

High Coverage Mitogenomes and Y-Chromosomal Typing Reveal

Ancient Lineages in the Modern-day Székely Population in Romania

Noémi Borbély^{1,2}, Orsolya Székely¹, Bea Szeifert^{1,2}, Dániel Gerber^{1,2}, István Máthé³, Elek Benkő⁴, Balázs Gusztáv Mende¹, Balázs Egyed⁵, Horolma Pamjav⁶, Anna Szécsényi-Nagy¹

1 Institute of Archaeogenomics, Research Centre for the Humanities, Eötvös Loránd Research Network, 1097 Budapest, Tóth Kálmán street 4, Hungary

2 Doctoral School of Biology, Institute of Biology, ELTE Eötvös Loránd University, 1117 Budapest, Pázmány Péter sétány 1/C, Hungary

3 Sapientia Hungarian University of Transylvania (Cluj-Napoca), Faculty of Economics, Socio-Human Sciences and Engineering, Department of Bioengineering, 530104 Miercurea-Ciuc, Piața Libertății 1, Romania

4 Institute of Archaeology, Research Centre for the Humanities, Eötvös Loránd Research Network, 1097 Tóth Kálmán street 4. Budapest

5 Department of Genetics, Faculty of Natural Sciences, ELTE Eötvös Loránd University, 1117 Budapest, Pázmány Péter sétány 1/C, Hungary

6 Laboratory of Reference Samples Analysis, Institute of Forensic Genetics, Hungarian Institute for Forensic Sciences, 1087 Budapest, Mosonyi street 9, Hungary

Correspondence: szechenyi-nagy.anna@abtk.hu; borbely.noemi@abtk.hu

Keywords

Székelys; Hungarians; Transylvania; mitogenomes; Y-chromosome; genetic diversity; phylogeography; phylogeny

Abstract

Here we present 115 whole mitogenomes and 92 Y-chromosomal STR and SNP profiles from a Hungarian ethnic group, the Székelys (in Romanian: Secuii, in German: Sekler) living in southeast Transylvania (Romania). The Székelys can be traced back to the 12th century in the region, and numerous scientific theories exist as to their origin. We carefully selected sample providers that had local ancestors inhabiting small villages in the area of Odorheiu Secuiesc/Székelyudvarhely in Romania. The results of our research and the reported data signify a qualitative leap compared to previous studies, since complete mitochondrial DNA sequences and Y-chromosomal data containing 23 STRs have not been available from the region until now. We evaluated the results with population genetic and phylogenetic methods, in the context of the modern and ancient populations that are either geographically or historically related to the Székelys. Our results demonstrate a predominantly local uniparental make-up of the population that also indicates limited admixture with neighbouring populations. Phylogenetic analyses confirmed the presumed eastern origin of certain maternal (A, C, D) and paternal (Q, R1a) lineages and, in some cases, they could also be linked to ancient DNA data from Migration Period (5th-9th centuries AD) and Hungarian Conquest Period (10th century AD) populations.

Introduction

The Székelys (also known as Szeklers or Seklers) are a Hungarian-speaking minority that has been living in Transylvania (Romania) for more than 800 years. Several theories have been elaborated about the origin of the Székelys over time, but it is still an unresolved question to this day. They have been identified as descendants of Hunnic, Avar, Kabar, Volga Bulgarian (Onogur) and Hungarian ethnic groups. The story of their Hunnic origin was elaborated by mediaeval Hungarian chroniclers (who, by doing so, increased the authority of the Árpád dynasty and created the legal basis for the Hungarian conquest), so the Székelys' own Hunnic 'tradition' seems to have developed secondarily as a result of this and, due to the lack of evidence, modern historiography and archaeology have not considered the Székelys to be of Hunnic origin for some time [1]. According to some scholars, the ancestors of the Székelys were ethnic groups separated from the Volga Bulgarians, and were thus of Turkish origin. According to this idea, the accession of these Bulgarian tribes to the Hungarians would have taken place even before the Hungarian conquest of the Carpathian Basin in 895 AD. The theory, however, that attempted to connect the Székely folk name with the Askal/Äskäl tribe of the Bulgarians turned out to be linguistically incorrect [2]. There are scholars who consider the Székelys to be the remnants of the late Avar population who, according to the scholars' assumption, spoke the Hungarian language. Other scholars assume that the Székelys were originally Hungarian ethnic groups who guarded the various border sections of the early Hungarian Kingdom, primarily at the western ends, and later, in the 12th–13th centuries, the majority of them were resettled in Transylvania in order to stop the Cuman and later Tatar incursions that threatened the eastern borders [3]. The first written mention of the Székelys originates from the 12th century, mentioning them as military auxiliaries of the Hungarians along the Pechenegs, then still in the western border region.

At the moment, the research faces a serious contradiction. On the one hand, we see a Székely population with its own name and traditions, which, after observing typical military service to the king, may seem like an 'auxiliary people' who joined the Hungarians, whose territorial organisation was not the usual county of the rest of the Hungarians, but the district/*sedes* typical of foreign ethnic groups; the seemingly distant connections of this population towards Asia and the late Avar period have recently been raised by physical anthropological research [4]. On the other hand, their archaeological findings from the Árpád period do not differ from the findings of the rest of the Hungarian population and do not show special oriental features, just as their Hungarian dialect does not indicate a change of language and the subsequent acquisition of the Hungarian language. Their placement in small patches at critical points of the country's border reflects the conscious organisation of the early Árpáadian kings, for the sake of border and land protection. Their significant mediaeval privileges were based on their continuous military service, which made them important actors of the time, and this continued in the early modern period, in the territory of the independent Principality of Transylvania. To this day, their mother tongue is the clearly spoken Hungarian.

A branch of the Hungarian ethnic group known as the Csángó is also related to the Székelys, the Csángós of Ghimeş/Gyimes, who moved from the area of Ciuc/Csík district to the valley

of the Trotuş/Tatros river on the border of Transylvania and Moldavia from the early modern period (or perhaps earlier) and whose language is thus closely related to the Csík dialect [5]. Due to their close relationship with the Székelys, the Gyimes csángós are also analysed in this study from published sources. They should not be confused with other Csángó groups living in other areas of Moldavia and speaking a different dialect, whose origins are also different.

In recent decades, molecular genetics studies have described the genetic make-up of some of the urban Székely (Miercurea Ciuc/Csíkszereda and Corund/Korond) and Ghimeş Csángó groups, investigating maternally inherited mitochondrial DNA (mtDNA) [6–8] and paternal Y-chromosomal [9–11] lineages. Most of these studies lacked thoroughly planned and executed sample collection; thus, we cannot be sure that all sample donors had local ancestors. The studies revealed an increased number of Central or Eastern Asian lineages in the Székely population compared to other Hungarian-speaking populations. In addition to former uniparental studies, genome-wide genotype data from 24 Székely individuals from the commune of Corund were analysed and compared with the genotype data of Hungarians [12]. The genetic research on the Székely population does not currently have databases containing complete mitochondrial genomes, that would have been based on accurate sampling. Both of these have great importance in evaluating genetic continuity between present-day and ancient populations.

Besides Csángós and Székelys, the population of Hungary has also been investigated [13]. Hungarian paternal lineages from Hungary were reported by Völgyi et al. in 2009 and by Pamjav et al. in 2017 [14,15]. In the latter study, all samples were collected in the Bodrogek, which is a geographical area in the Upper-Tisza region in north-eastern Hungary bordered by the Bodrog and Tisza rivers. Due to its isolated nature, the dataset from Bodrogek has been treated separately from the Hungarian data in our study. The authors assumed that its present-day population is likely to preserve ancient markers and lineages, as its former inhabitants had a better chance of surviving both Tatar (Mongol) and Ottoman invasions than groups living in some of the other affected regions [15]. There is scarce genetic data available on the Romanian population—only one mitochondrial DNA sampling has been reported in Romania, and haplogroup results, which are based on mtDNA control region and coding marker data, were made accessible by Cocos et al. in 2017 [16].

In this study, we aimed to reconstruct the uniparental gene pool of the Székely population that existed 100–150 years ago by finding elderly sample donors living in isolated villages and carefully documenting their genealogies. Furthermore, we aimed to monitor any regional genetic structure discrepancies of the Hungarian-speaking population and to confirm preliminary uniparental genetic studies that revealed an increased number of Eastern Eurasian lineages in isolated populations, compared to populations of larger cities nearby. We present new genetic data containing 115 newly sequenced whole mitochondrial genomes, and 92 Y-chromosomal Short Tandem Repeat (STR) haplotypes and haplogroups of a Székely population that has not been sampled before, and compare them to recent Eurasian and available ancient DNA (aDNA) data to gain further knowledge about their genetic history.

Material and Methods

DNA samples, extraction, amplification and sequencing

Samples were collected with buccal swabs by researchers from ELRN RCH Institute of Archaeogenomics, the ELTE University of Budapest and the Sapientia Hungarian University of Transylvania. The samples were taken from 115 mostly unrelated individuals of the Székely population of Transylvania, Romania. The selected individuals spoke Hungarian as their mother tongue and had Hungarian surnames. All sampled individuals agreed and gave their written consent to the anonymous use of their samples in this study. Their ancestors had been documented for two generations, and these ancestors were born in the same region of Transylvania and had declared themselves as Székelys. The following villages in Udvarhely county were included in the sampling, near the town Odorheiu Secuiesc, which has appeared in written records from the 14th century onward [17]: Inlăceni/Énlaka, $n = 9$; Firtănuș/Firtosmartonos, $n = 7$; Ulieș/Kányád, $n = 21$; Mugeni/Bögöz, $n = 13$; Goagiu/Gagy, $n = 11$; Avrâmesti/Szentábrahám, $n = 13$; Cechești/Csekefalva, $n = 9$; Dobeni/Székelydobó, $n = 12$; Văleni/Patakfalva, $n = 7$; Forteni/Farcád $n = 13$ (Figure 1, Supplementary Table S1).

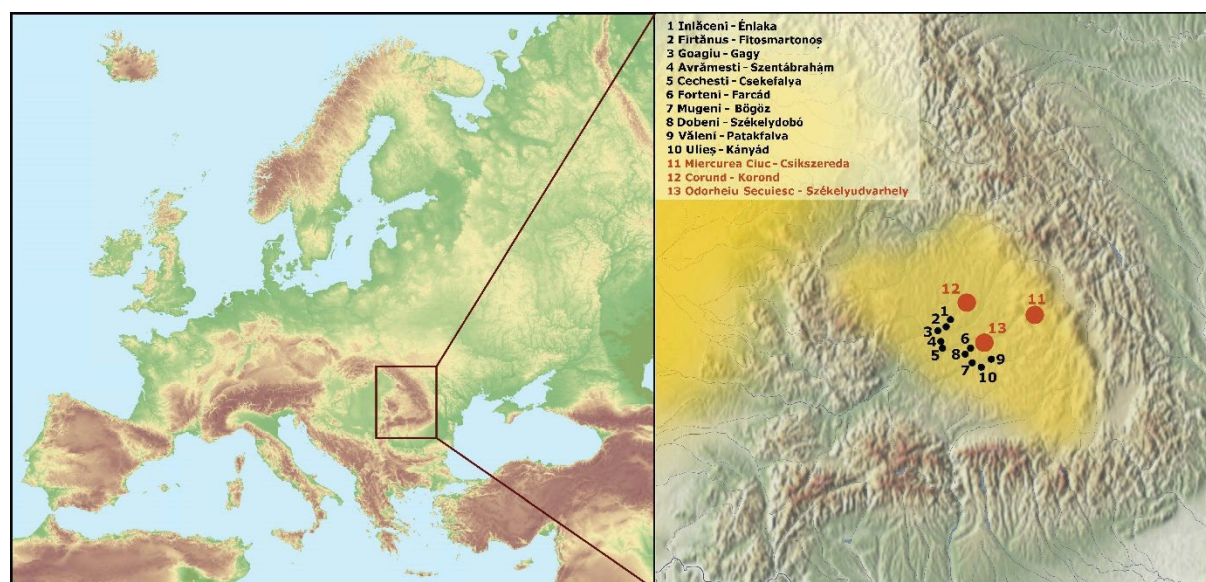


Figure 1. Map of Europe and the Transylvanian part of Romania showing the Székely villages where the DNA samples were collected (in black). The yellow shadings indicate the settlement areas of the Hungarian-speaking populations, including the Székelys. Red circles indicate previously collected and published Székely datasets [6], [8] and the city of Odorheiu Secuiesc. The map of Europe was downloaded from MAPSWIRE [18], licensed under CC BY 4.0. The map of the Carpathian Basin is owned by the Institute of Archaeogenomics, Research Centre for the Humanities, Eötvös Loránd Research Network; modifications were made in Adobe Acrobat Pro DC and Inkscape 1.1.1.

DNA was extracted with QIAamp DNA Mini Kit (Qiagen) according to the producer's buccal swab spin protocol. The concentration of the samples was measured with Qubit™ dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific).

The amplification of the whole mtDNA was performed with the Expand™ Long Range dNTPack kit (Sigma Aldrich) according to Fendt et al. [19] (primer sequence 5'–3', forward 'A' (FA): AAATCTTACCCCGCCTGTTT; reverse 'A' (RA): AATTAGGCTGTGGGTGGTTG; forward 'B' (FB): GCCATACTAGTCTTTGCCGC; reverse 'B' (RB): GGCAGGTCAATTTCACTGGT). We amplified the mtDNA in two fragments, and modified the PCR program according to the length of the fragments. Conditions used for long-range PCR consisted of an initial denaturation step of 2 min at 92°C followed by 10 cycles of denaturation at 92°C for 10 s, annealing at 60°C for 15 s and elongation at 68°C for 8 m 30 s, 10 cycles of denaturation at 92°C for 10 s, annealing at 60°C for 15 s and elongation at 68°C for 8 m 50 s, 15 cycles of denaturation at 92°C for 10 s, annealing at 60°C for 15 s and elongation at 68°C for 9 m 10 s, and a final elongation step at 68°C for 7 min. The amplification reaction was checked on 0.8% agarose gel and visualised after EcoSafe staining with UV transillumination. We pooled the two separately amplified fragments, then purified the amplicons with the QIAquick PCR Purification Kit (Qiagen). The concentration of the PCR products was measured with Qubit™ dsDNA Broad Range Assay Kit (Thermo Fisher Scientific).

NEBNext Ultra II FS DNA Library Prep Kit for Illumina kit was used for the preparation of the mtDNA libraries. The products were checked with the Agilent D1000 ScreenTape Assay on the 4200 TapeStation system. Next-generation sequencing was performed on the Illumina Miseq System (Illumina) using Illumina Miseq Reagent Kit V2 (2 x 150 cycles) sequencing kit. Indexed libraries' final concentrations were adjusted to 4 nM. Samples were pooled together with regard to the calculated desired coverage. Five percent of PhiX was used to increase the heterogeneity of samples.

