

ARCHAEOGENETIC DATA REVEAL THE GENETIC COMPOSITION OF 10TH-CENTURY CONQUERING HUNGARIANS

Gergely I. B. VARGA*

Introduction¹

The Conquering Hungarians arrived in the Carpathian Basin from the Steppe in the second half of the 9th century. Their archaeological horizon is well-documented and can be reliably identified.² However, their origins and genetic composition remain a subject of debate. The medieval chronicle tradition refers to them as relatives of the Huns, whereas linguistic research has clearly demonstrated that the Hungarian language belongs to the Finno-Ugric language family.³ Historical scholarship, based on the limited and chronologically restricted written sources and relying primarily on the results of archaeological research, seeks to reconstruct the history of the Hungarians prior to their migration into the Carpathian Basin.

* University of Szeged, Department of Genetics, Szeged, Hungary; ORCID ID: 0000-0001-9073-5788; vgisti@gmail.com

¹ The presentation and its written version present the results of research carried out by a team of numerous researchers. The researchers involved in the project are: Gergely I. B. Varga, Zoltán Maróti, Endre Neparácski, Oszkár Schütz, Kitty Maár, Bence Kovács, Tibor Kalmár, Emil Nyerki, István Nagy, Dóra Latinovics, Balázs Tihanyi, Antónia Marcsik, György Pálfi, Zsolt Bernert, Zsolt Gallina, Ciprián Horváth, Sándor Varga, László Költő, István Raskó, Péter L. Nagy, Csilla Balogh, Albert Zink, Frank Maixner, Anders Götherström, Robert George, Csaba Szalontai, Gergely Szenthe, Erwin Gáll, Attila P. Kiss, Bence Gulyás, Bernadett Ny. Kovácsóczy, Szilárd Sándor Gál, Péter Tomka, and Tibor Török. For the results of the research, see Maróti et al., 2022.

² Only a limited number of publications in Turkish are available on the archaeological record of the Conquering Hungarians. For the Hungarian conquest of the Carpathian Basin, their earlier settlement areas, often referred to as the “homeland,” as well as the archaeological research related to the conquering Hungarians, see Petkes–Sudár–Türk, 2017; Balogh, 2023. For a general characterization of the anthropological material of the Conquering Hungarians, see Marcsik–Tihanyi, 2021.

³ For the most comprehensive discussion in Turkish on the Hungarian Hun tradition, see Aydemir, 2021, 101–127.

To clarify their biological origins, we applied archaeogenetics to address the following key questions: (1) Can the first generation of immigrants to the Carpathian Basin be identified? (2) If so, where did they come from, and who were their ancestors? (3) Was their genetic makeup homogeneous, or did they form a heterogeneous population? (4) Furthermore, how were they related to earlier populations migrating into the Carpathian Basin from the east, such as the Huns and Avars, as well as the local population?

To explore these questions, we sequenced the whole genomes of a large number of individuals from 10th–11th century Carpathian Basin burials. Our dataset includes 48 skeletons from so-called elite cemeteries – i.e., burial sites of the military elite – and 65 individuals from village or commoner cemeteries. We intentionally overrepresented elite samples, as we expected to find a higher proportion of first-generation immigrants among them.

We compared our genome data to a large dataset of Eurasian samples. Our analysis included 1,400 individuals from 179 modern Eurasian populations and over 2,700 ancient genomes from approximately 200 archaeological cultures across Eurasia.

First, we identified close family relationships using the `correctKin` software. Next, we classified our samples into genetically homogeneous groups using various clustering methods based on genetic similarities. Finally, we modeled the genomic composition of the studied individuals using multiple advanced biostatistical approaches. Among these, the `qpAdm` method proved to be the most powerful, as it optimally reconstructs test genomes using potential source genomes. Additionally, we estimated admixture dates with the `DATES` algorithm.

