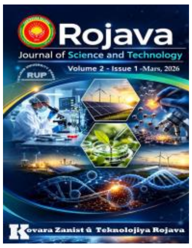
Rojava Journal of Science and Technology

2026, VOL. 2, No 1, [1-15]



The Phylogeny of Kurdish Populations: A Genetic and Historical Review

Pablo Gomez-Prieto[1*]

[1]**Department of Clinical Laboratory, Institute of Science and Modern Technology – Rojava University. Syria**

Abstract

The Kurdish people represent one of the largest ethnolinguistic groups in the Middle East. Despite their demographic significance, their genetic origins and phylogenetic relationships remain incompletely characterized. This review synthesizes findings from uniparental genetic markers, HLA allele frequency studies, and ancient genomic analyses to reconstruct the phylogenetic history of Kurdish populations and their relationships with neighboring groups.

The available evidence converges on a model of deep genetic continuity rooted in the Neolithic and Bronze Age populations of the Northern Fertile Crescent and Zagros Mountains. Modern Kurds carry approximately 75–80% indigenous Near Eastern ancestry, with a minor Western Steppe component attributable to Iron Age Indo-Iranian migrations. Uniparental marker analyses are consistent with descent from early Anatolian and Zagrosian farmers, alongside mitochondrial haplogroups H and U reflecting ancient Near Eastern hunter-gatherer and farming substrata. HLA studies place Kurds within a Mediterranean–Middle Eastern genetic cluster, with closest affinities to Iranians, Armenians, and Georgians, and suggest a possible ancestral link to the Hurrian populations of the Bronze Age. Significant intra-Kurdish genetic heterogeneity exists, shaped primarily by geographic isolation and tribal endogamy rather than linguistic boundaries.

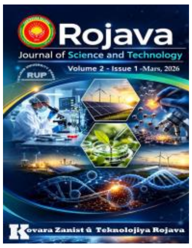
These findings refute models of wholesale population replacement by Central Asian migrants and instead support a continuity model in which Indo-Iranianization represented a linguistic and cultural overlay upon a pre-existing indigenous population. The phylogenetic stability of Kurdish populations since at least the Iron Age has important implications for medical genetics, including HLA-associated disease susceptibility and pharmacogenomic applications.

ARTICLE HISTORY
Received 12 Jan. 2026
Revised 3 Mar. 2026
Accepted 10 Mar. 2026

Keywords
Biology;
Anthropology;
Genetics; Phylogeny;
Kurdistan; History;
Review

---

[*] - Corresponding Author: p632129155@gmail.com



INTRODUCTION

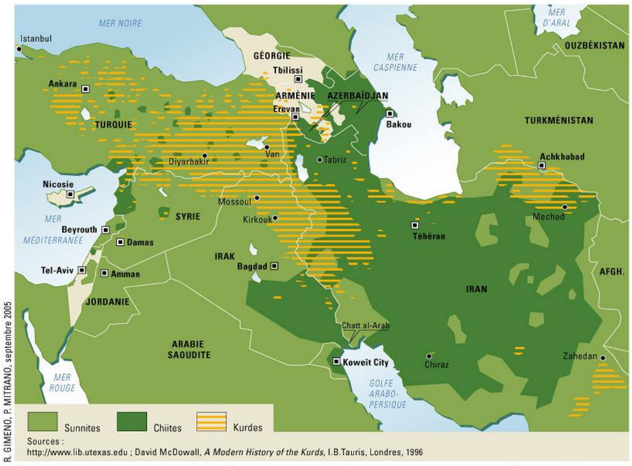
Kurdistan is a vast geographical and cultural region situated south of the Caucasus and north of Mesopotamia (Fig. 1). It has been continuously inhabited by Kurdish populations since at least 1000 BC [1]. Estimates place the modern Kurdish population between 30 and 38 million people, largely defined by their spoken Indo-European languages [2]. The distribution of present-day Kurds is widespread, with major populations in Turkey (15–20 million), Iran (10–12 million), Iraq (8–8.5 million), and Syria (3–3.6 million), alongside a significant diaspora in Europe and the Caucasus [3][4].

Kurds are considered an ethnic group since the medieval period [5]. Genetically, Kurds are part of an ancient Mediterranean, Middle Eastern, and Caucasian substratum. Unlike many neighboring groups, they retain a tangible link to Upper Paleolithic foragers and exhibit a slightly higher percentage of Neanderthal DNA, likely due to millennial-scale geographical isolation in highland refugia [6][7][8] .

The last decade has seen a paradigm shift in the study of human population history, driven largely by advances in ancient DNA (aDNA) recovery and high-throughput sequencing. The sequencing of 10,000-year-old genomes from the Zagros Mountains [9] and the broader archaeogenomic synthesis of the Southern Arc [10] have provided an unprecedented resolution of the demographic processes that shaped modern Near Eastern populations, including Kurds. These developments, combined with growing datasets from uniparental markers and high-resolution HLA typing, make a synthesis of Kurdish phylogenetics both timely and necessary.

These studies have been done with Kurds from all other three parts of Kurdistan but never from Rojava.

