middle of the 5th century, sometimes even after the Roman abandonment of the area.[76–78] These sites represent small, rural communities with similarities in both age and sex distributions and burial customs. However, they show differences in terms of spatial organization of the burials and due to chronological differences in archaeological material (Figure 2). The three sites chronologically cover the second part of the 5th century with Fonyód dated to the middle third, Hács to the second half, and Balatonszemes to the end of the 5th century.

The most interesting feature of the Fonyód site is its unique structure: 19 graves forming six clusters lying at roughly equal distances of 50–60 m from each other (Figure 2A). Some of the burial groups lay along the boundaries of the investigated area, so the site cannot be considered completely excavated. The 19 graves contained 14 individuals with 7 adult females, 2 adult males, 5 children, and 5 graves without any human remains. Out of the 11 well-preserved skulls at least 7 showed signs of artificial cranial deformation (henceforth ACD), that includes both adult male, female, and child burials. Curiously enough, adult male and adult female burials never occurred in the same cluster; however, given the low number of burials, no far-reaching conclusions can be drawn from this pattern. Most of the graves are simple pits with two exceptions found in Group VI: the pit of Fonyod_489 followed the contours of the human body, while Fonyod_666 was a side niche grave. The site is dated to the middle of the 5th century, to the peak of the Hunnic movement based on the presence of certain artefact types, such as the so called 'nomadic mirrors' of the Čmi-Brigetio type from Fonyod_444 and Fonyod_489 or the silver pins from Fonyod_491.

The cemetery at Hács contains at least 29 burials with a very unbalanced sex ratio (13 identifiable adult females and only 3 adult males) (Figure 2B). The burials form several clusters, but since the cemetery has not been completely excavated and at least four graves were destroyed, its exact layout cannot be reconstructed. Burials were more concentrated in the northern section, while they were more dispersed in the southern part, lying at distances of 5–10 m from each other or forming pairs. The site is dated to the second half of the 5th century based on the Béndekpuszta-type brooches found in Hacs_19 and Hacs_20, while polyhedral earrings and double-sided combs give a wider chronological frame.[5] The most remarkable finds of the cemetery are the delicate lead sheet fragments—possibly used as an amulet—bearing a text inscribed with the Gothic uncial and Gothic cursive script, probably found in the heavily disturbed Grave 5.[79]

At Balatonszemes, 19 graves were unearthed with five adult females, one adult male, two unidentifiable adults, five children, and six pits lacking any human remains (Figure 2C). A large grave cluster with twelve graves is arranged into a fairly irregular north to south row. Three burials (Bal_267, Bal_268, Bal_269) lying some 7-8 m to its south form a separate cluster. Yet another burial (Bal_26) was uncovered some 200 m west of the other graves, whose association with the cemetery remains unclear, even though the grave goods indicate contemporaneity.[25] While most burials barely contained any artefacts, four burials (two adult females and two girls including one girl with ACD) were remarkably richly furnished. Archaeological dating is again based on female jewelry and brooches found in the four richly furnished burials are the best chronological indicators pointing to the end of the 5th century.[3,4]

Torino-Lavazza (Italy)

In the close vicinity of a long-known, late Roman necropolis (to the north of ancient Turin, on the other side of the Dora river) that was used between the 1st to 4th centuries, an early Christian funerary complex including burials, mausolea, and a church (with single nave and semi-circular apse) had been discovered in 2013.[73] The complex developed as an organic extension of the late Roman necropolis with the first burials and several mausoleums appearing at the site during the 3rd-4th centuries. This phase is dated based on late Roman artefacts (most notably glass) of the period, but also based on the structure and masonry technique of the buildings. One of the mausolea—an apsidal hall—was probably turned into a funerary church sometimes later attested by the presumed later dating of burials both in the apse and the hall.

Later, in the second half of the 4th century or during the 5th century a large building with a single hall and semi-circular apse facing to the west, was erected above the foundations of the earlier mausolea in the northern part of the site. The building has been interpreted as a suburban funerary basilica. Funerary use of the site continued well into the 5th century with burials appearing inside and around the church suggesting an intense and systematic funerary use. Most graves of this phase are similar to the late Roman tombs: rectangular pits with masonry walls (re)using Roman bricks, while their bottom is often paved with joined bricks or tiles. Due to the lack of artifacts in the burials and the absence of epigraphic evidence, burials of this phase are dated based on their stratigraphy, their relative position to the earlier and later burials. The sampled burials belong to this phase, however due to the lack of absolute dating methods, their dating can only be given as 5th-7th centuries. Later phases also show the systematic and accurate exhumation of the skeletal remains of the masonry tombs that suggest the partial abandonment of the site and degradation of the buildings probably used as stone quarries, although there is some evidence in written records that suggests the survival of the church into the 11th-12th centuries.