The most widely used approach for grouping genomes by similarity is Principal Component Analysis (PCA). This method allowed us to create a genetic map from thousands of modern genomes, onto which we projected the studied samples. In the two-dimensional PCA plot, genetically similar genomes cluster closely together, while more divergent genomes are positioned farther apart. Notably, the genetic map also reflects geographic distribution: European populations appear on the left, East Asian populations on the right, and the second axis primarily separates populations along a north-south gradient.

The projected ancient genomes cluster near the most genetically similar modern and ancient samples. The majority of the 10th-century samples align with European populations (*Fig. 1*). They form a genetic continuum (European cline) ranging from Northern to Southern European genome types, indicating that they represent the local population. However, a significant number of individuals shift eastward, toward Asia, forming a distinct Conqueror cline. This eastward shift suggests that these individuals carry a substantial proportion of Asian genetic components – the greater the Asian ancestry, the farther they are positioned along this axis.

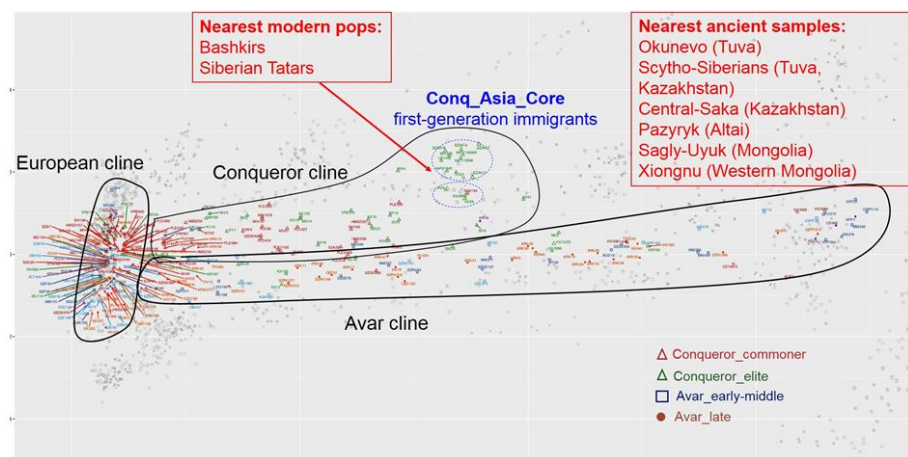


Figure 1. The PCA “genome map” of our newly sequenced genomes

At the easternmost end of the Conqueror cline, we identified a group of 12 elite Conquerors with the highest proportion of eastern genetic elements, which we term the Conqueror Asia Core (Conq_Asia_Core). These individuals occupy an intermediate position between Europe and Asia, suggesting a roughly equal mix of European and Asian ancestry. Importantly, our analysis shows that Conq_Asia_Core individuals lack recent European admixture, with their European ancestry tracing back to the Bronze Age. This finding indicates that they were first-generation immigrants.

The 12 Conq_Asia_Core genomes cluster closely together on the PCA plot, highlighting their genetic similarity and suggesting they belonged to the same population. Their genomes most closely resemble those of modern Bashkirs and Siberian Tatars, while among ancient populations, they show

the strongest affinity to Scytho-Siberian groups. These data suggest that the Conquerors likely originated from a region between the Ural Mountains and the Altai.

Since genome similarity alone does not indicate origin, we next investigated the ancestral origins of the Conq_Asia_Core group. We first modeled their ancestry using f_3 statistics, a method that identifies the populations most closely related to a test group by measuring shared genetic drift. This analysis included all available ancient and modern populations from genetic databases. The results showed that among the hundreds of tested populations, the closest genetic relatives of Conq_Asia_Core were Uralic-speaking peoples, with the Mansi – the closest linguistic relatives of modern Hungarians – showing the strongest genetic affinity. This striking parallel between linguistic and genetic evidence suggests that the Conquering Hungarians were likely responsible for bringing the Hungarian language to the Carpathian Basin.

Since the Mansi and the Conquerors share common ancestry, we included the Mansi in further analyses.

Ancestry modeling with qpAdm

Next, we applied qpAdm analysis to model the genetic composition of both the Conq_Asia_Core and the Mansi populations. This method determines the most plausible ancestral sources for a given population and quantifies their contributions.