We analysed 92 male samples in the Laboratory of Reference Sample Analysis of the Department of Genetics, Directorate of Forensic Expertise, Hungarian Institute for Forensic Sciences in Budapest. DNA was surveyed for STR variation using the Promega PowerPlex Y23 for the Székely population, including 23 Y-STR loci. ABI3130 Genetic Analyzer and GeneMapper ID-X v.1.2 software were used for fragment analyses of PCR products. The results of the Y-chromosomal Short Tandem Repeat (STR) analyses were verified by haplogroup-defining Single Nucleotide Polymorphism (SNP) markers (see Supplementary Table S2) on ABI 7500 Real-time PCR instrument using SDS.1.2.3 software.

Pre-processing of the sequencing data

A custom in-house analytic pipeline was applied on the Illumina sequencing data [20]. Paired-end reads were merged together with SeqPrep master [21]. At a maximum of one mismatch, the one base with higher base quality was accepted and the overlapping reads with two or more mismatches were discarded. The pre-processed reads were mapped to the rCRS reference

sequence using BWA v.0.7.5 [22] with MAPQ of 30. Samtools v.1.3.1 [23] was used for further data processing, such as indexing, removing PCR duplications and creating bcf files.

Mitochondrial haplogroup determinations were performed by HaploGrep2 [24], which uses Phylotree mtDNA tree Build 17 [25,26].

Y-haplogroups were assigned based on Y-STR data using nevgen.org, as well as based on Y-SNP genotyping by TaqMan assay on a Real-time PCR platform. Terminal Y-SNPs were verified on the Y tree of ISOGG version 15.34 [27].

We created and visualised the median-joining (MJ) network of the whole mitochondrial genomes of our dataset with the PopArt program [28]. The input file of PopArt was made by DnaSP [29].

Phylogenetic analysis of the mtDNA

We used all publicly available data to compile datasets for each mtDNA haplogroup with sequences that belonged to the same haplotype or were closely related to it. For each phylogenetic tree, 50–150 mtDNA sequences were used. For every haplotype, all available closely related mitochondrial genome sequences in NCBI were downloaded. Multiple alignments for each haplogroup were performed with ClustalO within SeaView [30]. The alignments were checked and corrected manually where necessary. Next, neighbour-joining (NJ) trees were generated by PHYLIP version 3.6 [31]. The phylogenetic trees were then drawn using Figtree version 1.4.2 [32].

Most of the data used for the neighbour-joining mitochondrial phylogenetic trees are from the NCBI database; IDs and sources of other data are available in Supplementary Tables S5.

Population genetic analysis

Principal component analysis was performed based on the mtDNA haplogroup frequencies of 56 modern and two ancient populations. In the PCA of the modern populations, we considered 36 mitochondrial haplogroups. The PCAs were carried out using the prcomp function in R v4.0.0. [33] and visualised in two-dimensional plots with two principal components (PC1 and PC2 or PC1 and PC3).

For a Ward-type hierarchical clustering we involved the same population datasets as for PCAs. Based on the mtDNA haplogroup frequencies, we calculated PC-scores in R v4.0.0 [33], then applied PC1–PC6 scores using the Euclidean distance measurement and ward.D method. We visualised the results as a dendrogram with the hclust library.

To calculate the inter-population variability of the mtDNA genetic profiles characteristic of the three Székely populations, we performed analysis of molecular variance (AMOVA) using Arlequin v3.5.2.2 software [34].

We calculated population pairwise F_{ST} and linearised Slatkin F_{ST} values based on the whole mitochondrial genome sequences of 3981 modern-day individuals (classified into 21 groups) and 362 ancient individuals (classified into 7 groups) using Arlequin v3.5.2.2. [34] with the following settings: Tamura & Nei substitution model with 10,000 permutations, significance level of 0.05, gamma value of 0.3.

We used the same linearised Slatkin F_{ST} values for clustering in Python using the seaborn [35] clustermap function (parameters: metric = 'correlation', method = 'complete').

Analysis of Y-STR variations

MJ networks were constructed using Network v10.1.0.0 and the results were visualised with Network Publisher v2.1.2.5 [36]. The following settings were used in Network v10.1.0.0: Network Calculation: Median-Joining [37], Optional Post-processing: MP calculation [38] (selected option: Network containing all shortest trees, and list of some of the shortest trees sufficient to generate the network) and in Network Publisher, shortest tree visualisation was applied and coloured according to the haplogroups and sample provenance.

Results

The sample pool of this study consisted of 92 male and 23 female participants, all Székely individuals from Transylvania, Romania (for detailed information see Supplementary Table S1). We performed whole mitochondrial genome enrichment and next-generation sequencing (NGS) to obtain 115 complete mitogenomes. In addition, we investigated the Y-STR profiles (23 STRs) and Y-SNP data of the 92 male individuals (see Supplementary Table S2).

Maternal lineages in the dataset

Haplogroup-based analyses

115 high-coverage mitochondrial genomes were obtained with NGS methods (from 111.46x to 276.83x coverage), with a mean coverage of 233.56x. The 115 complete mitochondrial genomes were classified into 72 different haplotypes. These mitochondrial haplotypes are mainly present in European regions, but there were several haplotypes predominantly found in present-day Asian or Near Eastern populations. The new dataset consisted of the following macro haplogroups: A, C, D, H, HV, HV0, I, J, K, T, U, N, R, V, W and X. A list of the mtDNA sub-haplogroups found in the Székely population is in Table 1. The overall haplogroup composition of the investigated Székely population was similar to the formerly described Székely populations in Miercurea Ciuc and Corund, and to the Hungarian population in Hungary as well [7,8]. Around Odorheiu Secuiesc, most of the individuals belonged to haplogroup H—as expected in a European population [39], and as had been observed in the case of earlier studied Székely, Hungarian and Csángó populations (see Figure 2). Compared to the populations of Miercurea Ciuc and Corund, some differences were conspicuous. We observed a higher proportion of haplogroups I, T2, HV and W, and a lower proportion of haplogroup K than previous studies; furthermore, no T1 was present in our dataset. The proportion of the Eastern Eurasian-derived haplogroups was higher in the population of Miercurea Ciuc than around Odorheiu Secuiesc and in Corund. The Csángó population from Ghimeș stood out slightly in the comparison due to its high proportion of the haplogroup K and lower proportion of H [7]. All three Székely populations had a significant proportion of mitochondrial haplogroups with Eastern Eurasian prevalence (groups A, B, C, D, G and Y).

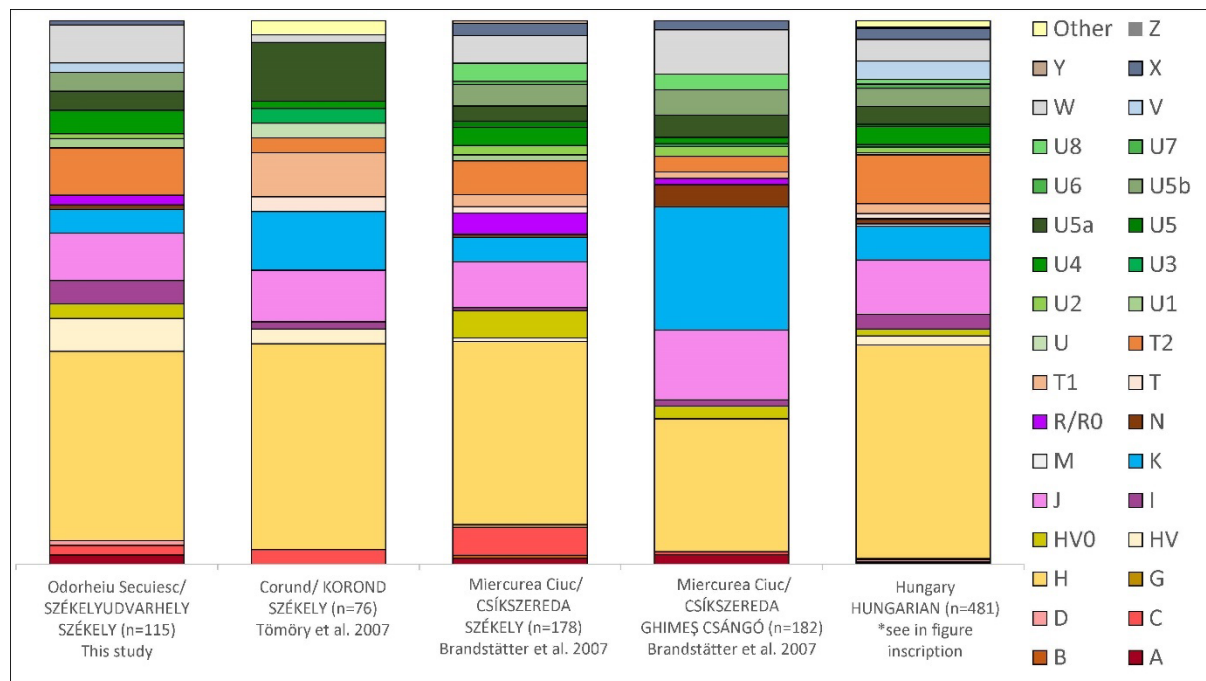


Figure 2. Haplogroup composition of the investigated Székely population around Odorheiu Secuiesc (n = 115) compared to other Székely populations [7,8], the Csángó population [7] and Hungarians living in Hungary (n = 481).

Observed mtDNA haplogroup in this study	n	Frequency in the Székely population	Modern-day presence/geographic distribution peaks of the haplogroups
A+152+16362	1	0.87%	Eastern Asia (Han Chinese, Japanese)
A12a	1	0.87%	East-Asia (Evenk, Yakut, Mansi)
C4a1a3	1	0.87%	Northeastern-Asia, Central Siberia (Even, Evenk, Yukaghir, Yakut, Altai-Kizhi, Buryat)
C5c1a	1	0.87%	Middle East (Iranian), Europe
D4e4	1	0.87%	Northeastern-Asia, Central Siberia (Tungusic, Evenk, Pamir, Uyghur, Buryat)
H and subclades (H, H1, H1a3b, H1b1a, H1c, H1c1b, H1c20, H1q, H1u2, H24a, H26a1, H2a1a, H2a3, H3, H3h, H4a1, H5, H5'36, H55, H5a1, H6a1a, H11a1, H13a1c)	39	33.91%	Europe
HV, HV0 and subclades (HV+16311, HV6, HV11, HV0a1, HV0f)	10	8.70%	Caucasus (Armanians), Middle East (Turks, Iranians), Europe
I2	1	0.87%	Northern Europe (English, Irish, Norwegian), Caucasus
I5 and subclades (I5a1, I5a4)	4	3.48%	Europe, Middle East (Turks)
J and subclades (J1b1a1, J1c2, J1c2p, J1c10, J2b1)	10	8.70%	Europe (Finnish, English), Middle East (Iranian)
K and subclades (K1a+195, K1a2, K1b2a2, K1c1a, K2b1a4)	5	4.35%	Europe, Eastern Europe (Russian), Baikal Region, Caucasus
N1b1a	1	0.87%	Caucasus, Middle East, Eastern Europe
R0a2	3	2.61%	Middle East (Iranians), Arabian Peninsula, Horn of Africa
T, T2 and subclades (T2a1b1ba1, T2b, T2b25, T2b4, T2f1a1, T2g1a)	10	8.70%	Europe, Caucasus, Central Siberia, Middle East
U subclades (U1a1a1, U1a1a1a, U2e1a1, U4a, U4a2, U5a1a1, U5a1b1a, U5b1a, U5b2b1b)	16	13.91%	Middle East (Iranian), Europe, Russia
V and subclade (V, V1a1b)	2	1.74%	Europe (Finnish), Eastern Europe (Volga Tatars)

W and subclades (W, W1e1a, W1i, W3a1d, W5a1a)	8	6.96%	Europe, Europe, North-Caucasus, Central-Asia, India, Iran
X2e1	1	0.87%	Europe

Table 1. Modern-day distribution peaks of the mtDNA sub-haplogroups found in the Székely population based on the IanLogan database [40]. Column ‘n’ gives the absolute frequency of the given haplogroup in the studied Székely population around Odorheiu Secuiesc/Székelyudvarhely.

We used Principal Component Analysis (PCA) in order to comparatively visualise the mtDNA haplogroup composition of the different populations (see Supplementary Table S3). The investigated Székely population (Odorheiu Secuiesc) was positioned among European populations, closest to other Székelys, Csángós, Croatians, Bosnians, modern Czech populations and Transylvanian Romanians. It was not possible to further examine all connections at the complete mitogenome sequence level due to the lack of whole mitogenome data in some populations.

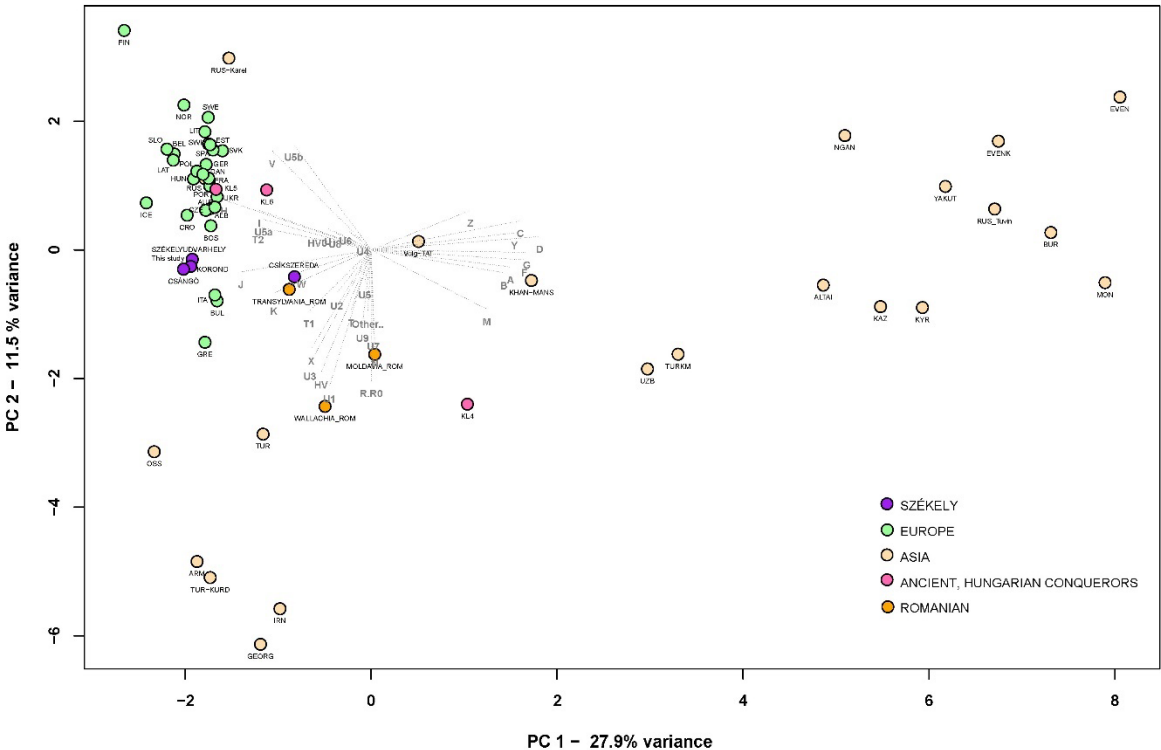


Figure 3. PCA plot with 56 modern and three ancient populations (36,803 samples), representing first and second principal components (39.4% of the total variance): PCA analysis based on haplogroup frequencies in Eurasian modern populations and three ancient populations. The selected ancient populations are the Hungarian Conquest Period (10th century AD) populations of the Carpathian Basin [41–44]. KL4-5-6 groups indicate different cemetery types in the Hungarian Conquest Period, as used in Szeifert et al. 2022 [45]. The investigated Székely population and previously examined Székely

groups are marked in purple, the ancient populations from Hungary in pink, modern-day Romanian populations in orange, other modern-day Europeans in green and Asian populations in beige. The PCA shows a clear separation of Eastern (right side of the plot) and Western (left side of the plot) populations. For further information see Supplementary Table S3.