Bardonecchia (Italy)

In 2005 a small cemetery was excavated near the today's commune of Bardonecchia, 90 kms west from Turin, Italy at an intersection of multiple valleys and surrounded by mountains.[74,75] Altogether 16 graves—four of them without any human remains, while others containing multiple individuals—came to light, but as it is located on a steep, south-oriented slope, it is probable that an unknown number of graves were destroyed by erosion. The east-west oriented graves are simple pits with occasional stone slab or stone lining around the edges, and they form irregular north-south rows.

From male burials belt buckles and mounts, knives, a double-sided comb, and a scramasax came to light. Female artefact types, such as jewelry, are only known from the burial of a child, possibly a young girl, while other female burials lacked any grave goods or dress accessories. Based on the artifacts the site can be dated from the second half of the 6th and to the 7th century, but the site might have survived into the 8th century, as most of the burials are undatable due to the lack of grave goods. C14 dating of multiple burials attest the wide dating range between the 6th and 8th centuries.

Based on the abundance of males and much fewer females, the presence of a weapon and the relatively large number of fractures and cranial trauma, it has been suggested that the community using the Bardonecchia cemetery might have had a military role in controlling the region and the mountain passes leading through the Susa valley.

Artificial cranial deformation

Artificial cranial deformation (ACD) was found at all three of the 5th century sites we studied, but it was most prominent at Fonyód. Out of the 11 well-preserved skulls at Fonyód, at least 7 showed signs of ACD, including adult male, female, and child burials. Based on the techniques (i.e., placement and number of bandages) used, three different types of artificially deformed skulls were observed: Fonyod_304, 305, 336, 490, 491 constitute Type 1; Type 2 is Fonyod_489 and Type 3 is Fonyod 666.

The first type is *moderate fronto-occipital ACD with two bandages*, where the first pressure bandage had been tied around the forehead and the occipital bone and was adjusted more tightly than the second one. The second bandages were situated behind the bregmatic region and either under the mandible or around the nuchal region. This is the most common type at the site (Fonyod_304, 305, 336, 490, 491). Fonyod_468 shows similar characteristics, but the deformation is so mild that it can be the result of natural variation as well.

The second type is *severe oblique cranial deformation with one bandage*: the modification was made by applying one bandage fastened between the frontal and the occipital region producing an obliquely conical skull shape. This type is only present in Fonyod_489.

The third type is *severe oblique artificial cranial deformation with two bandages*. The modification is the result of a strongly fastened bandage encircling the frontal and occipital bone resulting in an oblique elongation and conical shape of the skull. The second bandage had been placed in a vertical position around the bregmatic region and either under the mandible or around the nuchal region. This type is only present in Fonyod_666.

While it was possible to distinguish three types of ACD based on the techniques used, the second (Fonyod_489) and third (Fonyod_666) types resulted in very similar, elongated skulls that would have been indistinguishable to people living at the time, so only two distinctive forms (three including not-deformed) can be recognized. Additionally, Fonyod_468 has marks that might indicate a slight modification, but any such deformation would have been unrecognizable during their lifetime.

There is a clear connection between the spatial organization of the site and ACD. All individuals with recognizable deformed skulls were found in Burial group V and VI, where all preserved skulls showed signs of this tradition. Fonyod_489 and Fonyod_666 are both found in Burial group VI. Individuals from Burial group I, II, III, or IV did not show any signs of ACD with the exception of the female with the possibly slightly modified skull from Fonyod_468.

We conducted a logistic regression comparing ACD to our PCA results among the 10 sequenced individuals with well-preserved skulls and coverage greater than 0.1×; we coded Fonyod_468 as without ACD. We found significant associations between PCA results and ACD.

At the other two-fifth century sites ACD is less prominent, but is still present in Hács (Hacs_23) and Balatonszemes (Bal_268). Hacs_23 is an adult female (who was not sequenced), and Bal_268 is a young girl. Both of these individuals had deformed skulls with similar shapes to type 2 and/or 3. Another interesting factor is that in our fastNGSadmix analyses, Bal_268 (unlike all other individuals from Balatonszemes and Hács) has significant proportions of Asian ancestry components in both the fastNGSadmix analyses using 1000 Genomes Project and penecontemporaneous reference panels (Figure 3). Interestingly, her maternal half-sister (Bal_267) lacks both ACD and these modern/historic Asian components.

METHOD DETAILS

Ancient DNA lab work and sequence processing

Petrous bone samples were collected from individuals from Fonyód, Hács, and Balatonszemes (as well as two penecontemporaneous Italian reference sites, Bardonecchia and Torino Lavazza), with the exceptions of Fonyod_304 and Fonyod_305 where tibial fragments were collected instead. Illumina libraries were created and processed for all DNA samples. For Hács, Balatonszemes, Bardonecchia, and Torino Lavazza exploratory shotgun sequencing was conducted. Based on endogenous DNA preservation, libraries either underwent a target enrichment for ~1.2 million genome-wide SNPs (1240K)[28–30]) before sequencing or instead direct whole deep genome sequencing (WGS). Alternatively, for Fonyód, all libraries directly went to 1240K target enrichment, omitting shotgun sequencing.