First, we identified the distant Bronze Age ancestors of Conq_Asia_Core and the Mansi separately. The algorithm reconstructed the Conq_Asia_Core genome from three sources: 52% Mezhovskaya (a Late Bronze Age population), 13% Nganasan, 35% Late Bronze Age Mongolia. In comparison, the Mansi genome was modeled as: 48% Mezhovskaya, 44% Nganasan, 8% Kazakhstan Eneolithic (Botai).

The shared Mezhovskaya-Nganasan ancestry indicates a common ancestral population, reinforcing the f_3 statistics results. However, the differing third components suggest distinct population histories after their divergence.

Modeling recent admixture events

Next, we modeled the Conq_Asia_Core genome using more recent sources, including Iron Age and younger populations. Based on prior results, we also included modern Mansi as a potential source. The results were highly consistent, showing a genetic composition of: 50% Mansi, 35% Sarmatian, and 15% Scythian- or Xiongnu-related ancestry.

To determine whether the Hun-related or Scythian-related ancestry was the more accurate source, we estimated the timing of admixture using the DATES algorithm. The results confirmed that the admixture occurred between 290–461 CE, corresponding to the late-Xiongnu/early Hun era, thereby excluding the Iron Age Scythians as a major ancestral source. Meanwhile, the Sarmatian admixture was dated to 484–112 BCE, aligning with the early Sarmatian period in the Ural region.

These findings outline a prehistory in which the ancestors of Conq_Asia_Core initially formed a "Proto-Ugric" community with the ancestors of the Mansi during the Late Bronze Age. Over time, the two populations diverged, with the ancestors of Conq_Asia_Core intermixing first with early Sarmatians and later with late-Xiongnu and early Huns.

Validating late-Xiongnu admixture

To further verify the late-Xiongnu admixture, we conducted a two-dimensional F4-statistics analysis, measuring the presence of two characteristic late-Xiongnu genome components: SlabGrave and Han Chinese populations.

These two components varied simultaneously in the Conqueror genomes, indicating they entered the gene pool in a single admixture event, further supporting the qpAdm results (see *Fig. 2*).

The broader genetic landscape of the conquering Hungarians

Finally, we examined the origins of the remaining Conquering-period individuals. Most of those with European genomes came from commoner cemeteries, suggesting they were local inhabitants of the Carpathian Basin.

Individuals falling within the Conqueror cline (as seen in *Fig. 1*) exhibited a mix of European and Asian genetic components. We tested whether

these individuals could be explained as admixtures between local populations and the immigrant Conq_Asia_Core group. Indeed, most could be modeled as descendants of local inhabitants and first-generation Conquerors, suggesting they represent the later generations of the Conquerors in the Carpathian Basin.

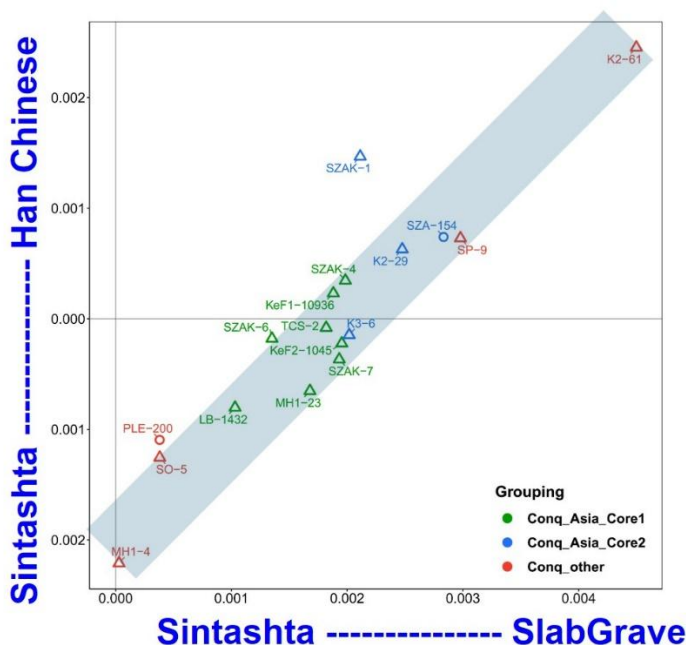


Figure 2. Major late-Xiongnu components of Han-Chinese and SlabGrave change together in Conq_Asia_Core according to f_4 -statistics