This review aims to: (i) summarize the methodological approaches used to reconstruct Kurdish population history, including Y-chromosome and mtDNA haplogroup analyses, HLA allele frequency studies, and ancient genomics; (ii) characterize the major findings regarding Kurdish origins, internal genetic structure, and relationships with neighboring populations; and (iii) discuss the medical and evolutionary significance of these findings, with particular attention to the implications for HLA-associated disease susceptibility and pharmacogenomics in this underrepresented group.



R. GIMENO, P. MITRANO, septembre 2005

Sources :
http://www.lib.utexas.edu ; David McDowall, *A Modern History of the Kurds*, I.B.Tauris, Londres, 1996

© *Questions internationales*, numéro 16, La Documentation française, Paris, novembre-décembre 2005.
http://www.ladocfrancaise.gouv.fr/revues/qi/sommaires/16/sommaire16.shtml

Fig 1. Extension of the Kurdish population in the Middle East

METHODOLOGIES IN PHYLOGENETIC RECONSTRUCTION

To accurately map the phylogeny of the Kurdish people, researchers rely on several methods that track different aspects of ancestral history [11].

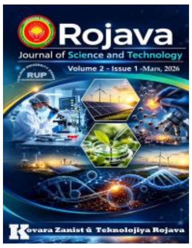
1. Uniparental Markers

Tracking ancestry through lineage-specific genes allows scientists to peer into the deep past because this DNA does not recombine like the rest of the genome.

Y-Chromosome (Y-DNA)

---

* - Corresponding Author: p632129155@gmail.com



Passed exclusively from father to son, 95% of the Y-chromosome remains identical over generations. Single Nucleotide Polymorphisms (SNP) define broad haplogroups, while Short Tandem Repeats (STR) track more recent tribal movements. The mutation rate of these markers can help us estimate when a common paternal ancestor lived (coalescence time).

There are several reasons why this method is appropriate for understanding the ancestry of Kurds:

- The presence of haplogroup R1a-Z93 in Kurds provides a direct paternal signature of the Indo-Iranian (Steppe) groups that brought the Kurdish language to the region.
- The high frequency of J2 and G markers connects Kurdish men directly to the Neolithic farmers of the Zagros Mountains from 10,000 years ago.
- Because many Middle Eastern cultures are patrilocal (men stay in their ancestral village while women move), Y-DNA often shows much more localized "tribal" patterns than other DNA types.

Mitochondrial DNA (mtDNA)

Passed exclusively from the mother, the hyper-variable "control region" (D-loop) serves as a molecular clock for tracking female migrations.

Why it's relevant to Kurds:

- Kurdish maternal lineages are extremely ancient and diverse[12]. High frequencies of haplogroup U suggest Kurds carry the DNA of the original Paleolithic hunter-gatherers of the Near East.
- While Kurdish paternal lines look more like neighboring West Asian groups (like Iranians/Turks), their maternal lines often show closer ties to European and Mediterranean groups. This "mismatch" reveals a complex history where indigenous women mixed with arriving male-dominated migrant groups.
- In groups like the Hawrami, low mtDNA diversity proves they have been a closed maternal community for a very long time, acting as a genetic isolate.

2. The HLA System and High-Resolution Typing

The Human Leukocyte Antigen (HLA) system on chromosome 6 is the most polymorphic genetic system in humans. The HLA system plays a critical role in our immune system's ability to distinguish "self" from "non-self." Because HLA gene frequencies show remarkable geographical correlation, they are indispensable for tracing ethnic composition and ancient migrations.

Certain combinations of contiguous alleles due to the strong linkage disequilibrium between HLA neighbouring loci show a characteristic frequency or are distinctive in many present day populations[13]. The HLA genetic system is a unique tool for studying the origins of relatively isolated groups, like some Kurdish populations.

HLA alleles are linked to diseases and also to response to drug treatments, which accomplishes relationships of certain HLA alleles with at least 15 different pathologies, including AIDS [14]. This is important in personalized treatment design.

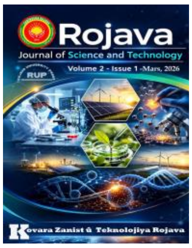
3. Ancient genomics

In 2016 10,000-year-old genomes from the Zagros Mountains (Wezmeh Cave) were sequenced [15][16]. This ancient DNA (aDNA) opened the door to the field of palaeogenetics, or ancient genomics.

The last decade of ancient genomics has revolutionized the study of human prehistory and evolution[17]. We review new insights based on

---

* - Corresponding Author: p632129155@gmail.com



prehistoric modern human genomes, including greatly increased resolution of the timing and structure of the earliest expansions of populations into Eurasia and America. Prehistoric genomes now document population transformations on every inhabited continent and record a history of natural selection that shapes present-day phenotypic diversity.

4. **Procedures**

Sampling

Blood is collected from individuals of genetically isolated Kurdish populations. DNA is extracted from the white blood cells, purified and fed to a PCR which amplifies a certain gene or group of genes. The chosen strand is usually from the mitochondrial DNA (maternal lineage), from the Y chromosome (paternal lineage) or from chromosome 6, at position 6p21.3. This region is one of the most gene-dense and polymorphic areas of the human genome.

100 healthy unrelated volunteers from the study area are typed for mtDNA, Y-DNA, or class I and class II HLA. All subjects as well as their four grandparents must be born in the study area. Subjects must not be related to one another.

Samples are preserved at room temperature in sterile tubes with DMSO in a saturated solution of NaCl.