Molecular work for the specimens from Hács, Balatonszemes, Bardonecchia, and Torino Lavazza and from Fonyód was carried out at the University of Florence and the Institute of Archaeogenomics, Eötvös Loránd Research Network in Budapest (respectively). Initial steps were identical between all three sites. Work was conducted in dedicated ancient DNA clean room facilities, applying strict criteria to prevent contamination during all experimental procedures.[80,81] A silica-based protocol was used for DNA extraction and purification.[82] Moreover, blank controls were processed along with the samples during DNA extractions and libraries preparation to

monitor for contamination in reagents. Before sampling the bone powder, the outer layer of the temporal and petrous bones was brushed with disposable tools and irradiated by ultraviolet light (254 nm) for 30 min, to remove external contaminants. To maximize the recovery of well-preserved endogenous DNA, powder was collected from the densest part of the pars petrosa (inner ear) as described in Pinhasi et al.[83] (or from the tibia in case of Fonyod_304 and Fonyod_305), using a low-speed micro drill equipped with a disposable disk saw and dental burs. Subsequent steps differ between data from Hács, Balatonszemes, Bardonecchia, and Torino Lavazza (which were processed in 2018) and data from Fonyód (which were processed in 2021). Also, for the four sites processed in 2018 double-stranded, non-UDG treated libraries were generated; alternatively for Fonyód single-stranded libraries that were generated and UDG treated in a manner with results functionally similar to partial UDG treatment or UDG-half.[84]

For Hács, Balatonszemes, Bardonecchia, and Torino Lavazza, libraries were were prepared at the University of Florence, Italy and 1240K library captures were conducted at the Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, Germany (Table S1). For Fonyód, all libraries were prepared and underwent 1240K capture at the Max Planck Institute for Evolutionary Anthropology (MPI-EVA) Leipzig, Germany. All 1240K enrichment was conducted using the wetlab protocols from Fu et al.,[28] Haak et al.,[29] and Mathieson et al.[30]

Hács, Balatonszemes, Bardonecchia, and Torino Lavazza

Libraries from Hács, Balatonszemes, Bardonecchia, and Torino Lavazza were sequenced in 2018, while the libraries from Fonyód were sequenced separately in 2021. For the sites processed in 2018, all sequencing was performed at the New York Genome Center (NYGC) (New York City, USA). Initially, 40 double-stranded libraries were shotgun sequenced at Stony Brook University using an Illumina benchtop sequencer (MiSeq). All 40 libraries passed screening and were chosen for further sequencing; the libraries that had the most reads mapping to the human reference genome (GRCh37) were chosen for whole genome sequencing (WGS), whereas the other libraries underwent 1240K capture sequencing for 1.24 million SNPs. In total, eleven were chosen for WGS, nine from Hács, one from Balatonszemes, and one from Bardonecchia. For 1240K sequencing from Hács (n = 5), Balatonszemes (n = 10), Bardonecchia (n = 7), and Torino Lavazza (n = 7), double-stranded libraries were sequenced using paired-end 125bp sequencing on an Illumina HiSeq 2500 sequencer at NYGC. For WGS sequencing, double-stranded libraries were sequenced using single-end 100bp sequencing on an Illumina NovaSeq S4 sequencer at NYGC with each WGS library sequenced over four different flow cell lanes.

Between 46 and 55 mg of bone powder per sample were used for DNA extraction at the Laboratory of Molecular Anthropology and Paleogenetics at the University of Florence (Italy). Extraction was conducted using a silica-based protocol that allows ancient DNA molecules to be efficiently recovered even if highly fragmented.[82] For these four sites, Illumina sequencing libraries without enzymatic damage repair were prepared from 20 μL of each extract using a double-strand and double-indexing protocol optimized for ancient samples.[85,86] After quality control on Agilent 2100 Bioanalyzer (DNA 1000 chip), libraries were pooled in equimolar amounts and sent to the New York Genome Center (NYGC) for sequencing where all 40 double-stranded libraries were shotgun sequenced using an Illumina benchtop sequencer (MiSeq) along with libraries from other sites (not discussed here). All libraries passed screening.

Genomic libraries were subsequently sent to the Max Planck Institute for the Science of Human History (MPI-SHH) where libraries were enriched for endogenous human DNA using capture probes targeting ~1.24 million SNPs (1240K)[28–30] using a protocol similar to Fonyód described below. These 1240K libraries were sent to NYGC for paired-end 125bp sequencing on an Illumina HiSeq 2500 sequencer. Nine libraries from Hács, one from Balatonszemes, and one from Bardonecchia (Hacs_1, Hacs_4, Hacs_5, Hacs_10, Hacs_13, Hacs_17, Hacs_18, Hacs_21, Hacs_22, Bal_268, and Bard_T1), which performed best in the screening, also had WGS libraries prepared at the University of Florence. WGS libraries were sent to NYGC for single-end 100bp sequencing on an Illumina NovaSeq S4 sequencer with each WGS library sequenced over four different flow cell lanes.