Data on the PCA were also displayed using the hierarchical ward-clustering method (see Supplementary Information Figure S1). The clustering confirms the connection of the studied Székely group with Europeans, but also separates the Miercurea Ciuc, Ghimeş Csángó and Corund groups from Odorheiu Secuiesc. This difference from the PC1-2 plot is also visible on the PC3, where the Odorheiu Secuiesc group becomes distant from the others (Supplementary Information Figure S2).

Since the other two Székely groups were only analysed for Hypervariable Region I of the mitochondrial genome, the sequence-based comparison of the three groups with limited conclusions is discussed in the Supplementary Information.

Whole mitogenome sequence-based evaluations

F_{ST} analyses

We analysed the whole mitogenomes (16,569 base pairs) at the DNA sequence level, and calculated Slatkin F_{ST} values (see Supplementary Table S4). A heatmap with clustering of F_{ST} values was created to visualise the genetic differentiation of the examined populations (Supplementary Information Figure S3). The Székelys cluster on the European branch with Hungarians, but the Serbians and the Conquest Period Hungarians are the most similar to them.

Whole mitochondrial data are missing from Romania, and the Slovakian and Czech datasets are also limited; therefore, the resolution of that analysis is restricted. Among the ancient populations, the KL6 group, which was discussed by Szeifert et al. as comprising large village cemeteries opened in the 10th century and used until the 11th and 12th centuries in the Hungarian Kingdom [45], was the closest to the Székelys.

Phylogenetic analyses of the Székely maternal lineages

The median-joining network of the Székely mitogenomes showed the variable distribution of the maternal lineages among the sampled villages (see Figure 4). Most of the haplogroups were shared among the villages, and shared lineages (haplotypes) were also found among certain H, T, U4, U5 and W haplotypes.

Three individuals belonging to the described Eastern Eurasian haplogroups originated from the same village (Goagiu), although we detected Eastern Eurasian haplogroup types in the Avrămeşti and Inlăceni villages as well.

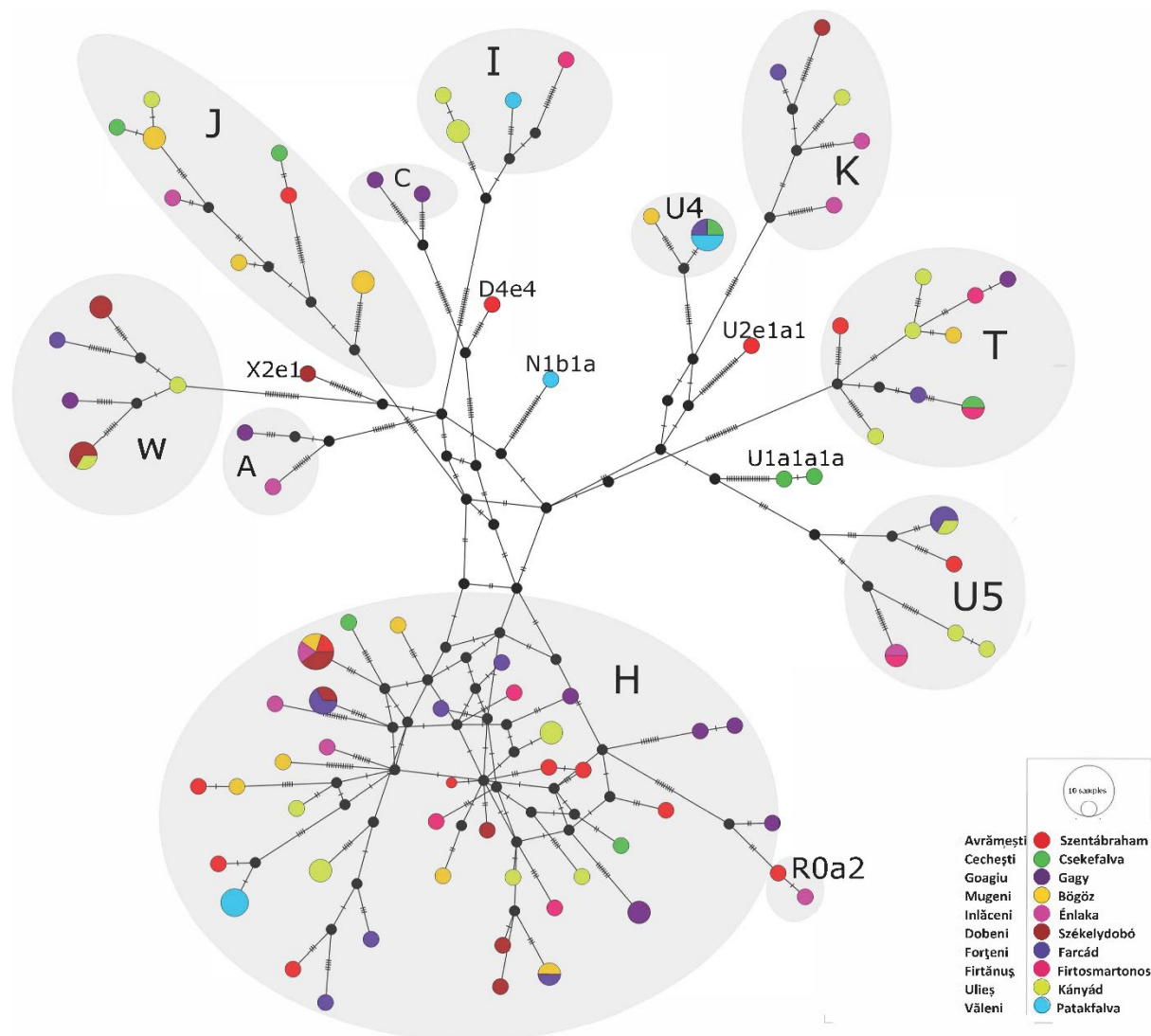


Figure 4. Median-joining network of 115 modern-day Székely mitogenomes. The sequences contained 484 variable sites and belonged to 72 haplotypes. The figure was created with the PopArt program.

Since analyses of mitogenomes in pools did not lead to differentiation of distant present-day European populations, we investigated individual maternal lineages in the following, in order to monitor the connections of the Székely maternal lineages to early Hungarian populations, among others.

On the A12a phylogenetic tree (Figure 5), a modern-day Hungarian sample and a Hungarian sample from the time of the Hungarian conquest (10th century, Harta_HC3), as well as two samples from the 9th–10th century Volga-Ural region (Bolshie Tigani RC8 and Uyelgi-No7, [42,45]), cluster together with the examined modern-day Székely sample. The Conquest Period and the Bolshie Tigani individuals had identical mitochondrial DNA sequences to the Székely individual and that lineage could also be found among today's Hungarian population.

Based on this tree, we assume that the phylogenetic lineage A12a came from the Volga region, and was also present at the time of the Hungarian conquest (late 9th–10th century). The newly reported samples within the A12a subgroup caused some changes in the nomenclature within

the A12a tree that we present in the Supplementary Information Figure S4. The Székely sample described here has been ordered to a new subgroup named A12a2b.

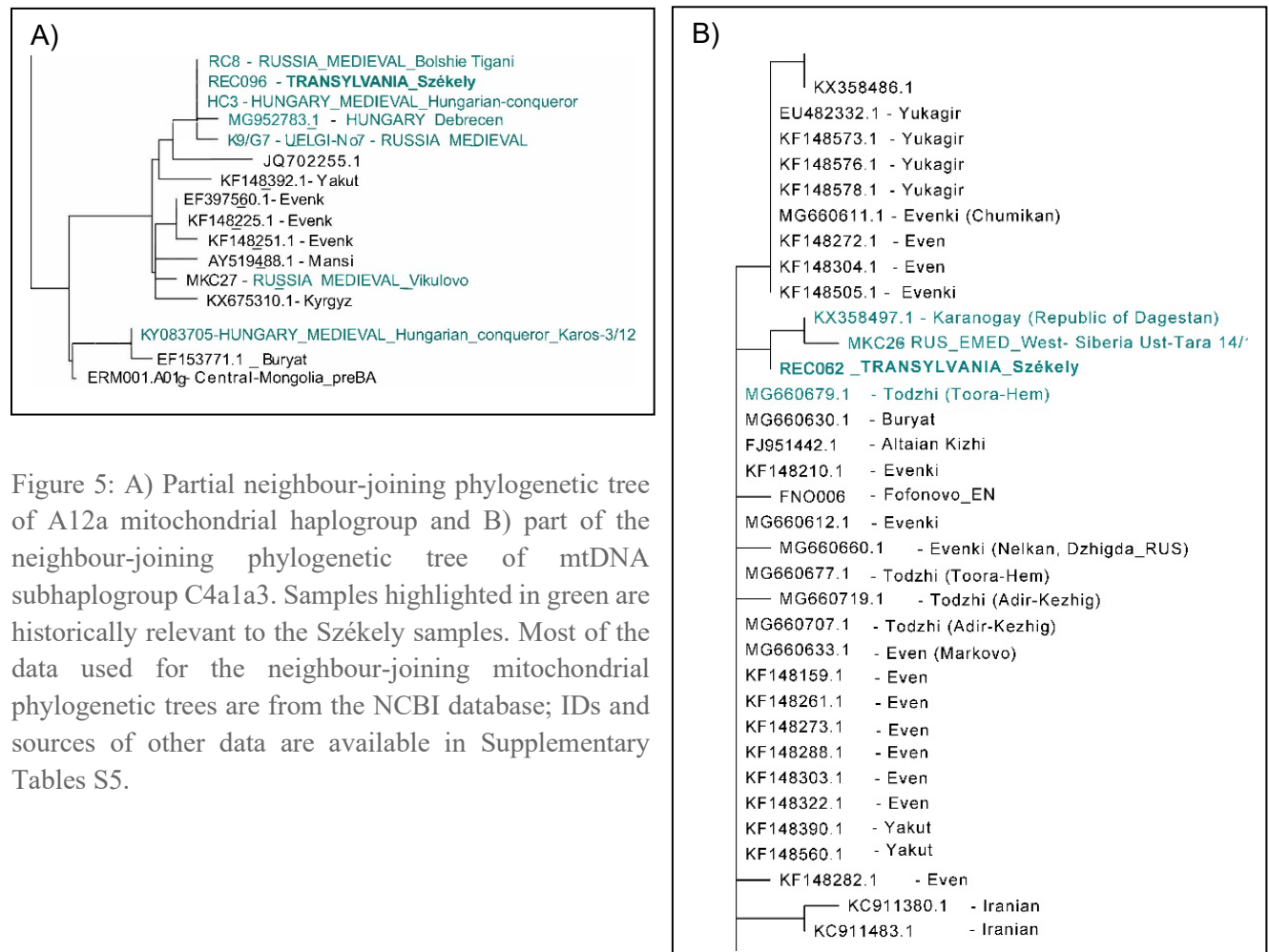


Figure 5: A) Partial neighbour-joining phylogenetic tree of A12a mitochondrial haplogroup and B) part of the neighbour-joining phylogenetic tree of mtDNA subhaplogroup C4a1a3. Samples highlighted in green are historically relevant to the Székely samples. Most of the data used for the neighbour-joining mitochondrial phylogenetic trees are from the NCBI database; IDs and sources of other data are available in Supplementary Tables S5.

On the C4a1a3 partial neighbour-joining phylogenetic tree (Figure 5B) the MKC26 refers to a sample that originated from the 6th–8th century West-Siberian Ust-Tara archaeological site of the Nizhneobskaya culture [45]. The population of this culture was probably proto-Ob-Ugric (Southern-Khanty), although it showed typical Hun-period cultural traits [46]. The other sample which shares a branch with the Székely sample originated from the Karanogay ethnicity (Turkic ethnic group), Dagestan, and the adjacent Todzhi sample was also from a Turkic ethnic group, being a group of Tuvans. This tree represents the mixed nature of the C4a1a3 lineage, which despite its prevalence in Turkic-speaking ethnic groups, may also have originated from Western Siberia in the Székely gene pool. Nevertheless, we do not have immediate proof of that hypothesis in the form of linking lineages from the Volga-Ural region.

Parts of the neighbour-joining phylogenetic tree of A+152+16362 and D4e4 mitochondrial haplogroups are shown in Supplementary Figures S5 and S6. On the A+152+16362 tree, a sample from Cis-Uralic Sukhoy Log cemetery (7th–8th centuries) [45], as well with the latter

contemporaneous, ‘Kimak’-reported individual from the Central Steppe [47], can be found in close proximity to the Székely sample. The relationships between samples from the Volga region Early Mediaeval sites Karanayevo, Bolshie Tigani, Gulyukovo, Tankeevka, from the Conquest Period Transdanubia and the Székely sample on the D4e4 phylogenetic tree suggest a possible connection between these sublineages via the conquerors. However, the Székely lineage has a basal character, and is identical to a lineage detected in the Bronze Age of Bolshoy Oleny island in Kola Bay. It is therefore possible that D4e4 is an originally Northeastern European maternal lineage that reached the Carpathian Basin via a different migration.