HLA genotyping

Modern studies utilize the PCR-SSOP-Luminex technique to perform high-resolution analysis of HLA class I (A and B) and class II (DRB1 and DQB1) alleles [18]

Generic HLA class I and high resolution HLA class II analysis consists of:

a) PCR using specific primer pairs. All the primers are 5’-biotined and they are specific to determinate sequences of exons 2 and/or 3 of HLA genes [19].

b) Hybridization: products of PCR biotin-labelled are denaturalized at 97 ºC and then hybridized to complementary DNA probes associated to microbeads.

c) Assignation of HLA alleles: complex resulting of hybridization is introduced a Luminex platform, this system identifies the fluorescent intensity of fluorophores on each oligobead that has hybridized with the biotin-labelled PCR product. Luminex Software assigns the HLA alleles for each DNA sample [4].

Statistical analysis

By calculating the linkage disequilibrium between neighboring loci, researchers can identify characteristic frequencies that are distinctive to the Kurdish population [20].

HLA allele frequencies, Hardy-Weinberg equilibrium and the linkage disequilibrium between n alleles at n different loci are calculated. The level of significance (p) for 2 x 2 comparisons is determined[21][22]. The most frequent complete extended haplotypes are deduced from the 2, 3, and 4 HLA loci haplotype frequencies, the haplotypes in other populations, and haplotypes if they appeared in two or more individuals and the alternative haplotype was well defined [4][23].

Phylogenetic trees (dendrograms) are constructed with the allelic frequencies using the Neighbour-Joining (NJ) method [24] with the genetic distances between populations (DA) [25].

Correspondence analysis in three dimensions and its bidimensional representation is carried out[26]. This geometric technique may be used for displaying a global view of the relationships among populations according to allele frequencies. This method is based on the DA variance among populations and a statistical visualization of the differences (Fig. 2).

---

* - Corresponding Author: p632129155@gmail.com

COMPARATIVE PHYLOGENY

By combining lineage, somatic, and genomic studies, we see that the “Kurdish braid" is made of two different strands: a paternal strand that reflects the arrival of Indo-European languages, and a maternal strand that reflects the deep, unchanging presence of the original people of the mountains [1].

Modern Kurdish phylogeny is best described as a Zagrosian-Hurrian core that was "re-branded" by a relatively small but influential migration of Indo-Iranian (Steppe) warriors during the Iron Age.

The genetic signature of Kurds include traits such as resistance to local pathogens, adaptation to seasonal climates, and high lactose tolerance [6].

The internal phylogenetic structure of Kurdistan is a patchwork of tribal endogamy (marrying within the tribe) and mountain-valley isolation. Distinct differences exist between Kurmanji (Northern), Sorani (Central), and other groups like the Hawrami or Zaza. These differences are often driven more by geography and historical bottlenecks than by linguistic boundaries alone [12] [27].

1. The Hurrian and Kura-Araxes Connection

During the Bronze Age (3,500–1,200 BCE), the population integrated with Hurrian and Kura-Araxes cultures. Many geneticists believe the Kurds represent the "Indo-Iranianized" descendants of these ancient Hurrians [28][1]. While their language changed, their genetic core remained rooted in these pre-Indo-European populations.

There is likely a bottleneck effect in certain migrant groups, such as the Georgian Kurds, who show signs of isolation and increased genetic drift compared to their Anatolian and Mesopotamian counterparts [27] [29] .

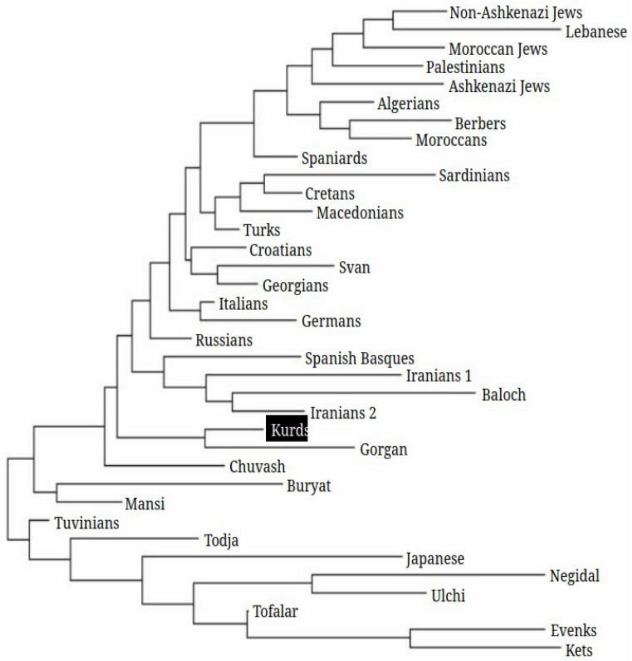


Fig 2. NJ dendrogram showing relations between Kurds and other world populations. DA were calculated using HLA-DRB1 (high resolution) [4]

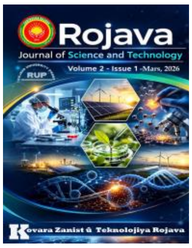
Y-DNA. Studies of Y-chrome STRs show that Kurdish populations exhibit high genetic diversity. In Iraqi Kurdistan, the genetic distance from Armenians is significantly shorter than from Assyrians[30]. Major Y-DNA haplogroups identified in Sorani Kurds include J2 (predominant at 28%), followed by E1b1b, J1, and G2a [31].

mtDNA. Most Kurdish mtDNA lineages originate from Western Eurasia, with high frequencies of haplogroups H (37.1%) and U (13.8%), which are typical for early Near Eastern farmers and hunter-gatherers [27].