Fonyód

For Fonyód, all available genomic libraries underwent 1240K capture sequencing without any screening phase. These libraries were single-stranded[87,88] and were UDG treated in a manner with results functionally similar to partial UDG treatment or UDG-half.[84] These libraries were sequenced using single-end 76bp sequencing on an Illumina HiSeq 4000 sequencer at the MPI-EVA.

DNA extraction and subsequent steps of sample preparation were performed in the Ancient DNA Core Unit of the Max Planck Institute for Evolutionary Anthropology (MPI-EVA), Leipzig, Germany. DNA was extracted from between 21.8 and 35.7 mg of sample material using the same silica-based method optimized for the recovery of short DNA fragments.[82] Briefly, lysates were prepared by adding 1 mL of extraction buffer (0.45 M EDTA, pH 8.0, 0.25 mg/mL proteinase K, 0.05% Tween 20) to the sample material in 2.0-mL Eppendorf Lo-Bind tubes and rotating the tubes at 37°C for approximately 16 h.[82,89] Using an automated liquid handling system (Bravo NGS Workstation B, Agilent Technologies), DNA was purified from 150 μL lysate using silica-coated magnetic beads and binding buffer D as described in Rohland et al.[89] Elution volume was 30 μL. Extraction blanks without sample material were carried alongside the samples during DNA extraction.

DNA libraries were prepared from 30 μL extract using an automated version of single-stranded DNA library preparation[90] described in detail in Gansauge et al.[91] *E. coli* Uracil-DNA-glycosylase and *E. coli* endonuclease VIII were added during library preparation to remove uracils in the interior of molecules. Libraries were prepared from both the sample DNA extracts and the extraction blanks, and additional negative controls (library blanks) were added. Library yields and efficiency of library preparation were determined using two quantitative PCR assays.[91] The libraries were amplified and tagged with pairs of sample-specific indices using AccuPrime Pfx DNA polymerase as described in Gansauge et al.[91] Amplified libraries were purified using SPRI technology[92] as described in Gansauge et al.[91]

Sample and control libraries were enriched for endogenous human DNA using capture probes targeting ~1.24 million SNPs (1240K).[28–30] Two consecutive rounds of in-solution hybridization capture were performed using the method of Fu et al.[28] automated on the Bravo NGS workstation B. Pools of up to 24 libraries (also of different projects) were created and sequenced in single-end mode for 76 cycles on 3 lanes of 2 independent runs of an Illumina HiSeq 4000 sequencer.

QUANTIFICATION AND STATISTICAL ANALYSIS

Processing of novel genomic dataset

Sequencing data were processed based on a pipeline from Kircher.[59] Reads were trimmed and merged (when necessary) and mapped to GRCh37 using samtools.[57] Duplicate reads were marked using Picard Tools,[60] and reads less than 30bp long were filtered out. All untrimmed reads were excluded from processing and analysis, as these reads are more prone to contain contaminant sequences. Alignment was conducted with BWA.[58] For WGS data, as each library was sequenced over four flow cell lanes, data from each lane were initially processed separately and BAM files were merged prior to marking duplicate reads. We used mapDamage 0.3.3 to assess postmortem DNA damage patterns.[63]

We calculated coverage for all individuals for the autosomal 1240K SNPs (as well as for the mitochondrial genome) using gatk v4.2 DepthOfCoverage; for novel WGS data, we also calculated full autosomal coverage using gatk v3.3 DepthOfCoverage in multi-threaded operation.[60] For each individual, we calculated the number of autosomal 1240K SNPs based on the number of these SNPS covered by at least one read with mapping and base quality scores greater than or equal to 30. This was calculated for each BAM file using samtools depth.[57] Individuals' genetic sex was identified using the Sex.DetERRmine pipeline based on relative X and Y chromosome coverage (https://github.com/TCLamnidis/Sex.DetERRmine).[70] Subsequently, we used Angsd to estimate nuclear contamination rates of genetic males based on the hemizygous X chromosome.[32] We also used Schmutzi to estimate mitochondrial contamination rates for all individuals.[31] For seven individuals, Schmutzi failed to complete (Bard_T2, Bard_T6, To_Lav_T10US67_ind2, To_Lav_T1US6, To_Lav_T2US16, To_Lav_T37US60, To_Lav_T38US344), likely due to insufficient non-duplicate reads, similar to the results from Gnecchi-Ruscone et al.[93] However, six of these seven were males with successful Angsd results suggesting low contamination (thus only To_Lav_T38US344 lacks any contamination estimates) (Table S1).