Paternal lineages in the new Székely dataset

The population genetic investigation of non-recombining Y-chromosomal markers like Y-STRs and Y-SNPs can be used to trace back paternal lines in time and describe phylogeographic structures and diversities of populations.

Haplogroup-based analyses

Y haplotypes from 23 Y-STR markers were obtained from 92 men out of the sampled 115 individuals. The haplogroup predictions were confirmed by selective Y-SNP typing (see methods and Supplementary Table S2). In this dataset, we found eight Y macrohaplogroups (E, G, H, I, J, Q, R and T), which included 21 different subhaplogroups based on SNP typing. Some of these Y subhaplogroups are predominantly found among Inner Asian (R1a-Z93) and South Asian/European Roma (H-M52), as well Northern Eurasian (Q-M242) people (in a total of 7 samples, 7.6% of all samples) the other 18 subhaplotypes are mostly referred to as European-derived types (Table 2).

Although some studies on the Székely male populations have been published previously, their comparability with our dataset is rather limited. In 2005, Egyed et al. studied 257 Székely individuals from Miercurea Ciuc, including 89 males typed for 12 Y-STR haplotype loci. In Csányi’s study from 2008, 13 Y haplogroups were determined in the Székely population from Corund [48]. The Y haplogroup diversity was 0.9157 in the latter Székely population, 0.9011 in the Székelys living in Miercurea Ciuc [7] and 0.8636 among the Hungarian male population in Hungary [14]. The Y-STR haplotype diversity in our studied Székely population was 0.9995, and the proportion of unique haplotypes was 97.8% using the PowerPlex Y23 System. Based on an investigation of 72 European populations comprising a total of 12,000 samples, the average haplotype diversity was higher ($H_d = 0.999992$) using the same Y23 System [49] than in the Székelys.

In 2015, Bíró et al. studied Székely haplogroups from Miercurea Ciuc (haplotypes published by Egyed et al.) [9,11] that had proven Central and Inner Asian genetic contributions: in their dataset, the possible maximum Central/Inner Asian admixture among the Székely male population was 7.4% [9,11]. In our results, this proportion was a comparable 7.6% in the population around Odorheiu Secuiesc. According to Bíró’s study of contemporary Hungarians from Hungary, this Central/Inner Asian admixture was estimated as only 5.1%, and 6.3% in

the Csángó male population from Ghimeș. Examining the Bodrogek Hungarian dataset, these Asian-derived haplogroups appeared in 6.9% of the population.

Y chromosome subhaplogroup observed in this study	n	Haplogroup frequency in this study	Modern-day geographic distribution of the haplogroup
E-M123	2	2.17%	Saudi Arabia, Yemen, Europe, Middle East, North Africa - TMRCA 17,700 ybp
E-M78	6	6.52%	Europe (Mediterranean region), Saudi Arabia, Somalia, Morocco - TMRCA 13,200 ybp
G2-L156 (G2a-P15)	4	4.35%	Russia, Saudi Arabia, Turkey, Caucasus - TMRCA 18,200 ybp
H-M52	1	1.09%	India, Pakistan, European Romany populations- TMRCA 15,900 ybp
I1-L22	15	16.30%	Europe (Sweden, Finland, United Kingdom), North-America - TMRCA 3,900 ybp
I1-M253	3	3.26%	Europe (Sweden, United Kingdom, Finland, Norway) - TMRCA 4,600 ybp
I2a-P37	18	19.57%	Eastern Europe (a typical paternal lineage of Mesolithic Europe), Russia - TMRCA 18,400 ybp
I2b-M223	2	2.17%	Europe (Germany, United Kingdom, Sweden) - TMRCA 14,600 ybp
J2 (M172, L24)	2	2.17%	Asia, Saudi Arabia, Caucasus, India, Mediterranean, Middle East - TMRCA 27,600 ybp
J2a-M67	7	7.61%	Russia, Caucasus, Arabian Peninsula, Mediterranean (Italy) - TMRCA 12,100 ybp
J2b-M12	3	3.26%	Europe (Italy, Germany, United Kingdom), Russia (Volga-Ural region), India, Anatolia, Yemen, Saudi Arabia- TMRCA 15,800 ybp
Q-M242	4	4.35%	Sweden, Americas, Asia (Russia, China, Mongolia) - TMRCA 28,700 ybp
R1a-M458	2	2.17%	Poland, Baltic region, Ukraine, China - TMRCA 4,800 ybp
R1a-Z280	6	6.52%	Asia, (Russia), Poland, Ukraine, Baltic region - TMRCA 4,600 ybp
R1a-Z93	2	2.17%	Asia (Russia, China), India, Saudi Arabia, Afghanistan, Pakistan, Kyrgyzstan, Turkey - TMRCA 4,600 ybp
R1b-M343*	2	2.17%	Europe, North-America - TMRCA 20,400 ybp

R1b-M412	1	1.09%	Europe (United Kingdom, Ireland, Germany, Portugal, France, Sweden) - TMRCA 5,700 ybp
R1b-P312	5	5.43%	Europe (United Kingdom, France, Spain, Italy, Sweden) - TMRCA 4500 ybp
R1b-U106	5	5.43%	Europe (Germany, Sweden, United Kingdom) - TMRCA 4600 ybp
T-M70	2	2.17%	Syria, Iraq, Saudi Arabia, Mediterranean (Italy) - TMRCA 15900 ybp

Table 2. Modern-day distribution peaks of the Y-chromosomal subhaplogroups were found in the Székely male population and TMRCA times (given in years before present), based on the Yfull database [50].

The comparative analyses of the Y haplogroups with other Székely populations showed some level of diversity among the Székely groups. However, while N1a [42,45,51,52] occurred among the Székelys in Miercurea Ciuc, the population of the Odorheiu Secuiesc region did not yield such a signal in our analysis, but resulted in higher proportions of haplogroups Q, I and I2a than in the previously studied Székelys (Figure 6).

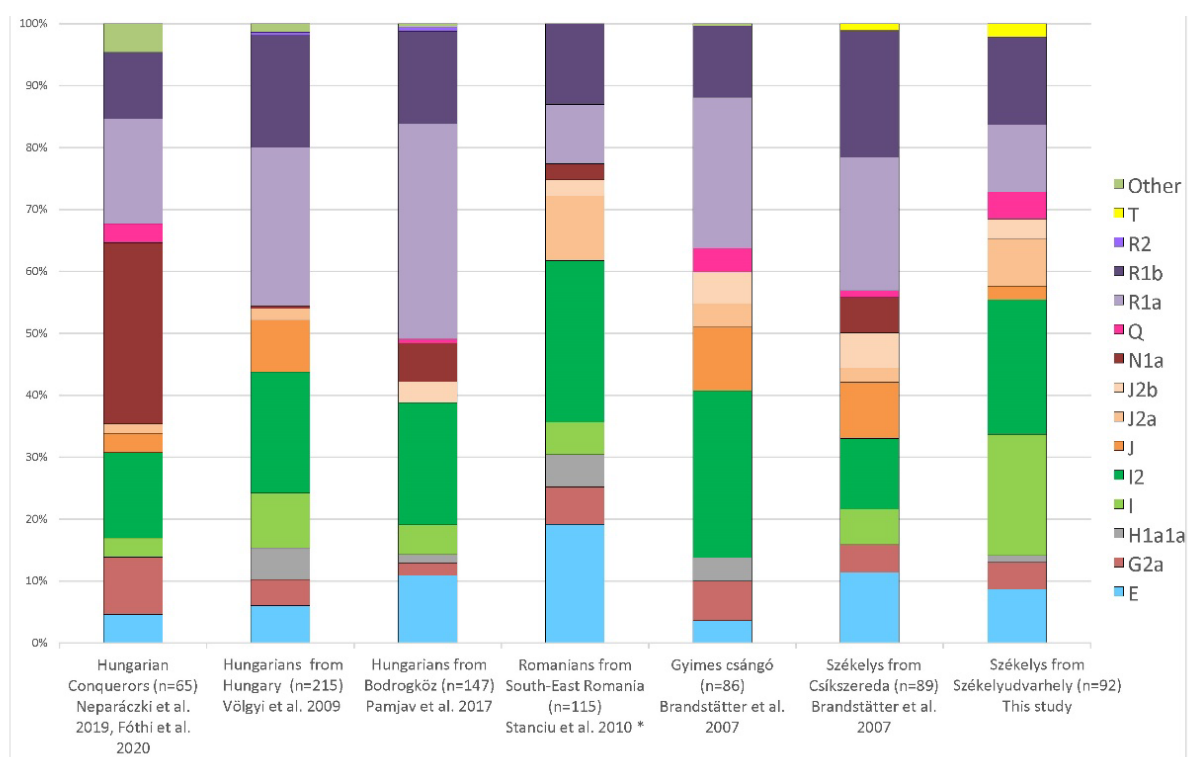


Figure 6. Diagram of the Y haplogroups in Székely, Ghimeş Csángó, Hungarian and Romanian populations. *Haplogroups from the Romanian population were predicted from 17 STR data using nevgen.org.

The haplogroup R1a comprised a substantially higher proportion of the haplogroups among the Hungarian populations living in Hungary than among either the Székely population of the

present study, or the Ghimeş Csángó and Romanian populations. In the Székely population we studied, R1a and R1b comprised only 25% of all haplogroups, while in the case of Hungarians in Hungary, they accounted for 45–50%. However, the ratio of I and J2 haplogroups was higher in the studied Székelys than in the Hungarians. Furthermore, the frequency of haplogroup J was higher in the Székely population of Miercurea Ciuc than in Odorheiu Secuiesc.

Székelys in Miercurea Ciuc showed roughly similar frequencies of G2a and E haplogroups to Székelys around Odorheiu Secuiesc; furthermore, the T haplogroup did not appear elsewhere in the comparison except for in the two Székely groups. T-M70 seems to have originated from the Fertile Crescent and possibly arrived in Europe in the Neolithic with the first farmers [53,54]; today it shows the highest frequency in East Africa and the Middle East. The Székely T-M70 samples belonged to the T1a2b1 subhaplogroup, which is rarely found in ancient data; this was made even more interesting given that it was found in the Hungarian Conquest Period horizon of the Western Hungarian Vörs-Papkert cemetery [55,56].

The I2a1a (I2a-P37) haplotype occurred in the highest number in the studied Székelys, comprising 20% of the total haplotypes. In Hungary, 16.74% of men carry this haplotype [14]. I2a-P37 and subgroups occur at high frequencies in Sardinia (38.9% [57]), and are also present at high frequency among Balkan populations [58]. The proportion of the I2a1a Y type in the Romanian population is 17.7%.

Another dominant haplogroup in our dataset is I1a1b1 (I1-L22), which is most frequent in Sweden and Finland, and represents a fairly large Nordic branch of I1. It was dispersed by the Vikings and nowadays can be frequently found in the Baltics, Britain, Poland and Russia [59].

The distribution of Y haplogroups between the villages did not show a characteristic pattern or patrilineal system; the observed haplogroups were mixed within the region, as is demonstrated in Supplementary Figure S7.

Y chromosome phylogenetic analyses

Our dataset contained four samples that belonged to the Y-chromosomal haplogroup G2a, which we analysed in further detail. In the absence of comparative data covering 23 STR, the network in Figure 7 is based on 17 STR. Two of the Székely G2a Y chromosomes clustered with the M406 subgroup of G2a, with individuals sampled in Tyrol, Austria [60]. This terminal SNP defines the Y-chromosomal subgroup G2a2b1 (ISOGG 2020 v15.73) [27], whereas some of the Székely samples most probably belong to G2a2b2a1a1b based on the L497 marker (ISOGG 2020 v15.73). The G-L497 sub-haplogroup likely originated from Central Europe and has been mostly prevalent in European populations since the Neolithic period [54][61]. The G-L497 lineage could potentially be associated with the *Linearbandkeramik* (LBK) culture of Central Europe. The G-M406 sub-cluster is most concentrated in Cappadocia and Anatolia in Turkey nowadays [62], and has been present in that area since the early Neolithic [63,64].

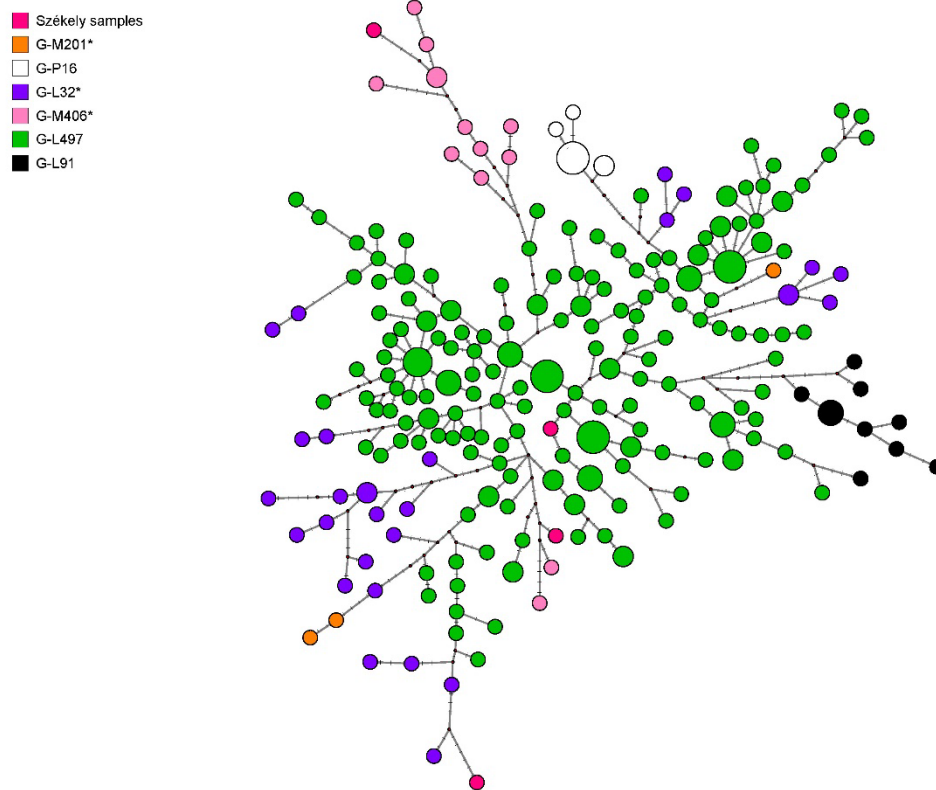


Figure 7. G2a median-joining network: The network was constructed based on 17 Y-STRs, the data used for the network are all from the Tyrolean region of Austria [60]. Circles represent distinct haplotypes, the size of the circles is proportional to the haplotype frequency (the smallest circle corresponds to one individual), the colouring corresponds to haplogroup subclades where the Székelys are highlighted.