HLA. HLA studies place Kurds within the Mediterranean cluster, showing strong affinities with Iranians, Georgians, and Palestinians. An ancient ancestor group for the Kurds may be the Hurrians (circa 2,000 BC) [29] [10].

---

* - Corresponding Author: p632129155@gmail.com



2. The Elite Dominance Model: The Median genetic myth vs. reality

While Kurds are often called "the descendants of the Medes," [32] new genomic research [33] offers a direct check. The Medes were likely a composite group. They were a mix of the incoming Steppe people and the local Mannean and Gutian populations. Because Kurds share the same genetic "profile" as these Iron Age Western Iranians, the claim is biologically accurate in terms of continuity, but the "Mede" part was just one chapter in a much older story [34] .

The arrival of Indo-Iranian tribes, such as the Medes, introduced Indo-European languages and the R1a-Z93 Y-haplogroup. However, recent research suggests this was a cultural and linguistic shift rather than a massive population replacement [33].

Haplogroup R1a is present in Kurds at about 12–20%. This "Steppe" marker confirms a legitimate Indo-Iranian branch but is significantly lower than in Eastern Iranian groups like Tajiks (68%).

Modern Kurds typically exhibit 75–80% indigenous Near Eastern DNA, with only a 10–18% Western Steppe component.

3. Tribal Heterogeneity and Regional Variation

The phylogeny of Kurds is not a single line, but a braided rope. The individual strands —the tribes— each preserve slightly different ratios of the Neolithic, Hurrian, and Indo-European inputs we discussed earlier.

Tribal structures have acted as genetic reservoirs, preserving ancient lineages.While "tribe" is a social construct, it has created genetic micro-clades [35]. For example, certain rare Y-STR mutations (like null alleles at DYS448) have been found in specific clusters in the Silêmanî province, likely tracing back to the founders of major tribal lineages like the Baban or Jaff.

The Barzani. A northern confederation (including the Sherwani, Muzuri, and others) exhibiting high Y-DNA diversity (FTDNA, 2023), reflecting a history of absorbing neighboring clans while maintaining high frequencies of Neolithic markers like **G-M201**. Historical records and genetic surveys suggest a shared foundation between the Barzani and Yezidi populations of the region, both showing deep-rooted North Mesopotamian ancestry before the 19th-century tribal shifts [1].

The Jaff. A central Zagros powerhouse showing a strong correlation with the "Zagros Neolithic" profile and specific HLA markers that link them to the Mediterranean genetic chain. Unlike migrant groups, the Jaff show a strong correlation with the "Zagros Neolithic" profile. Research on HLA markers in the Silêmanî region (the Jaff heartland) shows they are a primary link in the Mediterranean-Middle Eastern genetic chain [4].

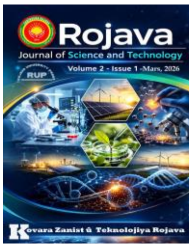
Interestingly, studies have noted that the Jaff and other Sorani tribes show an "intermediate" ABO blood group distribution between European and Asian patterns, but with a specific Kurdish frequency (O type at 37%) that suggests long-term regional preservation [27].

The Hawrami. Considered the "Gold Standard" for isolation, the Hawrami exhibit the lowest mtDNA diversity among Kurds, proving they have been a closed maternal community for millennia. Because of this lack of admixture, the Hawrami are considered one of the best proxies for the "Proto-Kurdish" genetic profile of the Bronze Age [35].

4. Tribal vs. General Population Phylogeny

---

* - Corresponding Author: p632129155@gmail.com



Tribal Group	Predominant Feature	Phylogenetic Significance
Barzani	High G-M201 & J2	Represents a confederation of indigenous mountain clans.
Jaff	High J2 & Mediterranean HLA	Key representative of the Central Zagros agricultural core.
Hawrami	Low mtDNA diversity	Evidence of extreme geographic isolation and ancient continuity.
Zaza	Similar to Kurmanji/Sorani	Refutes theories of a separate origin; confirms shared Kurdish ancestry.

5. Regional Subgroups

A primary focus in recent studies is the comparison of paternal and maternal lineages among these groups.

Paternal Lineages (Y-DNA)

- **Sorani Kurds**. Recent studies in the Silêmanî province of Iraq show a high frequency of haplogroup J2 (28%), with significant presence of E1b1b (16.5%) and J1 (14%). This suggests a strong connection to the Neolithic farmers of the Fertile Crescent [30] (Fig. 3).

- **Kurmanji Kurds.** While also showing high frequencies of J2, Kurmanji groups (especially in Turkey and the Caucasus) often show slightly higher frequencies of R1a and R1b compared to some Sorani samples. In the Caucasus, Kurmanji groups show signs of a bottleneck effect, with lower genetic diversity due to historical isolation [30][9].

Maternal Lineages (mtDNA)

- **European Affinities.** Interestingly, Kurmanji groups in Georgia and Turkey often show mtDNA lineages that are more closely related to European populations than to Caucasian populations. This suggests a distinct maternal history involving migrations from the West [29].