BAM files from these novel sites were called using in-house scripts (https://github.com/kveeramah). Non-UDG treated data was processed using a caller that incorporates damage patterns from mapDamage, while UDG treated data was instead processed using an indent caller that disregarded the first and last five bases of any given read. Output VCF files include diploid genotypes as well as genotype likelihoods. Information about all individuals newly processed and analyzed for this paper are presented in Table S1.

Mitochondrial and Y chromosome analyses

For mitochondrial analysis, BAM files from the four Balaton region sites (as well as Bardonecchia and Torino Lavazza) were filtered down to reads aligning to the mitochondrial genome as fastq files using samtools. Mitochondrial genomes from individuals with >1500 mitochondrial reads were assembled using Mapping Iterative Assembler[64] (which was designed for ancient mitochondrial genome assembly). Haplotype assignment was made using MitoTool 1.1.2.[66]

We also identified the Y chromosome haplogroups for the genetic males from Fonyód, Hács, Balatonszemes, Bardonecchia, and Torino Lavazza. We restricted the analysis to 1240K Y chromosome SNPs among 1240K-captured males but analyzed the whole chromosome for the five WGS males. Phylogenetic position of each Y chromosome SNP was ascertained based on published databases (https://isogg.org/)[40,94–96] and used to identify NRY haplogroups.

Comparative dataset

We analyzed our newly sequenced data alongside comparative datasets.[6–12,14–19,37] All individuals analyzed had date ranges overlapping with the 4th-8th centuries CE with the exception of four individuals from Olalde et al.,[11] where date cut-offs were broadened to increase the number of Iberian reference individuals (given our cut-off coverage of 0.1×). For all comparative individuals, coverage of the autosomal 1240K SNPs was calculated using gatk DepthOfCoverage,[60] and (with the exception of four individuals from Szólád) individuals with <0.1× coverage were excluded from our dataset and all analyses. In most cases, we used the BAMs published in the study, but in some cases we used BAMs that were re-processed in the Veeramah lab from published fastq files using the same type of pipeline as our novel data. BAMs were called using in-house genotype callers consistent with our novel data. Information on all comparative individuals analyzed are presented in Table S1.

All individuals (both reference and novel) were called for genotypes from the whole genome, imputed for all diploid, dinucleotide sites within the 1000 Genomes Project Phase 3 v5a VCF files using GLIMPSE v1.1[62] (using the 1000 Genomes Project data as the imputation reference),[40] and filtered down to the 1240K positions.

Biological relatedness assessment

We used the software package lcMLkin to identify close biological relatedness between the individuals from all four Balaton sites (as well as Bardonecchia and Torino Lavazza).[47] All individuals were included regardless of level of coverage. We used a modified version of lcMLkin that included external allele frequency data, similar to Amorim et al.[6] Analyses were performed using genotype likelihood data from 1,079,996 autosomal 1240K sites as input. We used allele frequency data from the 1000 Genomes CEU and TSI populations (as well as the merger of the two, CEU+TSI).[40] We ignored all relationships between low-coverage individuals

with minimal common SNP coverage (i.e., less than 5,000 shared SNPs in an lcMLkin analysis) as these are likely spurious (not real) relationships. We also conducted a READ analysis[48] using pseudohaploid genotypes for 164,127 transversions with minor allele frequencies greater than 0.05 in the CEU+TSI merged population. For this analysis, we used 87 individuals with coverage $\geq 0.1\times$ from the four Balaton sites (as well as Bardonecchia and Torino Lavazza). READ results were completely consistent with lcMLkin results for all first/second degree pairs.

A major aim of this study was to examine to what extent genetic connections between individuals observed at these post-Roman 5th century Pannonian sites were known and acknowledged by members of their communities and thus could be understood as a component of social kinship. We cannot assume that all genetic relationships were known or acknowledged by the communities, so these genetic connections do not necessarily manifest in social reality. But if they are reflected in the archaeological data, we can assume that they have held importance in the life of the individuals, as the community (or part of the community responsible for the funeral) chose to express them in the burial customs through the placement of the burials or similarities in grave goods, etc. Biological relatedness might form the basis of or play an important role in the construction of social kinship, but social kinship can also be organized based on shared space, imagined or real descent, economical connection, social agreement, etc. as it is an outcome of social actions, developed and expressed through culturally defined social practices, including funerals[97–99] While there is a slight chronological difference between the three sites) the sites show similar characteristics in demography, funerary customs, and burial representation, and at all three sites, biological kindreds consisted of very close, first- and second-degree relatives, mostly on the mother's side. This is in stark contrast to Szólád where the cemetery was organized largely around male biological relatives with a large extended pedigree.