Finally, a significant number of individuals – primarily from the elite – could not be modeled as descendants of Conq_Asia_Core. Instead, they were best explained as admixtures of other Hun- and Alan-related steppe groups with the local population. This suggests that these individuals were integrated members of the elite who did not originate from Proto-Ugric-related sources, but rather from Hun-related steppe populations.

Statistical analysis of population distribution indicates that the majority of the 10th–11th century Carpathian Basin population consisted of local inhabitants, primarily commoners buried in large village cemeteries. The immigrant population had two main origins:

1. The majority were descendants of a Proto-Ugric-related community that had undergone admixture with Sarmatians and Huns.

2. A substantial proportion consisted of other steppe groups, primarily the descendants of the earlier Hun Empire, who joined the Conquering Hungarians in significant numbers.

References

- Aydemir, H. (2021). *Türk tarihinin kaynakları olarak eski Türk boy adları. Köken-bilgisel bir inceleme*, Cilt 1. Çanakkale.
- Balogh, Cs. (2023). Yurt tutan Macarların giyim kuşamı üzerine (On the clothings of the conquering Hungarians). Balogh, Cs. – Korucu Yağız, F. (Ed.), *Moda, kültür, sanat – Arkeoloji ve sanat tarihi çalışmaları — Fashion, culture, art – Archaeology and art history studies*. Çanakkale, 33–90.
- Marsik, A. – Tihanyi, B. (2021). Macar Ovası'nın Avar ve macar Yurt Tutuş Dönemi'ndeki Tarihsel ve Antropolojik Görünümü Kısa Bir özet. Çirkin, S. (Haz.), *Kökler: Yay Çeken Kavimlerin Şafağı. Dil, Arkeoloji, Tarih, Antropoloji ve Etnografya Işığında Altay Halklarının Kökeni*, Cilt III: *Antropoloji ve Etnografya*. İstanbul, 99–115.
- Maróti et al. (2022). The genetic origin of Huns, Avars, and conquering Hungarians. *Current Biology* 32, 2022, 1–13.
<https://doi.org/10.1016/j.cub.2022.04.093>
- Petkes, Zs. – Sudár, B. – Türk, A. (2017). *Ural'dan Karpat Havzası'na Macarların Göçü ve Yurt Tutması*. Çev. Şalıklolu, E. İstanbul.

Arkeogenetik verilerin ışığında 10. yüzyıl Macar fatihlerinin genetik bileşimi Genişletilmiş özet

Macarlar, 9. yüzyılın ikinci yarısında bozkır bölgesinden Karpat Havzası'na ulaşmışlardır. Arkeolojik ufukları iyi belgelenmiş olup güvenilir biçimde tanımlanabilmektedir. Orta Çağ kronik geleneği onları Hunların akrabaları olarak zikrederken, dilbilimsel araştırmalar Macarca'nın Ural dil ailesinin Fin-Ugor koluna ait olduğunu açık biçimde ortaya koymuştur. Tarihsel araştırmalar ise, sayıca az ve kronolojik açı-

dan sınırlı yazılı kaynaklara ve esas olarak arkeolojik bulgulara dayanarak, Macarların Karpat Havzası'na göçlerinden önceki tarihini aydınlatmayı amaçlamaktadır.