- **Sorani & Iranian Kurds.** These groups exhibit high frequencies of haplogroups H (37.1%) and U (13.8%), which are typical for the earliest Near Eastern farmers [12] (Fig. 3).

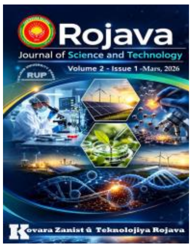
Kurdish Group	Geographic Focus	Key Genetic Finding
Kurmanji	Bakûr, Armenia, Georgia	Paternal ties to the Caucasus; Maternal ties to Europe. Evidence of genetic drift. [36]
Sorani	Başur (Silêmanî)	Strong J2 dominance; closely clustered with Iraq (Arabs), Lebanon, and Iran. [36]
Hawrami	Rojhilat	Lower mtDNA diversity suggests long-term isolation in mountainous regions. [27]
Zazaki	Bakûr	Genetically similar to other Kurdish groups, refuting theories of a distant North Iranian origin. [36]

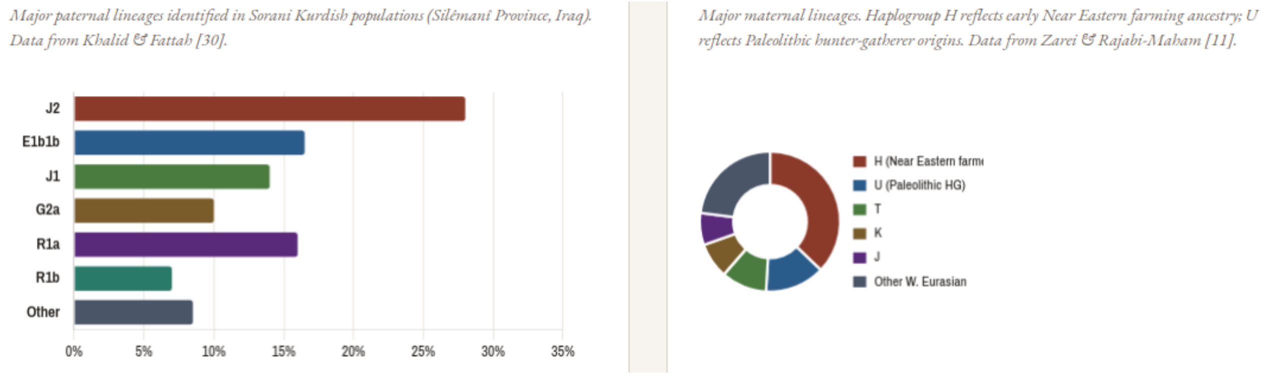
Insights on "Isolation by Distance"

While Kurds are a genetically cohesive group relative to distant populations (like Central Asians or Europeans), they are a "genetic mosaic" of the

---

* - Corresponding Author: p632129155@gmail.com




Major paternal lineages identified in Sorani Kurdish populations (Silêmanî Province, Iraq). Data from Khalid & Fattah [30].

Major maternal lineages. Haplogroup H reflects early Near Eastern farming ancestry; U reflects Paleolithic hunter-gatherer origins. Data from Zarei & Rajabi-Maham [11].

Fig 3. Paternal (left) an maternal (right) line haplogroup distribution in Sorani Kurds.

ancient Near East, with local variations caused by the rugged geography of the Zagros and Taurus mountains.

Geography, rather than language, seems to be the main force shaping Kurdish genetics. The Mantel test showed a strong "isolation by distance" effect. This means a Kurmanji speaker in Iraq is often genetically closer to a Sorani speaker in Iraq than they are to a Kurmanji speaker in the Caucasus [27].

Recent ancient DNA studies have revolutionized our understanding of Kurdish origins, shifting the focus from purely linguistic theories to a model of deep genetic continuity in the Zagros and Taurus mountains.

The prevailing scientific consensus is that modern Kurds primarily descend from indigenous Neolithic and Bronze Age populations of the Northern Fertile Crescent, rather than being late-arriving migrants from Central Asia [1][16].

6. Other Kurdish-related groups: genetic isolates

The phylogeny of Yezidi and Kurdish Jews confirms that "Kurdishness" is rooted in an ancient population of the Northern Fertile Crescent. These groups represent a genetic "time capsule" that links modern Middle Easterners directly to the Neolithic and Bronze Age civilizations of Mesopotamia.

Because both groups have historically practiced **endogamy**, they have acted as "genetic capsules," preserving lineages that date back to the Neolithic period.

The Yezidi

Recent studies [9][31] highlight two key phylogenetic findings:

5. Reproductive isolation due to strict religious laws has resulted in a notable reduction in Y-chromosome variability.
6. Despite this, they show significant genetic affinity with Iraqi Kurds and Armenians, confirming a common indigenous North Mesopotamian origin.

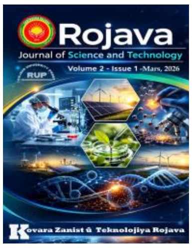
Some researchers view the Yezidi as the most direct descendants of the ancient inhabitants of Mesopotamia (Sumerians/Akkadians), having maintained their distinct identity while neighboring groups underwent arabization or turkification [16][35].