In Fonyód and Balatonszemes biologically related individuals are buried near each other suggesting that biological relatedness likely had social value to the community. Except for the three related females buried in richly furnished burials (Bal_267, Bal_268, and Bal_269) in Balatonszemes, these connections are not observable in the grave goods, probably as a result of differences in burial representation between the different genders. Adult male and child burials contain very few if any artefacts (notable exceptions to the latter are Bal_267 and Bal_268), while social differences are more clearly expressed in female burials.[49,50] At Hács, in contrast, the lack of spatial clustering and similarities in terms of burial representation may initially suggest a reduced role for biological relatedness among the three individuals directly related along the maternal line, a connection that is especially hard to hide among members of a small community. Previous stable isotope results[46] suggest that Hacs_4 was most likely raised locally, but probably gave birth to her daughter elsewhere, as the 87Sr/86Sr ratio of Hacs_20 falls completely outside of the regional values of the Lake Balaton area. Ancestry clustering analysis shows that across the generations of this kindred that SCAND proportions are being replaced with NGBI proportions (Figures 2B and 4), suggesting that. Hacs_4 and Hacs_20 had children with men who had predominantly NGBI ancestry. This signal of replacement is not visible when using the 1000 Genomes Project reference panel. The fact that both women came back to this community later during their lives and were buried at Hács indicates that their connection to the community never ceased to exist and that their biological relatedness nevertheless reflected social ties as well.

Principal component analysis

We analyzed the data from the Balaton region sites alongside data from 492 penecontemporaneous comparative individuals. PCAs were conducted to place ancient individuals one-by-one onto a background of modern individuals. A Procrustes transformation[6,34] was used to merge 561 ancient Europeans onto a single PCA (with a background of modern individuals) using smartPCA.[35,36] PCAs were plotted using matplotlib.[65] We used modern European populations from the POPRES dataset[33] as imputed by Veeramah et al.[37] We conducted analyses using all overlapping 328,670 SNPs. For each Balaton region site as well as for each geographic region, we also calculated covariance confidence ellipses based on Pearson correlation coefficients; ellipses were given radii corresponding to 1.5 standard deviations.

For this analysis, we converted our diploid 1240K VCF files to pseudohaploid transposed Plink[67,68] datasets using a custom, in-house script. Heterozygous genotype calls were converted to calls for the allele with greater allele depth; when both alleles had the same depth, an allele was chosen at random. All analyzed POPRES genotypes were also made pseudohaploid by randomly choosing one allele for all heterozygous genotypes.

We used an automated script to use smartPCA[35,36] to perform PCAs on the POPRES dataset including one ancient individual at a time (https://github.com/ShyamieG/). We then used an in-house script to conduct a Procrustes transformation,[6,34] merging all ancient individuals (with 1240K coverage of at least $0.1\times$) onto a single principal component analysis (Figures 3 and S1). In our PCA plots for these analyses, we divide POPRES populations into regions (i.e., CE, EE, NE, NEE, NWE, SE, SEE, and WE) as demarcated in Veeramah et al.[37]

We also use normalized PC1 and PC2 results from the PCA for our logistic regressions for testing association between artefacts (or ACD) and genetic ancestry (Table S4). Logistic regressions testing association of PCA results with dress accessories and ACD were calculated using R v4.1.2[69]; power analyses were conducted using the WebPower package.[72] Hosmer and Lemeshow's R^2 and chi-square p values for each regression were calculated within R (Table S4).

Gene flow modeling analysis

We used FEEMS[38] in order to model gene flow in a spatial context across Europe using our Lake Balaton and penecontemporary individuals. We used GLIMPSE imputations described above. FEEMS has a native imputation feature for missing data, but this simply relies on the mean allele frequencies across individuals for a given SNP, and does not account for low coverage. FEEMS also relies on

population allele frequencies, which are likely to be estimated more accurately for low coverage data than individual genotypes. Of the 566 4-8th century Europeans available to us, 26 were removed due to inferred 1-3rd degree relatedness between other individuals due to the reliance on unbiased population frequencies. We further set a maximum of n = 100 for individuals represented from a particular macro-region (for example British Isles) in order to mitigate as much as is possible against biases that result from oversampling in one region and undersampling in others. For macro-regions with more than 100 individuals, we randomly chose 100 representatives. We filtered out SNPs with a minor allele frequency less than 0.1 in the CEU 1000 Genomes populations[40] and performed linkage disequilibrium filtering using Plink 1.9 using –indep-pairwise 50 5 0.2.[67,68] Thus the final dataset consisted of 371 4-8th century individuals from across Europe with imputed diploid genotypes at 76,669 SNPs. We note that this compares favorably to the example use case of FEEMS in Marcus et al.[38] using 111 grey wolves from across North America genotyped at 17,729 SNPs (though that dataset has a more even geographic spread and relies on true diploid genotypes rather than imputation). An approximate polygon of Europe encompassing the geographic location of our sampleset was obtained using a Google Maps API tool (https://www.birdtheme.org/useful/v3tool.html). FEEMS was run under default parameters under three different lambda values (20, 2, 0.2). Decreasing lambda decreases the smoothing but can lead to overfitting, but we found a lambda of 0.2 to provide the best resolution of meaningful geographic barriers, with larger numbers leading to overly smooth surfaces that were not informative, and thus we provide visualizations from using this value. We conducted this analysis using all Lake Balaton individuals simultaneously (Figure 3C) and for each of the four populations individually (Figure S2).