Karpat Havzası'na yerleşen Macarların biyolojik kökenlerini ve genetik bileşimlerini aydınlatmak amacıyla, temel araştırma sorularına yanıt aramak üzere arkeogenetik yöntemler uyguladık: (1) Karpat Havzası'na göç eden ilk nesil tanımlanabilir mi? (2) Eğer öyleyse, nereden gelmişlerdir ve ataları kimlerdir? (3) Genetik yapıları homojen miydi, yoksa heterojen bir nüfus mu oluşturuyorlardı? (4) Ayrıca, doğudan Karpat Havzası'na daha önce göç eden Hunlar ve Avarlar gibi topluluklarla ve yerel nüfusla ne tür ilişkiler içindeydiler?

Bu soruları incelemek amacıyla, 10–11. yüzyıla tarihlenen Karpat Havzası mezarlarından elde edilen çok sayıda bireyin tüm genomunu diziledik. Veri setimiz, askerî elit gömülerinden 48 iskelet ile halk (teba) mezarlıklarından 65 bireyi içermektedir. Genom verilerimizi geniş kapsamlı bir Avrasya karşılaştırma veri tabanıyla karşılaştırdık. Analizimiz, 179 günümüz Avrasya popülasyonundan 1.400 bireyi ve yaklaşık 200 arkeolojik kültüre ait 2.700'den fazla antik genomu kapsamaktadır.

İlk adım olarak, correctKin yazılımı yardımıyla yakın akrabalık ilişkilerini tespit ettik. Ardından, örneklerimizi genetik benzerliklere dayalı çeşitli kümeleme yöntemleri kullanarak genetik açıdan homojen gruplara ayırdık. Son olarak, incelenen bireylerin genomik bileşimini birden fazla gelişmiş biyostatistiksel yöntem uygulayarak modelledik. Bunlar arasında qpAdm yöntemi en etkili olanı olarak öne çıkmıştır; zira bu yöntem, incelenen genomları olası kaynak genomlar kullanarak en uygun biçimde yeniden yapılandırabilmektedir. Ayrıca, DATES algoritması aracılığıyla karışım (admixture) olaylarının zamanlamasını da tahmin ettik.

Genomların benzerliklerine göre sınıflandırılmasında en yaygın kullanılan yöntem başlıca bileşenler analizi (Principal Component Analysis, PCA)'dır. Bu yöntemle, binlerce modern genom temelinde bir genetik harita oluşturduk ve incelenen örnekleri bu harita üzerine projekte ettik. İki boyutlu PCA grafiğinde genetik olarak benzer genomlar birbirine yakın kümelenirken, daha uzak akrabalık ilişkisine sahip olanlar daha uzak konumlanır. Dikkat çekici biçimde, bu genetik harita coğrafi dağılımı da yansıtmaktadır: Avrupa popülasyonları sol tarafta, Doğu Asya popülasyonları sağ tarafta yer alırken, ikinci eksen esas olarak kuzey–güney yönündeki farklılıkları temsil etmektedir.

Projeksiyonla yerleştirilen antik genomlar, genetik olarak en benzer modern ve antik örneklerin yakınında kümelenmektedir. 10. yüzyıla ait örneklerin büyük çoğunluğu Avrupa popülasyonlarıyla örtüşmektedir (*Fig. 1*). Bunlar, Kuzey Avrupa'dan Güney Avrupa'ya uzanan bir genetik süreklilik (Avrupa klini) oluşturmakta olup yerel nüfusu temsil ettiklerini göstermektedir. Bununla birlikte, önemli sayıda birey doğuya, Asya yönüne doğru kayarak belirgin biçimde ayrılan bir “Fetihçi klini” oluşturmaktadır.

Fetihçi kliniğin doğu uç noktasında, doğu kökenli genetik bileşenlerin en yüksek oranını taşıyan 12 bireyden oluşan elit bir grup tespit ettik; bu grubu “Conqueror Asia Core” (Conq_Asia_Core) olarak adlandırdık. Bu bireyler, Avrupa ile Asya arasında ara bir konumda yer almakta olup yaklaşık olarak eşit oranlarda Avrupa ve Asya kökenine işaret etmektedir. Önemli bir bulgu olarak, bu bireylerde yakın döneme ait Avrupa kökenli genetik karışım görülmemektedir; Avrupa kökenli bileşenleri Bronz Çağı’na kadar izlenebilmektedir. Bu durum, onların birinci nesil göçmenler olduğunu göstermektedir.