Kurdish Jews

Kurdish Jews and Kurdish Muslims are more closely related to each other than either is to Semitic-speaking Arabs [37]. A high frequency of the Eu9 (haplogroup J2) chromosome is found in both Jewish and Muslim Kurds. This marker is most prevalent in the Northern

---

* - Corresponding Author: p632129155@gmail.com



Fertile Crescent (modern-day Kurdistan, Turkey, and Armenia), suggesting a common ancestor from this region.

Interestingly, Kurdish Jews showed almost no genetic admixture with their host Muslim population over the 2,500 years they lived in Kurdistan, yet they remained genetically nearly identical to them. This suggests both groups descended from a single, ancient population that lived in the region before the arrival of modern religious and linguistic boundaries.

Medical & Evolutionary Significance

Studies on beta-thalassemia among Kurdish Jews [38] identified 13 different mutations, some of which are unique to this group. This "mutational diversity" in a small isolate suggests that central Kurdistan was a site of multiple independent mutational events over thousands of years, further proving the deep antiquity of the population in that specific geography.

7. Comparative phylogeny with neighbors

Genetic distance between a Kurd from Van and an Armenian is often smaller than the distance between that same Kurd and an Arab from Basra (Fig. 4).

The "Genetic Triangle"

The "Northern Highland" of the Middle East forms a distinct biological zone. **Kurds, Armenians, and Assyrians** have cohabited the area for millennia, and their DNA acts as a map of the Bronze Age and Iron Age transitions in Upper Mesopotamia and the South Caucasus.

Scientific consensus suggests that while they are linguistically diverse (Indo-European, Afro-Asiatic, and Isolate-turned-Indo-European), they share a massive genetic substratum rooted in the ancient Near East [39] .

Recent "Southern Arc" data [16] and regional studies in 2024 provide a breakdown of how "Indo-European" Kurds actually are compared to their neighbors. Kurds show 75-80% frequency of indigenous Near Eastern DNA, a 10-18% Western Steppe (MLBA) component, and 5-10% Levantine/Mediterranean ancestry. The Central Asian/East Asian part is negligible (<1%), in contrast to Turks, who show ~10–20% of this component [40].

Armenians and Kurds show the highest degree of genetic overlap, often clustering indistinguishably in PCA plots. Both populations carry high frequencies of haplogroup R1b and J2.

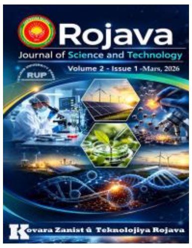
Assyrians. The presence of Y-DNA haplogroup J2 indicates an ancient connection to Assyrians. The version of J2 common in northern Mesopotamia implies that the Kurds, through their male lineage, carry markers of this ancient imperial past. Studies using admixture software like Admixture or TreeMix have shown cultural integration between highland herders and imperial administrators [4][6].

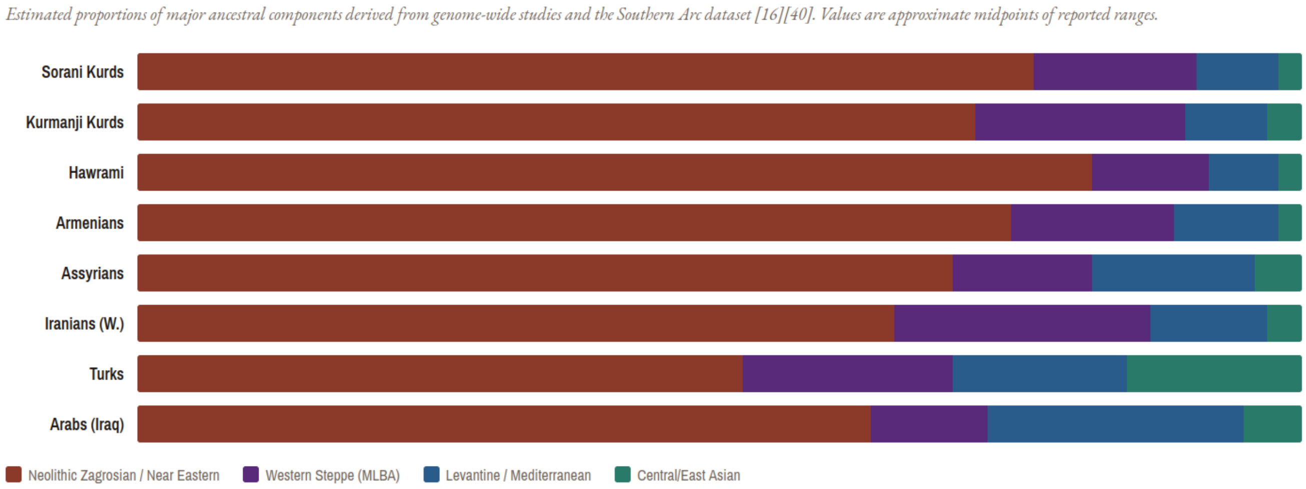
While sharing J2 and G markers, Assyrians show higher frequencies of J1, reflecting deeper Semitic roots in ancient Mesopotamia (Akkadians/Assyrians).

The "Mountain Refuge". All three groups lack the significant East Asian/Turkic admixture (typically <1%) found in modern Turkic populations (~10–20%), confirming that they are the primary descendants of the pre-Islamic inhabitants of the region.

---

* - Corresponding Author: p632129155@gmail.com





Estimated proportions of major ancestral components derived from genome-wide studies and the Southern Arc dataset [16][40]. Values are approximate midpoints of reported ranges.