Model-based clustering analyses

We conducted supervised ancestry clustering analyses using fastNGSadmix.[39] In our study, we took a new approach to supervised model-based clustering analyses. Instead of solely using modern data for reference panels, we also conducted analyses using penecontemporary reference populations. All penecontemporary individuals used in reference panels were imputed for the entire genome (based on the 1000 Genomes Project Phase 3 v5a vcf files)[40] using Glimpse v1.1[62] and then filtered down to the autosomal 1240K positions. Due to a bias in coverage in favor of the more northern reference populations in our set (Table S1), imputation was necessary. Preliminary, non-imputed analyses found considerable bias towards the northern components as there was substantially more genotype data from the higher coverage data from northern Europe.[9,10,12] Furthermore, this bias also made it impossible to filter imputed genotypes based on genotype probabilities (as these are closely linked to coverage).

For panel construction, we attempted to construct penecontemporaneous reference panels[7,9–12,14–19] to mirror the 1000 Genomes Project populations (as used by Amorim et al.[6]). We constructed seven reference panels of imputed penecontemporaneous individuals representing Mediterranean Europe (Italy/Iberia; MEDEU, n = 40), Northern Germany/Britain (NGBI, n = 40), Scandinavia/Estonia (SCAND, n = 40), East Asia (EASIA, n = 16), South Asia (SASIA, n = 17), North Africa (NAFRICA, n = 20), and sub-Saharan Africa (SUBSAHARAN, n = 8).[7,9–12,14–19] We could not create a working historic analog for the IBS population; thus, instead we merged the penecontemporaneous Italian and Iberian groups into a single MEDEU panel.

For the three European panels (MEDEU, NGBI, and SCAND), we picked individuals with the highest coverage while preventing any kin pairs from being put into a panel together.[7,9–12] We also included two Asian and two African panels: EASIA (n = 16 from what is now Hanben, Taiwan), NAFRICA (n = 20 from the island of Kulubnarti in what is now Sudan), SASIA (n = 17 from Roopkund Lake in what is now India), and SUBSAHARAN (n = 8 from various sites).[14–19] The reduction in SUBSAHARAN size owes directly to the rarity of ancient DNA from sub-Saharan Africa, especially from our time window. Given the closer relationships between the European reference panels, we used larger sample sizes (n = 40) for these panels, while using fewer for the non-European panels. We also conducted analyses using modern 1000 Genomes Project populations, CEU+GBR, FIN, IBS, TSI, YRI, EAS, and SAS; as per Amorim et al.,[6] we merged the CEU and GBR populations into a single population, as the two populations are not properly distinguishable by these types of analyses.

We generated the Beagle PL files for the genotype likelihoods used by fastNGSadmix for all individuals with coverage $\geq 0.1\times$ for 1,091,054 autosomal 1240K sites using vcftools.[71] For each individual, we ran the fastNGSadmix analysis 50 times and selected the run with greatest (i.e., least negative) likelihood. For the selected run, we re-ran the analysis using the exact same random seed to replicate the final result and calculate 100 bootstraps for Lake Balaton individuals and ten bootstrap for penecontemporary individuals. This approach was necessary as calculating bootstraps 50 times per individual would have been too computationally taxing, and even more so on lower coverage individuals who take longer to converge. All fastNGSadmix results were plotted using ggplot2.[61] In Figure 4 of the main text, we have presented our analyses for the four Lake Balaton sites using both modern and penecontemporary panels. In Figure S3, we have also presented our results from the penecontemporaneous individuals. We found that for the most part the bootstrap analyses were largely consistent with the converged results. When we analyzed the reference individuals, we found essentially perfect recall (i.e., each individual is assigned the ancestry that they are a reference for) (Figures S3B and S3C). It is important to note that while the reference panel was imputed, the data put into fastNGSadmix were raw, un-imputed genotype likelihoods; thus, this supports the accuracy of our imputations.

We also conducted a version of the analysis where we used separate IBERIA (n = 25) and ITALY (n = 40) reference panels (Figure S3C). The IBERIA panel was smaller due to fewer available individuals; outside of Olalde et al.,[11] we were unable to find additional published, penecontemporary Iberians and even extended our 4th-8th century date cut-offs to include I8339, I10866, I10892, and I10895. Based on analyzing bootstraps the IBERIA component frequently appeared and disappeared within individuals between bootstraps. We particularly found these cases in Fonyod_304, Fonyod_316, and Fonyod_489 as well as several individuals from the datasets from Veeramah et al.[37] and Gretzinger et al.[9] We also noticed the lack of IBERIA in individuals such as Bard_T11,

FN2, I3056, and IND006, that were clustered with Iberians in our PCAs and contained significant IBS proportions in the 1000 Genomes Project analysis (Figures 3 and S3). Based on these findings, we found that the IBERIA component could not be fully differentiated, and thus opted to merge IBERIA and ITALY into MEDEU for our primary analyses.