PCA grafiğinde 12 Conq_Asia_Core genomu birbirine yakın biçimde kümelenebilmekte olup bu durum genetik benzerliklerini vurgulamakta ve aynı popülasyona ait olduklarını göstermektedir. Genomları modern Başkurtlar ve Sibirya Tatarlarına en yakın benzerliği gösterirken, antik popülasyonlar arasında en güçlü yakınlığı İskit-Sibirya gruplarıyla sergilemektedir. Bu veriler, fetihçi Macarların kökeninin büyük olasılıkla Ural Dağları ile Altaylar arasındaki bölgeye dayandığını göstermektedir.

Salt genomik benzerliğin tek başına kökeni göstermemesi nedeniyle, bir sonraki aşamada Conq_Asia_Core grubunun kökenini inceledik. İlk olarak, sözde f3-istatistiği kullanılarak kökenleri modellendi; bu yöntem, paylaşılan genetik sürüklenmenin derecesine dayanarak incelenen gruba en yakın popülasyonları belirler. Analize, genetik veri tabanlarında mevcut olan tüm antik ve modern popülasyonlar dâhil edildi. Sonuçlar, yüzlerce test edilen popülasyon arasından Conq_Asia_Core grubunun en yakın genetik akrabalarının Ural dil ailesine mensup topluluklar olduğunu göstermiştir; bunlar arasında ise günümüz Macarlarının en yakın dilsel akrabaları olan Mansiler en güçlü genetik benzerliği sergilemektedir. Dilbilimsel ve genetik veriler arasındaki bu dikkat çekici paralellik, fetihçi Macarların Macarca’yı Karpat Havzası’na getirmiş olabileceğini düşündürmektedir.

Mansiler ile fetihçi Macarlar ortak bir kökene sahip olduklarından, Mansileri sonraki analizlere de dâhil ettik.

Bunu takiben, Conq_Asia_Core ve Mansi popülasyonlarının genetik bileşimini modellemek amacıyla qpAdm analizini uyguladık. Bu yöntem, belirli bir popülasyonun en olası atasal kaynaklarını belirler ve bunların katkı oranlarını nicel olarak hesaplar.

İlk aşamada, Conq_Asia_Core ve Mansi gruplarının uzak, Bronz Çağı’na ait atalarını ayrı ayrı tespit ettik. Algoritma, Conq_Asia_Core genomunu üç kaynaktan türetmiştir: %52 Mezhovskaya (Geç Bronz Çağı popülasyonu), %13 Nganasan ve %35 Geç Bronz Çağı Moğolistan bileşeni. Buna karşılık, Mansi genomu şu şekilde modellenmiştir: %48 Mezhovskaya, %44 Nganasan ve %8 Kazakistan Eneolitik (Botai).

Ortak Mezhovskaya–Nganasan kökeni, f3-istatistiği sonuçlarını da teyit ederek, ortak bir atasal popülasyona işaret etmektedir. Buna karşılık, üçüncü bileşenler-

deki farklılıklar, iki popülasyonun ayrışmalarının ardından farklı gelişim süreçlerinden geçtiğini göstermektedir.

Bunu takiben, Conq_Asia_Core genomunu daha yeni, Demir Çağı ve sonraki dönemlere ait kaynak popülasyonlar kullanarak modelledik. Önceki sonuçlara dayanarak modern Mansileri de olası bir kaynak olarak analize dâhil ettik. Elde edilen sonuçlar yüksek derecede tutarlılık göstermiş ve şu genetik bileşimi ortaya koymuştur: %50 Mansi, %35 Sarmat ve %15 İskit veya Xiongnu kökenli bileşen.