Our dataset comprised six samples classified into the Y chromosome R1a1a1b1a2 (R1a-Z280) haplogroup, which we visualised on a median-joining network (Figure 8). For comparison, we collected Y-STR data from the Family Tree Y-DNA database R1a page and we filtered the samples for Y4459 SNP based on nevgen.org prediction. In addition, we included data with the same haplogroup classification from Hungary, Bodrogek [15] and three early mediaeval samples from the Volga-Ural region (Novo Hozyatovo, Gulyukovo—Chiyalik culture), which may have incorporated Hungarians who remained in the East after the Hungarians Westward migration [65][45].

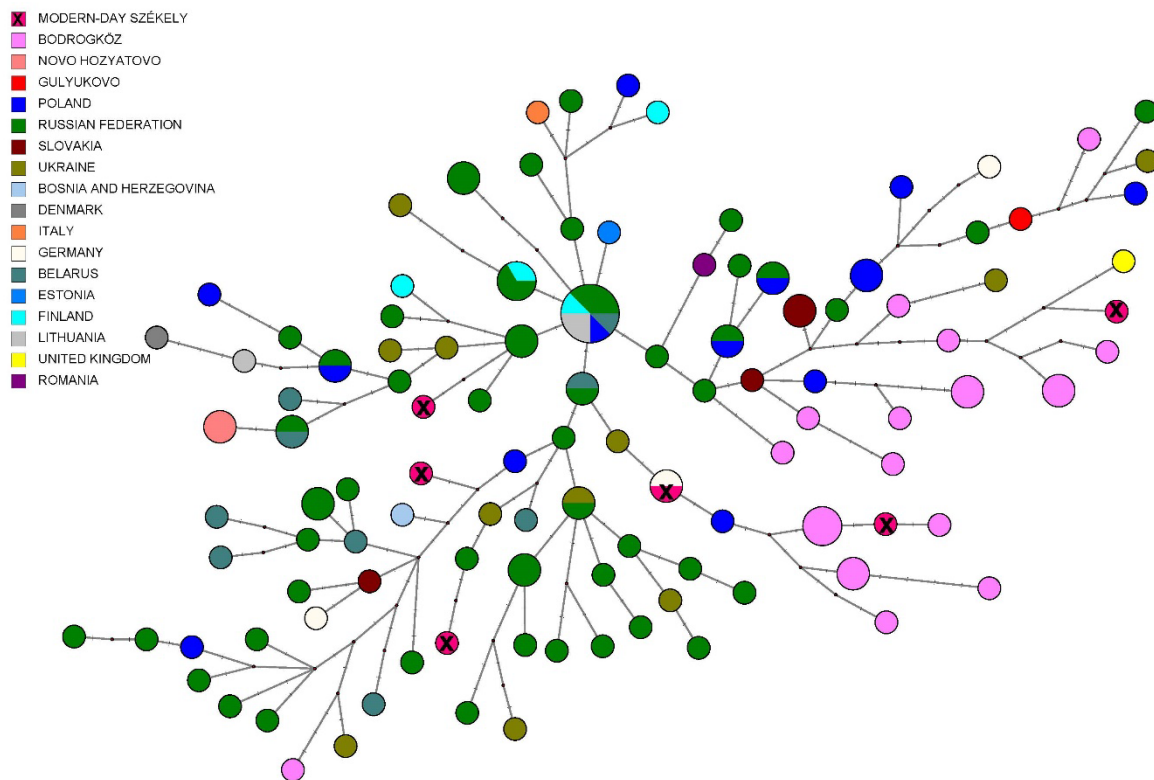


Figure 8. R1a-Y4459 median-joining network. The analysis was performed based on 15 STRs. No larger clusters are seen on the figure, but close relationships can be detected between the samples. Near the Székely samples, individuals from the Bodrogeköz region, Russia, Germany and Poland can be found.

We present further median-joining networks in the Supplementary Data, Figures S8-S10, for haplogroups Q-M242, R1a1a1b1a1a (R1a-M458) and R1a1a1b2 (R1a-Z93). The network of Q-M242 placed the present Székely samples among the Bukovina-Székelys, whereas the two R1a networks did not show geographically relevant patterns.

There were four samples in our dataset that could be classified into the Y-chromosomal macrohaplogroup Q, three of which belonged to the Q1a-F1096—probably to subhaplogroups Q1a2-M25 and Q1a2a1-L715—and one which belonged to the subgroup Q1b1a3-L330; these subgroups are interesting, from our perspective, due to their Central Asian origin (Supplementary Figure S8). Q1a2-M25 is known from the second half of the 5th century AD near Sângeorgiu de Mureş, Romania [66]. Based on the discovered grave goods of the buried man at this site, and his Asian cranial features as well as artificial cranial deformation, strong Hun period traditions have been pointed out [66]. The Q1a2-M25 lineage was also demonstrably present in the Carpathian Basin in the first half of the 7th century AD from a richly furnished, high-status Avar horseman warrior's grave in the Transztisza region, belonging to subhaplogroup Q1a2-M25 [67,68]. Ancient individuals with the same Y-chromosomal haplogroup are known from the Early Middle Bronze Age Okunevo and from

the Baikal Early Bronze Age (Shamanka and Ust’Ida sites), and the Tian Shan Hunnic [47] and Hungarian Sarmatian cultural context [68] as well.

The Q1b1a3-L330 subhaplogroup was also present from the middle third of the 7th century AD in the Carpathian Basin; it was identified from a richly furnished early Avar grave [67,68], and according to a median-joining network (Figure S5 in [67]), this male had a probable Altaian or South Siberian (Tuvanian) paternal genetic origin. The Q1b1a3-L330 lineage was present in the Proto-Ob-Ugric group (Ust’-Ishim culture—Ivanov Mis and Panovo sites), which corresponds to its Altai or Siberian origin [45,69]. The genetic imprint of the Avars in the Székely population can have multiple origins, as their 8th-century settlements were scattered throughout the central Carpathian Basin and along the Maros River in Transylvania, and part of them probably persisted after the Avar–Frankish wars as well [1].

The R1a-M458 median-joining tree shows notable separation of the R1a-M458 types within the Hungarian Bodrogeköz population, and the connection of the Székely lineages to some parts of it. On the R1a-Z93 median-joining tree Hungarian King Béla III and other skeletal remains originating from the Royal Basilica of Székesfehérvár [70,71] show a great genetic distance from the Székely samples, just like the Bashkirian Mari males (Figure S10).

Examining the R_{ST} values based on Y-STR profiles (see Supplementary Table S6, Supplementary Information Figure S11), the closest population to the Székelys was the Slovenian group, with non-significant R_{ST} p -value; the populations with significant R_{ST} p -values were the Greek, Hungarian, Hungarians from Bodrogeköz, Serbian and Croatian datasets (Figure 9). All these reflected a strongly Southeastern European-funded base population of the Székelys with a limited proportion of surviving eastern elements.

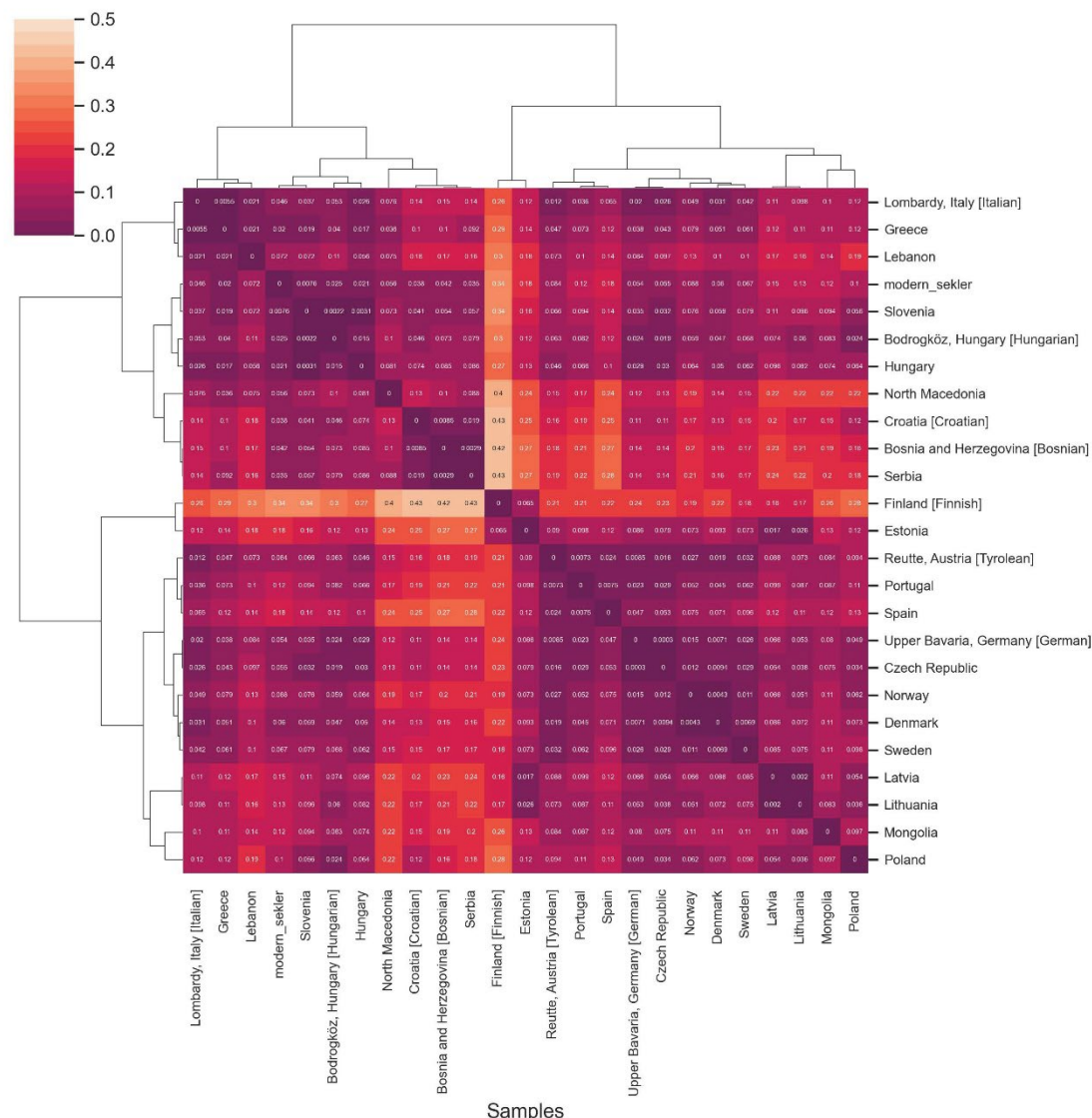


Figure 9. Heatmap of pairwise R_{ST} values with clustering applied for the modern Székely group and populations from Europe (colour scale ranging from yellow to dark purple). We calculated it in Python using the seaborn clustermap function with parameters ‘correlation’ distance metric and ‘complete linkage’ method [35].

Comparison of paternal lines with ancient data

Most of the paternal lines of the early mediaeval sites in the Western-Siberian and Volga-Kama regions—which are suspectedly linked to Hungarian ethnogenesis—belong to N1a1-M46, Q1b- M346-L330, G2a2b-M406, R1a and J2a1-Z6046 [45].

According to the publications of Fóthi et al. [52], Neparáczi et al. [51] and Csányi et al. [48], the N1a1-M46, R1b-U106 and I2a-M170 lines were the most widespread among the conquering Hungarians they examined. This suggests that the conquerors were of diverse

origins and, while N1a1-M46 originated from the east, the R1b-U106 (R1b1a1b1a1a1) lineage is known from the Late Copper Age/Early Bronze Age transition in Europe [72] and is most prevalent in Germanic-speaking people nowadays [73].

These observations fit to the genomic data published by Maróti et al. [56]; among the group they define as the ‘conquering elite’, the haplogroups N1a1a1a1a2a1c (Y13850), N1a1a1a1a4 (M2128), I2a1a2b1a1a (YP189) and R1a1a1b2a (Z94) can be found in notable frequencies.

According to all previous studies, the N1a1 line is the most characteristic of the Hungarian conquerors—almost 30% of the lines belong to the N1a group, and also appear with lower frequencies (4–6%) in the Bodroghöz and Miercurea Ciuc populations; however, this line is completely missing from our new Odorheiu Secuiesc region dataset. The second most widespread haplogroup among the Hungarian conquerors is the R1a-M198 (16.9% of all haplogroups) which is quite common nowadays in the Odorheiu Secuiesc region (10.9%). R1b and I2 haplogroups are present in a relatively higher proportion both in the conqueror (10.8% and 13.8%, respectively) and Székely groups (14.1% and 21.7%, respectively). Most of the other haplogroups listed above—like R1b-U106, G2a and Q1b- M346-L330—can also be found among the Székelys, at a maximum frequency of 5%. The direct comparison of the data is limited, however, by the numerous allelic dropouts in the ancient STR analysis [52], and by the different levels of haplogroup resolutions obtained.

Discussion

In this study, we presented the maternal and paternal genetic composition of a Székely group, a Hungarian-speaking minority living around the city of Odorheiu Secuiesc in Transylvania (Romania). We carefully selected 115 sample providers with local ancestors inhabiting small villages in the area. Altogether, 115 complete, high-coverage mitochondrial genomes were produced with next-generation sequencing methods, which revealed 89 unique haplotypes and could be classified into 72 different subhaplogroups. These mitochondrial haplogroups are mainly present in European regions, but there are also some Asian- and Near Eastern-derived lineages, like A+152+16362, A12a, C4a1a3, C5c1a and D4e4.

In this new Székely dataset, the discovery of an Asian maternal lineage (A12a) completely identical to that found in a male with typical Hungarian conqueror artefacts from the 10th-century cemetery in Harta, Hungary [74], and in the early Hungarian cemetery of Bolshie Tygani in the Volga-region, is a robust sign that some lineages in the Székely population are shared with Hungarian conquerors and are thus most probably of common origin.

The 92 paternal lines investigated in the dataset were mainly composed of European haplogroups, but some lineages (I1-L22, T-M70, J2a-M67) stood out or showed a different distribution in their proportions than in other surrounding Székely ethnic groups. The performed Y-STR networks allowed detailed observations on paternal lineages G2a-L156, R1a-Z280, Q-M242, R1a-M458 and R1a-Z93.