Analyses using qpAdm

In order to validate the results from our fastNGSadmix and FEEMS results suggesting an increase in post-Fonyód northern European gene flow into Lake Balaton, we also conducted a qpAdm analyses[29] using admixtools (https://github.com/DReichLab/AdmixTools).[56] For this analysis, we used pseudohaploid genotype data. As qpAdm can be sensitive to coverage differences between individuals, we conducted these analyses using pseudohaploid genotype calls generated with a random read indent caller that disregarded the first and last eight bases of any given read (https://github.com/kveeramah/). This calling method is consistent with the approach (https://github.com/DReichLab/adna-workflow) used in the lab where qpAdm was designed. We used a dataset of 144,109 biallelic transversions from the autosomal 1240K sites.

Each of the four Balaton communities were successively set as the target population. Consistent with the FEEMS analysis, we removed individuals with inferred 1-3rd degree relatedness with other individuals. For our "right" (reference) populations, we use 72 prehistoric individuals of Anatolian_Neolithic (n = 26), Steppe_Eneolithic (n = 18), Western Hunter-Gatherer (WHG) (n = 15), Iran_Neolithic (n = 9), and Morocco_Iberomaurusian (n = 4) origin (Table S1).[29,30,100–109] Individuals for these populations were chosen from the union of individuals used in prehistoric analyses by Amorim et al.[6] and Antonio et al.[7] We obtained published BAM files for all these individuals and called them for the autosomal 1240K SNPs consistent with all other individuals. For our "left" (source) populations, we used the penecontemporaneous reference individuals used in our fastNGSadmix panels. Due to a limit of four different source populations (when using five reference populations), for each Balaton community we ran six different sets of source populations. For the first four, we used the three European panels (NGBI, SCAND, and MEDEU) alongside one non-European panel (EASIA, SASIA, NAFRICA, or SUBSAHARAN). For the next set, we only included the three European panels (NGBI, SCAND, and MEDEU). For the final set, we merged NGBI and SCAND (i.e., NGBI+SCAND and MEDEU). We present the results from all the "best" models in Table S2 and we plot the model with the highest tail probabilities from the two European-only sets in Figure S4. It is important to remember that for qpAdm probability interpretation p values higher than 0.05 indicate that you cannot reject the model.[110]

It must be noted that the qpAdm approach was developed assuming a model involving the evolution of drift in discrete source populations with little or no gene flow followed by a single pulse of admixture to create the target population.[111] As such, it is has proved useful for analyses involving highly diverged source populations (for example such as estimating mixtures of WHG, Anatolia_Neolithic, and Steppe_Eneolithic in the earliest European farmers[105] or estimating mixtures of African, European, and Native American ancestry in African Americans[112]). As far as we are aware, it has not been shown to be applicable to studying closely related populations, such as modern Europeans or the penecontemporary populations examined here (who largely reflect the same level of diversity as modern populations). Using the program Angsd,[32] we estimated F_{ST} values for NGBI, SCAND, and MEDEU as well as WHG, Anatolia_Neolithic, and Steppe_Eneolithic. Among the three penecontemporaneous panels, we calculated unweighted F_{ST} values of 0.009 (SCAND vs. NGBI), 0.012 (SCAND vs. MEDEU), and 0.009 (NGBI vs. MEDEU), while for the prehistoric individuals (which we used for reference populations), we found values approximately an order of magnitude larger of 0.082 (WHG vs. Anatolia_Neolithic), 0.066 (WHG vs. Steppe_Eneolithic), and 0.057 (Anatolia_Neolithic vs. Steppe_Eneolithic). These F_{ST} demonstrate that the Late Antique/Early Medieval source populations we are studying are much more closely related than the typical qpAdm use case, which has likely resulted in some of our tail probabilities being very low for even the "best" models calculated by qpAdm, and the models generally showing poor stability from run to run and very coarse estimated admixture proportions.

Despite the fact that the assumptions of the qpAdm model does not fit our use case in almost all respects, it is still noteworthy that our results appear to broadly validate our findings from FEEMS of increased northern European ancestry post-Fonyód (Figures 3D and S2). However, studies of ancient populations involving spatially varying gene flow and with low genetic differentiation are likely to be a better fit for approaches such as FEEMS[38] (or EEMS[113] or MAPS[114]) as they do not require discrete source populations and instead explicitly model drift in a spatial context on which estimates of gene flow between neighboring populations/demes can be calculated.