Fig 4. Ancestral component admixture – Kurdish groups and neighboring populations

GENETIC HISTORY

Kurds have undergone remarkably little genetic change since the Sassanid era. These ancient Zagrosians were a distinct branch of humanity that separated from Anatolian farmers over 40,000 years ago (Fig. 5).

1. A Hurrian-Urartian Origin

Many geneticists point to the Kura-Araxes culture (3400–2000 BCE) as the common ancestor for these groups. The people of the Kura-Araxes culture are believed to be the speakers of Hurro-Urartian languages. As Indo-European (Armenian/Iranian) and Semitic (Aramaic) languages spread, these genetically similar people adopted different tongues but kept their biological heritage. This explains why a Kurd and an Armenian can be "genetic brothers" despite speaking languages from entirely different families.

The Hurrians and the Bronze Age

Linguistic and genetic markers suggest a link to the Hurrians (circa 3000–1000 BCE). Some researchers argue that the Kurds represent the "Iranianized" descendants of the Hurrians. While the Kurds now speak an Indo-European language, their genetic core is rooted in the pre-Indo-European Bronze Age populations of Mesopotamia and the Zagros[1][28].

Due to the rugged terrain of the Zagros and Taurus mountains, Kurdish groups acted as a "refuge" population, largely avoiding the massive genetic replacements seen in the flatter regions of the Middle East.

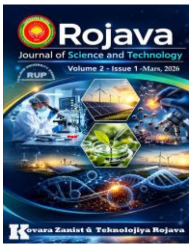
2. The Indo-European Layer (The Medes)

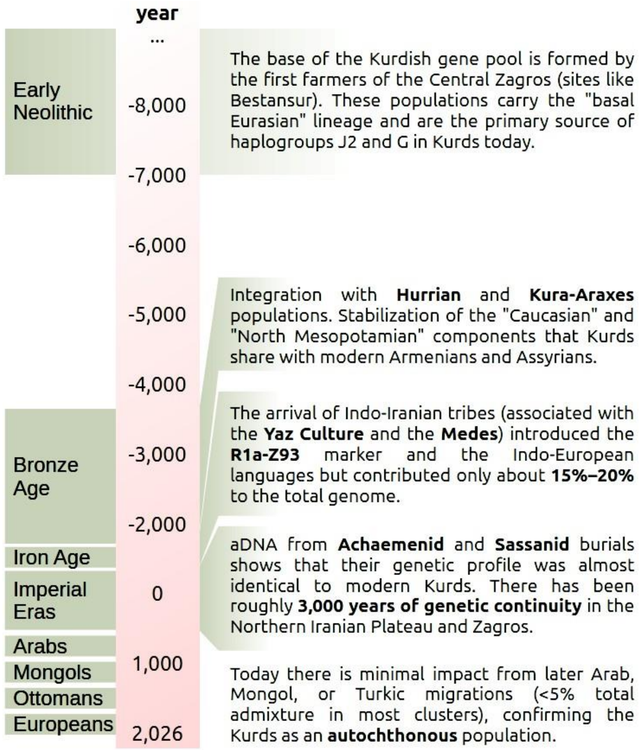
The traditional theory that Kurds are solely descendants of the Medes is now viewed with more nuance. Indo-Iranianization was likely a process where a smaller group of Central Asian migrants (carrying haplogroup R1a) imposed their language on the indigenous Karda/Hurrian majority.

While Kurds do carry the R1a marker associated with Indo-Europeans, the majority of their paternal (J2, G) and maternal (H, U) lineages are indigenous to the Near East.

---

* - Corresponding Author: p632129155@gmail.com





The base of the Kurdish gene pool is formed by the first farmers of the Central Zagros (sites like Bestansur). These populations carry the "basal Eurasian" lineage and are the primary source of haplogroups J2 and G in Kurds today.

Integration with **Hurrian** and **Kura-Araxes** populations. Stabilization of the "Caucasian" and "North Mesopotamian" components that Kurds share with modern Armenians and Assyrians.

The arrival of Indo-Iranian tribes (associated with the **Yaz Culture** and the **Medes**) introduced the **R1a-Z93** marker and the Indo-European languages but contributed only about **15%–20%** to the total genome.

aDNA from **Achaemenid** and **Sassanid** burials shows that their genetic profile was almost identical to modern Kurds. There has been roughly **3,000 years of genetic continuity** in the Northern Iranian Plateau and Zagros.

Today there is minimal impact from later Arab, Mongol, or Turkic migrations (<5% total admixture in most clusters), confirming the Kurds as an **autochthonous** population.

Fig 5. Kurdish genetic timeline based on bioarchaelogical and archaeogenetic research [26]

3. Genetic Continuity from the Iron Age

Comparing the phylogeny of modern Kurds with Ancient Iranian DNA from the Achaemenid (550–330 BCE) and Sassanid (224–651 CE) periods provides a clear look at how the "Kurdish" gene pool stabilized during the peak of the Persian Empires. Kurds have undergone remarkably little genetic change since the Sassanid era (Fig. 6).

The most recent research using samples from the Iranian Plateau and the Zagros mountains shows that Kurds are the primary genetic bridge between the ancient inhabitants of the Near East and the modern world [16][41] [42].