We detected a strongly Southeastern European base population of the Székelys. The genetic proximity of Balkan populations may be a consequence of the formerly inhabited areas of the Székelys in southern Hungary. Presumably, the signs of Glagolitic and Cyrillic origin included in their own old script are also connected with the Eastern European and Balkan connections of the Székelys. This alphabet was probably developed in the Carpathian Basin during the 10th century—using certain signs of the Turkish runic script too—and became suitable for writing short texts in Hungarian. It was used exclusively by the Székelys [75].

The Hunnic origin of the Székelys remains questionable in the light of the present genetic data, because we have scarce genetic data from the Hunnic period [51,68]; furthermore, the population living in the Carpathian Basin during the Hunnic period and in the Avar period (before the time of the Hungarian conquest) shows large heterogeneity [68]. Here we demonstrated connections of the Székelys to the 5th-7th centuries' population through Y-chromosomal Q Asian lineages, which, however, could have arrived repeatedly in the region in numerous epochs. The possible separation of the different immigrant waves requires larger comparative databases from the early mediaeval and mediaeval periods.

We found large among-village heterogeneity in both the maternal and paternal gene pool. The current Székely dataset completes the previous studies and is broadly in line with their observations. The genetic connections between the Székelys and Hungarians could be detected based on our uniparental genetic data using allele frequency analyses; moreover, genome-wide,

haplotype-based methods [12] confirm our observations. The results of this genome-wide, autosomal marker-based analysis show that Székelys from Corund and non-Székely Hungarians from Transylvania are strongly related to contemporary Hungarians living in Hungary [12].

Compared to previous uniparental studies involving Székelys [6–9,11,48], what is different and novel in our research is the sampling method, the selection of the participants and the careful documentation of the ancestors. The results of our research and the reported data are definitely a qualitative leap, considering that so far complete mitochondrial DNA data have not been available from the region and Y-chromosomal data containing 23 STRs have not been previously reported.

Besides revealing present-day diversity, it is of great importance to evaluate genetic continuity or transformation between present-day and ancient populations. To explore this, further mediaeval samples, regional genetic transects and complete genome analyses are aimed for. The follow-up project involves the study of mediaeval cemeteries from the same Odorheiu Secuiesc region to monitor the population history of the Székelys.

Author Contributions:

Conceptualisation, A.Sz-N. and B.G.M. and E. B and I.M.; Methodology, O.Sz. and N.B and B.Sz and D.G.; Formal Analysis, Writing – Original Draft Preparation, N.B. and O.Sz.; Writing – Review & Editing, E.B. and H.P and N.B. and A.Sz-N.; Visualisation, N.B. and O.Sz and A.Sz-N.; Supervision, A.Sz-N. and H.P and B.E.; Funding Acquisition, A.Sz-N.

Funding:

This paper was funded by the Hungarian National Research, Development and Innovation Office -FK 127938 project.

Institutional Review Board Statement:

For sampling, handling and storage of personal data and genetic samples, we considered the Hungarian 2008/XXI. law as guidelines. The Hungarian 2011/CXII. law provided us with rules about the information and self-determination rights of the sample providers. The research plan was forwarded to the Deputy State Secretary for National Medical Officers in the Ministry of Human Capacities, and to the Committee on Research Ethics for their consent.

Informed Consent Statement:

Informed consent was obtained from all subjects involved in the study.

Acknowledgement:

The authors would like to thank all voluntary sample donors and the organisers of the sampling for their contribution to the project.

Author Approvals:

All authors have seen and approved the manuscript, and it has not been accepted or published elsewhere.

Conflicts of Interest:

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses and interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations:

AMOVA: Analysis of Molecular Variance

HG: Haplogroup

MDS: Multidimensional Scaling

MJ: Median-Joining

NGS: Next-generation Sequencing

NJ: Neighbour-Joining

PCA: Principal Component Analysis

SNP: Single Nucleotide Polymorphism

STR: Short Tandem Repeat

Y-STR: Y chromosome Short Tandem Repeat

YHRD: Y-STR Haplotype Reference Database

ybp: years before present

- (1) Benkő, E.; Oborni, T. Székelyföld Története I. [The History of Szeklerland I.]; MTA Bölcsészettudományi Kutatóközpont; Erdély Múzeum-Egyesület; Haáz Rezső Múzeum: Székelyudvarhely, Romania, 2016; pp 94–103, 107–118.
- (2) Zimonyi, I. Muslim Sources on the Magyars in the Second Half of the 9th Century. In *East Central and Eastern Europe in the Middle Ages, 450–1450*; Brill Academic Publishers: Leiden–Boston, USA, 2015; pp 75–76.
- (3) Benkő, L.; Szabó T., Á. Die Szekler. Zur Siedlungsgeschichte Einer Ungarischen Volksgruppe. In *Ungarn Jahrbuch*; Budapest, Hungary, 1986; pp 207–224.
- (4) Benkő, E. A Középkori Székelyföld I. [The Medieval Székelyland I.]; MTA BTK Régészeti Intézet: Budapest, Hungary, 2012; pp 12–13.
- (5) Bárh, J. A Csíkszentmiklósi Havashasználat És a Tatros-Völgy Korai Népesége (The Usage of the Csíkszentmiklos Mountains and the Early Population of the Tatros Valley). In *A csíki Székely Múzeum évkönyve*; Csíkszereda, Romania, 2006; pp 17–36.
- (6) Egyed, B.; Brandstätter, A.; Irwin, J. A.; Pádár, Z.; Parsons, T. J.; Parson, W. Mitochondrial Control Region Sequence Variations in the Hungarian Population: Analysis of Population Samples from Hungary and from Transylvania (Romania). *Forensic Sci. Int. Genet.* **2007**, *1* (2), 158–162. <https://doi.org/10.1016/j.fsigen.2007.03.001>.
- (7) Brandstätter, A.; Egyed, B.; Zimmermann, B.; Duftner, N.; Padar, Z.; Parson, W. Migration Rates and Genetic Structure of Two Hungarian Ethnic Groups in Transylvania, Romania. *Ann. Hum. Genet.* **2007**, *71* (6), 791–803. <https://doi.org/10.1111/j.1469-1809.2007.00371.x>.
- (8) Tömör, G.; Csányi, B.; Bogácsi-Szabó, E.; Kalmár, T.; Czibula, Á.; Csősz, A.; Priskin, K.; Mende, B.; Langó, P.; Downes, C. S.; Raskó, I. Comparison of Maternal Lineage and Biogeographic Analyses of Ancient and Modern Hungarian Populations. *Am. J. Phys. Anthropol.* **2007**, *132* (4), 535–544. <https://doi.org/10.1002/ajpa>.
- (9) Egyed, B.; Füredi, S.; Padar, Z. Population Genetic Study in Two Transylvanian Populations Using Forensically Informative Autosomal and Y-Chromosomal STR Markers. *Forensic Sci. Int.* **2006**, *164* (2–3), 257–265. <https://doi.org/10.1016/j.forsciint.2005.10.020>.
- (10) Csányi, B.; Bogácsi-Szabó, E.; Tömör, G.; Czibula, Á.; Priskin, K.; Csősz, A.; Mende, B.; Langó, P.; Csete, K.; Zsolnai, A.; Conant, E. K.; Downes, C. S.; Raskó, I. Y-Chromosome Analysis of Ancient Hungarian and Two Modern Hungarian-Speaking Populations from the Carpathian Basin. *Ann. Hum. Genet.* **2008**, *72* (4), 519–534. <https://doi.org/10.1111/j.1469-1809.2008.00440.x>.
- (11) Bíró, A.; Fehér, T.; Bárány, G.; Pamjav, H. Testing Central and Inner Asian Admixture among Contemporary Hungarians. *Forensic Sci. Int. Genet.* **2015**, *15*, 121–126. <https://doi.org/10.1016/j.fsigen.2014.11.007>.
- (12) Ádám, V.; Bánfai, Z.; Sümegi, K.; Büki, G.; Szabó, A.; Magyar, L.; Miseta, A.; Kásler, M.; Melegh, B. Genome-Wide Marker Data-Based Comparative Population Analysis of Szeklers From Korond, Transylvania, and From Transylvania Living Non-Szekler Hungarians. *Front. Genet.* **2022**, *13* (March), 1–9. <https://doi.org/10.3389/fgene.2022.841769>.
- (13) Malyarchuk, B.; Derenko, M.; Denisova, G.; Litvinov, A.; Rogalla, U.; Skonieczna, K.; Grzybowski, T.; Pentelényi, K.; Guba, Z.; Zeke, T.; Molnár, M. J. Whole Mitochondrial Genome Diversity in Two Hungarian Populations. *Mol. Genet. Genomics* **2018**, *293* (5), 1255–1263. <https://doi.org/10.1007/s00438-018-1458-x>.
- (14) Völgyi, A.; Zalán, A.; Szvetnik, E.; Pamjav, H. Hungarian Population Data for 11 Y-STR and 49 Y-SNP Markers. *Forensic Sci. Int. Genet.* **2009**, *3* (2), 27–28. <https://doi.org/10.1016/j.fsigen.2008.04.006>.
- (15) Pamjav, H.; Fóthi, Á.; Fehér, T.; Fóthi, E. A Study of the Bodrogeköz Population in North-Eastern Hungary by Y Chromosomal Haplotypes and Haplogroups. *Mol. Genet. Genomics* **2017**, *292* (4), 883–894. <https://doi.org/10.1007/s00438-017-1319-z>.
- (16) Cocoş, R.; Schipor, S.; Hervella, M.; Cianga, P.; Popescu, R.; Banescu, C.;

- Constantinescu, M.; Martinescu, A.; Raicu, F. Genetic Affinities among the Historical Provinces of Romania and Central Europe as Revealed by an MtDNA Analysis. *BMC Genet.* **2017**, *18* (1), 1–11. <https://doi.org/10.1186/s12863-017-0487-5>.
- (17) Jakó, Z. Erdélyi Okmánytár. Codex Diplomaticus Transsylvaniae II.; Jakó, Z., Ed.; Magyar Országos Levéltár: Budapest, Hungary, 2004; Vol. 1, pp 413–413.
- (18) MAPSWIRE (accessed on: 27.10.2022). <https://mapswire.com/europe/physical-maps/>.
- (19) Fendt, L.; Zimmermann, B.; Daniaux, M.; Parson, W. Sequencing Strategy for the Whole Mitochondrial Genome Resulting in High Quality Sequences. *BMC Genomics* **2009**, *10*, 1–11. <https://doi.org/10.1186/1471-2164-10-139>.
- (20) Gerber, D.; Szeifert, B.; Székely, O.; Egyed, B.; Gyuris, B.; Giblin, J. I.; Horváth, A.; Palcsu, L.; Köhler, K.; Kulcsár, G.; Kustár, Á.; Szeverényi, V.; Fábrián, S.; Mende, B. G.; Bondár, M.; Ari, E.; Kiss, V.; Szécsényi-Nagy, A. Interdisciplinary Analyses of Bronze Age Communities from Western Hungary Reveal Complex Population Histories. *bioRxiv* **2022**, 2022.02.03.478968. <https://doi.org/10.1101/2022.02.03.478968>.
- (21) SeqPrep. <https://github.com/jstjohn/SeqPrep>.
- (22) Li, H.; Durbin, R. Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2010**, *26* (5), 589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
- (23) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25* (16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- (24) Weissensteiner, H.; Pacher, D.; Kloss-Brandstätter, A.; Forer, L.; Specht, G.; Bandelt, H. J.; Kronenberg, F.; Salas, A.; Schönherr, S. HaploGrep 2: Mitochondrial Haplogroup Classification in the Era of High-Throughput Sequencing. *Nucleic Acids Res.* **2016**, *44* (W1), W58–W63. <https://doi.org/10.1093/nar/gkw233>.
- (25) van Oven, M.; Kayser, M. Updated Comprehensive Phylogenetic Tree of Global Human Mitochondrial DNA Variation. *Hum. Mutat.* **2009**, *30* (2), 386–394. <https://doi.org/10.1002/humu.20921>.
- (26) Phylotree. <https://www.phylotree.org/> (accessed on: 27.10.2022).
- (27) International Society of Genetic Genealogy. *International Society of Genetic Genealogy. Y-DNA Haplogroup Tree 2019, Version: 15.73* (accessed on: 27.10.2022). 2020. <https://isogg.org/tree/>.
- (28) Leigh, J. W.; Bryant, D. POPART: Full-Feature Software for Haplotype Network Construction. *Methods Ecol. Evol.* **2015**, *6* (9), 1110–1116. <https://doi.org/10.1111/2041-210X.12410>.
- (29) Librado, P.; Rozas, J. DnaSP v5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics* **2009**, *25* (11), 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>.
- (30) Gouy, M.; Guindon, S.; Gascuel, O. Sea View Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol. Biol. Evol.* **2010**, *27* (2), 221–224. <https://doi.org/10.1093/molbev/msp259>.
- (31) Mansour, A. Phylip and Phylogenetics. *Genes, Genomes and Genomics* **2009**, *3*, 46–49.
- (32) Rambaut, A. *Figtree v 1.4.2*. (accessed on: 27.10.2022). 2014. <http://tree.bio.ed.ac.uk/software/figtree>.
- (33) *R: A language and environment for statistical computing* (accessed on: 27.10.2022). <https://www.r-project.org>.
- (34) Excoffier, L.; Lischer, H. E. L. Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Mol. Ecol. Resour.* **2010**, *10* (3), 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>.
- (35) Waskom, M. Seaborn: Statistical Data Visualization. *J. Open Source Softw.* **2021**, *6* (60), 3021. <https://doi.org/10.21105/joss.03021>.