Hun ya da İskit kökeninin hangisinin daha olası olduğunu belirlemek amacıyla, DATES algoritması kullanılarak genetik karışımın zamanlaması tahmin edilmiştir. Sonuçlar, karışımın MS 290–461 yılları arasında gerçekleştiğini göstermiştir; bu dönem geç Xiongnu / erken Hun evresiyle örtüşmekte olup, Demir Çağı İskitlerini başlıca kaynak olarak dışlamaktadır. Buna karşılık, Sarmat karışımının zamanı MÖ 484–112 arasına tarihlenmiş ve bu da Ural bölgesindeki erken Sarmat dönemiyle uyumlu bulunmuştur.

Bu bulgular, Conq_Asia_Core atalarının Geç Bronz Çağı'nda Mansi atalarıyla birlikte bir "proto-Ugor" topluluk oluşturduğunu göstermektedir. Daha sonraki süreçte bu iki popülasyon ayrılmış; Conq_Asia_Core ataları önce erken Sarmatlarla, ardından geç Xiongnu ve erken Hun gruplarıyla genetik olarak karışmıştır.

Geç Xiongnu kökenli katkıyı daha ileri düzeyde test etmek amacıyla, iki boyutlu F4-istatistik analizi gerçekleştirdik. Bu analiz, geç Xiongnu gen havuzuna özgü iki karakteristik bileşenin – SlabGrave ve Han Çinli popülasyonların – varlığını incelemektedir.

Bu iki bileşenin fetihçi genomlarda eşzamanlı olarak değişim göstermesi, bunların gen havuzuna tek bir genetik karışım olayı sonucunda dâhil olduğunu ortaya koymakta ve böylece qpAdm analizinin sonuçlarını daha da güçlendirmektedir (bkz. Gör. 2).

Son olarak, fetih dönemi nüfusuna ait diğer bireylerin kökenlerini de inceledik. Avrupa tipi genetik profile sahip bireylerin büyük çoğunluğu halk mezarlıklarından gelmekte olup, bunların Karpat Havzası'nın yerel nüfusunu temsil ettiğini göstermektedir.

Fetihçi klini içinde yer alan bireyler (bkz. Gör. 1), Avrupa ve Asya kökenli genetik bileşenlerin bir karışımını sergilemektedir. Bu bireylerin, yerel nüfus ile konar-göçer Conq_Asia_Core grubunun karışımı olarak açıklanıp açıklanamayacağını test ettik. Sonuçlar, çoğu durumda bunun mümkün olduğunu göstermiştir: bu bireyler, yerel halk ile birinci nesil fetihçilerin torunları olarak yorumlanabilir; yani Karpat Havzası'ndaki fetihçilerin daha sonraki nesillerini temsil etmektedirler.

Buna karşılık, önemli sayıda birey – özellikle elit tabakaya ait olanlar – Conq_Asia_Core grubunun doğrudan torunları olarak modellenememiştir. Bu bireyler en iyi şekilde, Hun ve Alan kökenli diğer bozkır toplulukları ile yerel nüfus arasındaki genetik karışım olarak açıklanabilmektedir. Bu durum, söz konusu bireylerin

proto-Ugor kökenli olmadığını, aksine Hun kökenli bozkır topluluklarından gelen ve elit yapıya entegre olmuş unsurlar olduğunu göstermektedir.

Nüfus dağılımına ilişkin istatistiksel analizler, 10–11. yüzyıl Karpat Havzası nüfusunun büyük çoğunluğunu yerel halkın oluşturduğunu, bunların da esas olarak geniş köy mezarlıklarına gömülen sıradan toplumsal kesimlere ait olduğunu ortaya koymaktadır. Göçmen nüfus ise iki ana kaynağa dayanmaktadır: 1. Çoğunluğu, Sarmatlar ve Hunlarla genetik olarak karışmış proto-Ugor kökenli bir topluluğun torunlarından oluşmaktadır. 2. Önemli bir kısmını ise diğer bozkır grupları, özellikle de eski Hun İmparatorluğu'nun torunları teşkil etmekte olup, bunlar fetihçi Macarlara önemli sayılarda katılmışlardır.