- (36) <http://www.fluxus-engineering.com>. Network V10.1.0.0. (Accessed on: 27.10.2022).
- (37) Bandelt, H. J.; Forster, P.; Röhl, A. Median-Joining Networks for Inferring Intraspecific Phylogenies. *Mol. Biol. Evol.* **1999**, *16* (1), 37–48.
<https://doi.org/10.1093/oxfordjournals.molbev.a026036>.
- (38) Polzin, T.; Daneshmand, V. S. On Steiner Trees and Minimum Spanning Trees in Hypergraphs. *Oper. Res. Lett.* **2003**, *31* (1), 12–20. [https://doi.org/10.1016/S0167-6377\(02\)00185-2](https://doi.org/10.1016/S0167-6377(02)00185-2).
- (39) Achilli, A.; Rengo, C.; Magri, C.; Battaglia, V.; Olivieri, A.; Scozzari, R.; Cruciani, F.; Zeviani, M.; Briem, E.; Carelli, V.; Moral, P.; Dugoujon, J. M.; Roostalu, U.; Loogväli, E. L.; Kivisild, T.; Bandelt, H. J.; Richards, M.; Villems, R.; Santachiara-Benerecetti, S. A.; Semino, O.; Torroni, A. The Molecular Dissection of MtDNA Haplogroup H Confirms That the Franco-Cantabrian Glacial Refuge Was a Major Source for the European Gene Pool. *Am. J. Hum. Genet.* **2004**, *75* (5), 910–918.
<https://doi.org/10.1086/425590>.
- (40) Logan, I. <http://www.ianlogan.co.uk/> (accessed on: 27.10.2022).
- (41) Neparáczi, E.; Kocsy, K.; Tóth, G. E.; Maróti, Z.; Kalmár, T.; Bihari, P.; Nagy, I.; Pálfi, G.; Molnár, E.; Raskó, I.; Török, T. Revising MtDNA Haplotypes of the Ancient Hungarian Conquerors with next Generation Sequencing. *PLoS One* **2017**, *12* (4), 1–11. <https://doi.org/10.1371/journal.pone.0174886>.
- (42) Csáky, V.; Gerber, D.; Szeifert, B.; Egyed, B.; Stégmár, B.; Botalov, S. G.; Grudochko, I. V.; Matveeva, N. P.; Zelenkov, A. S.; Sleptsova, A. V.; Goldina, R. D.; Danich, A. V.; Mende, B. G.; Türk, A.; Szécsényi-Nagy, A. Early Medieval Genetic Data from Ural Region Evaluated in the Light of Archaeological Evidence of Ancient Hungarians. *Sci. Rep.* **2020**, *10* (1), 19137. <https://doi.org/10.1038/s41598-020-75910-z>.
- (43) Maár, K.; Varga, G. I. B.; Kovács, B.; Schütz, O.; Maróti, Z.; Kalmár, T.; Nyerki, E.; Nagy, I.; Latinovics, D.; Tihanyi, B.; Marcsik, A.; Pálfi, G.; Bernert, Z.; Gallina, Z.; Varga, S.; Költő, L.; Raskó, I.; Török, T.; Neparáczi, E. Maternal Lineages from 10-11th Century Commoner Cemeteries of the Carpathian Basin. *Genes (Basel)*. **2021**, *12* (460). <https://doi.org/10.3390/genes12030460>.
- (44) Neparáczi, E.; Maróti, Z.; Kalmár, T.; Kocsy, K.; Maár, K.; Bihari, P.; Nagy, I.; Fóthi, E.; Pap, I.; Kustár, Á.; Pálfi, G.; Raskó, I.; Zink, A.; Török, T. Mitogenomic Data Indicate Admixture Components of Central-Inner Asian and Srubnaya Origin in the Conquering Hungarians. *PLoS One* **2018**, *13* (11), e0205920.
<https://doi.org/10.1101/250688>.
- (45) Szeifert, B.; Stashenkov, D. A.; Khokhlov, A. A.; Sitdikov, A. G.; Gazimzyanov, I. R.; Volkova, E. V.; Matveeva, N. P.; Zelenkov, A. S.; Poshekhonova, O. E.; Sleptsova, A. V.; Karacharov, K. G.; Ilyushina, V. V.; Konikov, B. A.; Sungatov, F. A.; Kolonskikh, A. G.; Botalov, S. G.; Grudochko, I. V.; Komar, O.; Samara, T. M.; Researchgroup, H. P. Tracing Genetic Connections of Ancient Hungarians to the 6-14th Century Populations of the Volga-Ural Region. *Hum. Mol. Genet.* **2022**, *31* (19), 3266–3280.
<https://doi.org/https://doi.org/10.1093/hmg/ddac106>.
- (46) Skandakov, I. E. – Danchenko, E. M. Burial Mound Ust-Tara-VII in the Southern Taiga of the Irtysh Region. *Humanit. Knowl. Ser. “Continuity” Yearb. Collect. Sci. Pap.* **1999**, No. 3, 160–186.
- (47) De Barros Damgaard, P.; Marchi, N.; Rasmussen, S.; Peyrot, M.; Renaud, G.; Korneliussen, T.; Moreno-Mayar, J. V.; Pedersen, M. W.; Goldberg, A.; Usmanova, E.; Baimukhanov, N.; Loman, V.; Hedeager, L.; Pedersen, A. G.; Nielsen, K.; Afanasiev, G.; Akmatov, K.; Aldashev, A.; Alpaslan, A.; Baimbetov, G.; Bazaliiskii, V. I.; Beisenov, A.; Boldbaatar, B.; Boldgiv, B.; Dorzhu, C.; Ellingvag, S.; Erdenebaatar, D.; Dajani, R.; Dmitriev, E.; Evdokimov, V.; Frei, K. M.; Gromov, A.; Goryachev, A.; Hakonarson, H.; Hegay, T.; Khachatryan, Z.; Khaskhanov, R.; Kitov, E.; Kolbina, A.; Kubatbek, T.; Kukushkin, A.; Kukushkin, I.; Lau, N.; Margaryan, A.; Merkyte, I.; Mertz, I. V.; Mertz, V. K.; Mijiddorj, E.; Moiyesev, V.; Mukhtarova, G.; Nurmukhanbetov, B.; Orozbekova, Z.; Panyushkina, I.; Pieta, K.; Smrčka, V.; Shevnina, I.; Logvin, A.; Sjögren, K. G.; Štolcová, T.; Tashbaeva, K.; Tkachev, A.; Tulegenov, T.; Voyakin, D.;

- Yepiskoposyan, L.; Undrakhbold, S.; Varfolomeev, V.; Weber, A.; Kradin, N.; Allentoft, M. E.; Orlando, L.; Nielsen, R.; Sikora, M.; Heyer, E.; Kristiansen, K.; Willerslev, E. 137 Ancient Human Genomes from across the Eurasian Steppes. *Nature* **2018**, 557 (7705), 369–374. <https://doi.org/10.1038/s41586-018-0094-2>.
- (48) Csányi, B.; Bogácsi-Szabo, E.; Tömöry, G.; Czibula, Á.; Priskin, K.; Csősz, A.; Mende, B.; Langó, P.; Csete, K.; Zsolnai, A.; Conant, E. K.; Downes, C. S.; Raskó, I. Y-Chromosome Analysis of Ancient Hungarian and Two Modern Hungarian-Speaking Populations from the Carpathian Basin. *Ann. Hum. Genet.* **2008**, 72 (4), 519–534. <https://doi.org/10.1111/j.1469-1809.2008.00440.x>.
- (49) Purps, J.; Sievert, S.; Willuweit, S.; Nagy, M.; Alves, C.; Salazar, R.; Angustia, S. M. T.; Santos, L. H.; Anslinger, K.; Bayer, B.; Ayub, Q.; Wei, W.; Xue, Y.; Tyler-Smith, C.; Bafalluy, M. B.; Martínez-Jarreta, B.; Egyed, B.; Balitzki, B.; Tschumi, S.; Ballard, D.; Court, D. S.; Barrantes, X.; Bäßler, G.; Wiest, T.; Berger, B.; Niederstätter, H.; Parson, W.; Davis, C.; Budowle, B.; Burri, H.; Borer, U.; Koller, C.; Carvalho, E. F.; Domingues, P. M.; Chamoun, W. T.; Coble, M. D.; Hill, C. R.; Corach, D.; Caputo, M.; D'Amato, M. E.; Davison, S.; Decorte, R.; Larmuseau, M. H. D.; Ottoni, C.; Rickards, O.; Lu, D.; Jiang, C.; Dobosz, T.; Jonkisz, A.; Frank, W. E.; Furac, I.; Gehrig, C.; Castella, V.; Grskovic, B.; Haas, C.; Wobst, J.; Hadzic, G.; Drobnic, K.; Honda, K.; Hou, Y.; Zhou, D.; Li, Y.; Hu, S.; Chen, S.; Immel, U. D.; Lessig, R.; Jakovski, Z.; Ilievska, T.; Klann, A. E.; García, C. C.; De Knijff, P.; Kraaijenbrink, T.; Kondili, A.; Miniati, P.; Vouropoulou, M.; Kovacevic, L.; Marjanovic, D.; Lindner, I.; Mansour, I.; Al-Azem, M.; Andari, A. El; Marino, M.; Furfuro, S.; Locarno, L.; Martín, P.; Luque, G. M.; Alonso, A.; Miranda, L. S.; Moreira, H.; Mizuno, N.; Iwashima, Y.; Neto, R. S. M.; Nogueira, T. L. S.; Silva, R.; Nastainczyk-Wulf, M.; Edelmann, J.; Kohl, M.; Nie, S.; Wang, X.; Cheng, B.; Núñez, C.; Pancorbo, M. M.; Olofsson, J. K.; Morling, N.; Onofri, V.; Tagliabracci, A.; Pamjav, H.; Völgyi, A.; Bárány, G.; Pawlowski, R.; Maciejewska, A.; Pelotti, S.; Pepinski, W.; Abreu-Glowacka, M.; Phillips, C.; Cárdenas, J.; Rey-Gonzalez, D.; Salas, A.; Brisighelli, F.; Capelli, C.; Toscanini, U.; Piccinini, A.; Piglionica, M.; Baldassarra, S. L.; Ploski, R.; Konarzewska, M.; Jastrzebska, E.; Robino, C.; Sajantila, A.; Palo, J. U.; Guevara, E.; Salvador, J.; Ungria, M. C.; Rodriguez, J. J. R.; Schmidt, U.; Schlauderer, N.; Saukko, P.; Schneider, P. M.; Sirker, M.; Shin, K. J.; Oh, Y. N.; Skitsa, I.; Ampati, A.; Smith, T. G.; Calvit, L. S.; Stenzl, V.; Capal, T.; Tillmar, A.; Nilsson, H.; Turrina, S.; De Leo, D.; Verzeletti, A.; Cortellini, V.; Wetton, J. H.; Gwynne, G. M.; Jobling, M. A.; Whittle, M. R.; Sumita, D. R.; Wolańska-Nowak, P.; Yong, R. Y. Y.; Krawczak, M.; Nothnagel, M.; Roewer, L. A Global Analysis of Y-Chromosomal Haplotype Diversity for 23 STR Loci. *Forensic Sci. Int. Genet.* **2014**, 12, 12–23. <https://doi.org/10.1016/j.fsigen.2014.04.008>.
- (50) Yfull (accessed on: 27.10.2022). www.yfull.com/tree.
- (51) Neparáczki, E.; Maróti, Z.; Kalmár, T.; Maár, K.; Nagy, I.; Latinovics, D.; Kustár, Á.; Pálfi, G.; Molnár, E.; Marcsik, A.; Balogh, C.; Lőrinczy, G.; Gál, S. S.; Tomka, P.; Kovácsóczy, B.; Kovács, L.; Raskó, I.; Török, T. Y-Chromosome Haplogroups from Hun, Avar and Conquering Hungarian Period Nomadic People of the Carpathian Basin. *Sci. Rep.* **2019**, 9 (1), 16569. <https://doi.org/10.1101/597997>.
- (52) Fóthi, E.; Gonzalez, A.; Fehér, T.; Gugora, A.; Fóthi, Á.; Biró, O.; Keyser, C. Genetic Analysis of Male Hungarian Conquerors: European and Asian Paternal Lineages of the Conquering Hungarian Tribes. *Archaeol. Anthropol. Sci.* **2020**, 12 (1). <https://doi.org/10.1007/s12520-019-00996-0>.
- (53) Mendez, F. L.; Karafet, T. M.; Krahn, T.; Ostrer, H.; Soodyall, H.; Hammer, M. F. Increased Resolution of Y Chromosome Haplogroup T Defines Relationships among Populations of the Near East, Europe, and Africa. *Hum. Biol.* **2011**, 83 (1), 39–53. <https://doi.org/10.3378/027.083.0103>.
- (54) Papac, L.; Ernée, M.; Dobeš, M.; Langová, M.; Rohrlach, A. B.; Aron, F.; Neumann, G. U.; Spyrou, M. A.; Rohland, N.; Velemínský, P.; Kuna, M.; Brzobohatá, H.; Culleton, B.; Daněček, D.; Danielisová, A.; Dobisíková, M.; Hložek, J.; Kennett, D. J.; Klementová, J.; Kostka, M.; Křišťuf, P.; Kuchařík, M.; Hlavová, J. K.; Limburský, P.;

- Malyková, D.; Mattiello, L.; Pecinová, M.; Petrišćáková, K.; Průchová, E.; Stránská, P.; Smejtek, L.; Špaček, J.; Šumberová, R.; Švejcár, O.; Trefný, M.; Vávra, M.; Kolář, J.; Heyd, V.; Krause, J.; Pinhasi, R.; Reich, D.; Schiffels, S.; Haak, W. Dynamic Changes in Genomic and Social Structures in Third Millennium BCE Central Europe. *Sci. Adv.* **2021**, 7 (35). <https://doi.org/10.1126/sciadv.abi6941>.
- (55) Költő, L. Honfoglalás Kori Tegezés Sír Vörsön. In *A Herman Ottó Múzeum Évkönyve*; Publisher: Herman Ottó Múzeum, Miskolc, Hungary; 1993; Vol. 30-31/2, pp 433–445.
- (56) Maróti, Z.; Neparáczki, E.; Schütz, O.; Maár, K.; Varga, G. I. B.; Kovács, B.; Kalmár, T.; Nyerki, E.; Nagy, I.; Latinovics, D.; Tihanyi, B.; Marcsik, A.; Pálfi, G.; Bernert, Z.; Gallina, Z.; Horváth, C.; Varga, S.; Költő, L.; Raskó, I.; Nagy, P. L.; Balogh, C.; Zink, A.; Maixner, F.; Götherström, A.; George, R.; Szalontai, C.; Szenthe, G.; Gáll, E.; Kiss, A. P.; Gulyás, B.; Kovacsóczy, B. N.; Gál, S. S.; Tomka, P.; Török, T. The Genetic Origin of Huns, Avars, and Conquering Hungarians. *Curr. Biol.* **2022**. <https://doi.org/10.1016/j.cub.2022.04.093>.
- (57) Grugni, V.; Raveane, A.; Colombo, G.; Nici, C.; Crobu, F.; Ongaro, L.; Battaglia, V.; Sanna, D.; Al-Zahery, N.; Fiorani, O.; Lisa, A.; Ferretti, L.; Achilli, A.; Olivieri, A.; Francalacci, P.; Piazza, A.; Torroni, A.; Semino, O. Y-Chromosome and Surname Analyses for Reconstructing Past Population Structures: The Sardinian Population as a Test Case. *Int. J. Mol. Sci.* **2019**, 20 (22). <https://doi.org/10.3390/ijms20225763>.
- (58) Rootsi, S.; Magri, C.; Kivisild, T.; Benuzzi, G.; Help, H.; Bermisheva, M.; Kutuev, I.; Barac, L.; Peričić, M.; Balanovsky, O.; Pshenichnov, A.; Dion, D.; Grobei, M.; Zhivotovsky, L. A.; Battaglia, V.; Achilli, A.; Al-Zahery, N.; Parik, J.; King, R.; Cinnioğlu, C.; Khusnutdinova, E.; Rudan, P.; Balanovska, E.; Scheffrahn, W.; Simonescu, M.; Brehm, A.; Goncalves, R.; Rosa, A.; Moisan, J. P.; Chaventre, A.; Ferak, V.; Füredi, S.; Oefner, P. J.; Shen, P.; Beckman, L.; Mikerezi, I.; Terzić, R.; Primorac, D.; Cambon-Thomsen, A.; Krumina, A.; Torroni, A.; Underhill, P. A.; Santachiara-Benerecetti, A. S.; Villems, R.; Semino, O. Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe. *Am. J. Hum. Genet.* **2004**, 75 (1), 128–137. <https://doi.org/10.1086/422196>.
- (59) Margaryan, A.; Lawson, D. J.; Sikora, M.; Racimo, F.; Rasmussen, S.; Moltke, I.; Cassidy, L. M.; Jørsboe, E.; Ingason, A.; Pedersen, M. W.; Korneliussen, T.; Wilhelmson, H.; Buš, M. M.; de Barros Damgaard, P.; Martiniano, R.; Renaud, G.; Bhérrer, C.; Moreno-Mayar, J. V.; Fotakis, A. K.; Allen, M.; Allmæ, R.; Molak, M.; Cappellini, E.; Scorrano, G.; McColl, H.; Buzhilova, A.; Fox, A.; Albrechtsen, A.; Schütz, B.; Skar, B.; Arcini, C.; Falys, C.; Jonson, C. H.; Błaszczyk, D.; Pezhemsky, D.; Turner-Walker, G.; Gestsdóttir, H.; Lundström, I.; Gustin, I.; Mainland, I.; Potekhina, I.; Muntoni, I. M.; Cheng, J.; Stenderup, J.; Ma, J.; Gibson, J.; Peets, J.; Gustafsson, J.; Iversen, K. H.; Simpson, L.; Strand, L.; Loe, L.; Sikora, M.; Florek, M.; Vretemark, M.; Redknap, M.; Bajka, M.; Pushkina, T.; Søvsø, M.; Grigoreva, N.; Christensen, T.; Kastholm, O.; Uldum, O.; Favia, P.; Holck, P.; Sten, S.; Arge, S. V.; Ellingvåg, S.; Moiseyev, V.; Bogdanowicz, W.; Magnusson, Y.; Orlando, L.; Pentz, P.; Jessen, M. D.; Pedersen, A.; Collard, M.; Bradley, D. G.; Jørkov, M. L.; Arneborg, J.; Lynnerup, N.; Price, N.; Gilbert, M. T. P.; Allentoft, M. E.; Bill, J.; Sindbæk, S. M.; Hedeager, L.; Kristiansen, K.; Nielsen, R.; Werge, T.; Willerslev, E. Population Genomics of the Viking World. *Nature* **2020**, 585 (7825), 390–396. <https://doi.org/10.1038/s41586-020-2688-8>.
- (60) Berger, B.; Niederstätter, H.; Erhart, D.; Gassner, C.; Schennach, H.; Parson, W. High Resolution Mapping of Y Haplogroup G in Tyrol (Austria). *Forensic Sci. Int. Genet.* **2013**, 7 (5), 529–536. <https://doi.org/10.1016/j.fsigen.2013.05.013>.
- (61) Lipson, M.; Szécsényi-Nagy, A.; Mallick, S.; Pósa, A.; Stégmár, B.; Keerl, V.; Rohland, N.; Stewardson, K.; Ferry, M.; Michel, M.; Oppenheimer, J.; Broomandkhoshbacht, N.; Harney, E.; Nordenfelt, S.; Llamas, B.; Mende, B. G.; Köhler, K.; Oross, K.; Bondár, M.; Marton, T.; Oszás, A.; Jakucs, J.; Paluch, T.; Horváth, F.; Csengeri, P.; Koós, J.; Sebok, K.; Anders, A.; Raczky, P.; Regénye, J.; Barna, J. P.; Fábrián, S.; Serlegi, G.; Toldi, Z.; Nagy, E. G.; Dani, J.; Molnár, E.; Pálfi, G.; Márk, L.; Melegh, B.; Bánfai, Z.

- Domboróczki, L.; Fernández-Eraso, J.; Mujika-Alustiza, J. A.; Fernández, C. A.; Echevarría, J. J.; Bollongino, R.; Orschiedt, J.; Schierhold, K.; Meller, H.; Cooper, A.; Burger, J.; Bánffy, E.; Alt, K. W.; Lalueza-Fox, C.; Haak, W.; Reich, D. Parallel Palaeogenomic Transects Reveal Complex Genetic History of Early European Farmers. *Nature* **2017**, *551* (7680), 368–372. <https://doi.org/10.1038/nature24476>.
- (62) Rootsi, S.; Myres, N. M.; Lin, A. A.; Järve, M.; King, R. J.; Kutuev, I.; Cabrera, V. M.; Khusnutdinova, E. K.; Varendi, K.; Sahakyan, H.; Behar, D. M.; Khusainova, R.; Balanovsky, O.; Balanovska, E.; Rudan, P.; Yepiskoposyan, L.; Bahmanimehr, A.; Farjadian, S.; Kushniarevich, A.; Herrera, R. J.; Grugni, V.; Battaglia, V.; Nici, C.; Crobu, F.; Karachanak, S.; Hooshier Kashani, B.; Houshmand, M.; Sanati, M. H.; Toncheva, D.; Lisa, A.; Semino, O.; Chiaroni, J.; Di Cristofaro, J.; Villems, R.; Kivisild, T.; Underhill, P. A. Distinguishing the Co-Ancestries of Haplogroup G Y-Chromosomes in the Populations of Europe and the Caucasus. *Eur. J. Hum. Genet.* **2012**, *20* (12), 1275–1282. <https://doi.org/10.1038/ejhg.2012.86>.
- (63) Mathieson, I.; Lazaridis, I.; Rohland, N.; Mallick, S.; Patterson, N.; Roodenberg, S. A.; Harney, E.; Stewardson, K.; Fernandes, D.; Novak, M.; Sirak, K.; Gamba, C.; Jones, E. R.; Llamas, B.; Dryomov, S.; Pickrell, J.; Arsuaga, J. L.; de Castro, J. M. B.; Carbonell, E.; Gerritsen, F.; Khokhlov, A.; Kuznetsov, P.; Lozano, M.; Meller, H.; Mochalov, O.; Moiseyev, V.; Guerra, M. A. R.; Roodenberg, J.; Vergès, J. M.; Krause, J.; Cooper, A.; Alt, K. W.; Brown, D.; Anthony, D.; Lalueza-Fox, C.; Haak, W.; Pinhasi, R.; Reich, D. Genome-Wide Patterns of Selection in 230 Ancient Eurasians. *Nature* **2015**, *528* (7583), 499–503. <https://doi.org/10.1038/nature16152>.
- (64) Skourtanioti, E.; Erdal, Y. S.; Frangipane, M.; Balossi Restelli, F.; Yener, K. A.; Pinnock, F.; Matthiae, P.; Özbal, R.; Schoop, U.-D.; Guliyev, F.; Akhundov, T.; Lyonnet, B.; Hammer, E. L.; Nugent, S. E.; Burri, M.; Neumann, G. U.; Penske, S.; Ingman, T.; Akar, M.; Shafiq, R.; Palumbi, G.; Eisenmann, S.; D’Andrea, M.; Rohrlach, A. B.; Warinner, C.; Jeong, C.; Stockhammer, P. W.; Haak, W.; Krause, J. Genomic History of Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. *Cell* **2020**, *181* (5), 1158–1175.e28. <https://doi.org/https://doi.org/10.1016/j.cell.2020.04.044>.
- (65) Kazakov, Y. P. Volzhskie Bolgari, Ugri i Finni v IX – XIV vv [Volga Bulgars, Ugrians and Finns: Problems of Interaction]. In *Problemy vzaimodeistviya*; Kazan, 2007.
- (66) Dobos, A.; Gál, S. S.; Kelemen, I.; Neparáczi, E. Attila’s Europe? Structural Transformation and Strategies of Success in the European Hun Period; Rácz, Z., Szenthe, G., Eds.; Hungarian National Museum, Eötvös Loránd University: Budapest, Hungary, 2021; pp 327–356.
- (67) Csáky, V.; Gerber, D.; Koncz, I.; Csiky, G.; Mende, B. G.; Szeifert, B.; Egyed, B.; Pamjav, H.; Marcsik, A.; Molnár, E.; Pálfi, G.; Gulyás, A.; Kovacsóczy, B.; Lezsák, G. M.; Lőrinczy, G.; Szécsényi-Nagy, A.; Vida, T. Genetic Insights into the Social Organisation of the Avar Period Elite in the 7th Century AD Carpathian Basin. *Sci. Rep.* **2020**, *10* (1), 948. <https://doi.org/10.1038/s41598-019-57378-8>.
- (68) Gneccchi-Ruscione, G. A.; Szécsényi-Nagy, A.; Koncz, I.; Csiky, G.; Rácz, Z.; Rohrlach, A. B.; Brandt, G.; Rohland, N.; Csáky, V.; Cheronet, O.; Szeifert, B.; Rácz, T. Á.; Benedek, A.; Bernert, Z.; Berta, N.; Czifra, S.; Dani, J.; Farkas, Z.; Hága, T.; Hajdu, T.; Jászberényi, M.; Kisjuhász, V.; Kolozsi, B.; Major, P.; Marcsik, A.; Kovacsóczy, B. N.; Balogh, C.; Lezsák, G. M.; Ódor, J. G.; Szelekovszky, M.; Szeniczey, T.; Tárnoki, J.; Tóth, Z.; Tutkovics, E. K.; Mende, B. G.; Geary, P.; Pohl, W.; Vida, T.; Pinhasi, R.; Reich, D.; Hofmanová, Z.; Jeong, C.; Krause, J. Ancient Genomes Reveal Origin and Rapid Trans-Eurasian Migration of 7th Century Avar Elites. *Cell* **2022**, *185* (8), 1402–1413.e21. <https://doi.org/https://doi.org/10.1016/j.cell.2022.03.007>.
- (69) Grugni, V.; Raveane, A.; Ongaro, L.; Battaglia, V.; Trombetta, B.; Colombo, G.; Capodiferro, M. R.; Olivieri, A.; Achilli, A.; Perego, U. A.; Motta, J.; Tribaldos, M.; Woodward, S. R.; Ferretti, L.; Cruciani, F.; Torroni, A.; Semino, O. Analysis of the Human Y-Chromosome Haplogroup Q Characterizes Ancient Population Movements

- in Eurasia and the Americas. *BMC Biol.* **2019**, *17* (1), 1–14.
<https://doi.org/10.1186/s12915-018-0622-4>.
- (70) Olasz, J.; Seidenberg, V.; Hummel, S.; Szentirmay, Z.; Szabados, G.; Melegh, B.; Kásler, M. DNA Profiling of Hungarian King Béla III and Other Skeletal Remains Originating from the Royal Basilica of Székesfehérvár. *Archaeol. Anthropol. Sci.* **2019**, *11* (4), 1345–1357. <https://doi.org/10.1007/s12520-018-0609-7>.
- (71) Nagy, P. L.; Olasz, J.; Neparáczi, E.; Rouse, N.; Kapuria, K.; Cano, S.; Chen, H.; Di Cristofaro, J.; Runfeldt, G.; Ekomasova, N.; Maróti, Z.; Jeney, J.; Litvinov, S.; Dzhaubermeszov, M.; Gabidullina, L.; Szentirmay, Z.; Szabados, G.; Zgonjanin, D.; Chiaroni, J.; Behar, D. M.; Khusnutdinova, E.; Underhill, P. A.; Kásler, M. Determination of the Phylogenetic Origins of the Árpád Dynasty Based on Y Chromosome Sequencing of Béla the Third. *Eur. J. Hum. Genet.* **2021**, *29* (1), 164–172. <https://doi.org/10.1038/s41431-020-0683-z>.
- (72) Dulas, K.; Foody, M. G.; Justeau, P.; Silva, M.; Martiniano, R.; Oteo-García, G.; Fichera, A.; Simão, R.; Gandini, F.; Meynert, A.; Donnelly, K.; Aitman, T.; Chamberlain, A.; Lelong, O.; Kozikowski, G.; Powlesland, D.; Waddington, C.; Mattiangeli, V.; Bradley, D.; Bryk, J.; Soares, P.; Wilson, J.; Wilson, G.; Moore, H.; Pala, M.; Ceiridwen, E.; Richards, M.; Aitman, T.; Zosia, M.; Nicola, W.; Alison, M.; Biankin, A.; Santoyo-Lopez, J. Ancient DNA at the Edge of the World: Continental Immigration and the Persistence of Neolithic Male Lineages in Bronze Age Orkney. *Proc. Natl. Acad. Sci.* **2022**, *119* (8), e2108001119. <https://doi.org/10.1073/pnas.2108001119>.
- (73) Myres, N. M.; Rootsi, S.; Lin, A. A.; Järve, M.; King, R. J.; Kutuev, I.; Cabrera, V. M.; Khusnutdinova, E. K.; Pshenichnov, A.; Yunusbayev, B.; Balanovsky, O.; Balanovska, E.; Rudan, P.; Baldovic, M.; Herrera, R. J.; Chiaroni, J.; Di Cristofaro, J.; Vilems, R.; Kivisild, T.; Underhill, P. A. A Major Y-Chromosome Haplogroup R1b Holocene Era Founder Effect in Central and Western Europe. *Eur. J. Hum. Genet.* **2011**, *19* (1), 95–101. <https://doi.org/10.1038/ejhg.2010.146>.
- (74) Langó, P. „Salamon Gyűrűi” - Pajzs Alakú, Kiszélesedő, Díszített Fejű Pántgyűrűk a X. Századi Kárpát-Medencei Emlékanyagban. In *Beatus homo qui invenit sapientiam*; Csécs, T., Takács, M., Eds.; Lekri Group Kft: Győr, Hungary, 2016; pp 387–408.
- (75) Benkő, E.; Sándor, K.; Vásáry, I. A Székely Írás Emlékei. Corpus Monumentorum Alphabeto Siculico Exaratorum; Bölcsészettudományi Kutatóközpont: Budapest, Hungary, 2021; pp 834–835.