- **Achaemenid Period.** Ancient DNA from this era shows a population that was already a mix of Neolithic Zagrosian farmers and the incoming Indo-Iranian Steppe tribes. Modern Kurds share a 90% overlap with these Achaemenid-era samples, particularly those found in Western Iran.
- **Sassanid Period.** By the Sassanid era, the "Kurdish" profile was essentially complete. The genetic distance (FST) between a 5th-century Sassanid soldier from the Zagros and a modern Sorani Kurd is nearly zero, indicating long-term stability.

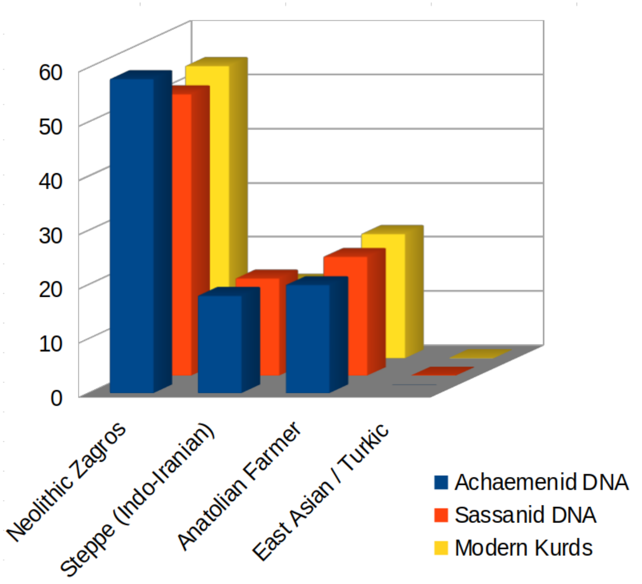


Fig 6. Kurds vs. Ancient Iranian Empires

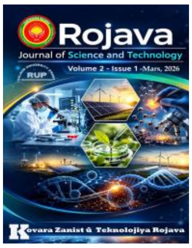
4. Kurds vs. Modern Persians (Western Iranians)

While both groups speak Iranian languages and share ancient roots, they have diverged due to different historical admixtures:

Modern Persians (particularly in Eastern and Central Iran) generally show a slightly higher "Steppe" (Indo-Iranian) and Central Asian component due to their proximity to the Silk Road and later Mongol/Turkic migrations.

---

* - Corresponding Author: p632129155@gmail.com



Kurds have maintained a higher percentage of the Neolithic Zagrosian and Caucasus Hunter-Gatherer (CHG) components. This makes Kurds "more ancient" in their genetic profile compared to some more urbanized or eastern Iranian populations [43].

5. Why the Gene Pool Hasn't Changed

The "phylogenetic stasis" of the Kurds since the time of the Persian Empires is attributed to three main factors:

1. **Topography.** The Zagros and Taurus mountains acted as a barrier against the "Genetic Replacement" seen in the plains of Mesopotamia (where arabization occurred) or the Iranian Plateau (where turkification occurred).
2. **Endogamy.** Social and tribal structures favored marriage within the community or tribe, preventing the dilution of the ancient gene pool.
3. **Refuge population.** During times of invasion (Mongol, Timurid, etc.), the mountains served as a safe haven, preserving the pre-invasion genetic signatures.

CONCLUSION

The phylogeny of Kurdish peoples is defined by a deep-rooted genetic stability forged in the highlands of the Zagros and Taurus mountains. This "mountain refuge" effect has acted as a biological preservative, maintaining ancient Near Eastern lineages through millennia of regional demographic turbulence. The convergent evidence from Y-chromosome, mtDNA, HLA, and ancient genomic studies paints a consistent picture: modern Kurds carry a Neolithic Zagrosian and Caucasus Hunter-Gatherer core (70–80%), augmented by a relatively modest Iron Age Indo-Iranian layer (15–20%). The result is a population of exceptional genetic depth, closely related to Armenians, Assyrians, and the populations of ancient Iranian empires, yet distinct from later-arriving groups whose genetic profiles bear stronger signatures of post-Bronze Age migrations.

The degree of this continuity is striking. Genomic comparisons between modern Kurds and ancient DNA samples from the Achaemenid and Sassanid periods suggest that a contemporary individual from Hewlêr or Sanandaj would be nearly indistinguishable, at the genomic level, from an inhabitant of the same region during the time of Cyrus the Great.

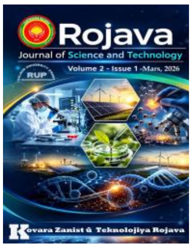
The most significant implication of recent data, including the 2025 archaeogenomic studies from the Northern Iranian Plateau, is the definitive refutation of mass migration models. Older hypotheses posited that Kurds arrived in their current territories during the Islamic era or the late Iron Age. Modern genomic science instead supports a model of in situ continuity, in which the Indo-Iranianization of the region represented a linguistic and political transformation overlaid upon a massive, pre-existing population of mountain farmers whose biological presence had remained essentially unbroken since the Neolithic.

It must be noted that the studies synthesized in this review draw primarily on samples from Başur, Bakûr, Rojhilat, Georgia, Armenia, and the Kurdish diaspora in Europe, with particular emphasis on the Zagros region. Genetic data from Rojava (Syrian Kurdistan) remains conspicuously absent from the literature. This is a significant gap, as western Kurdish populations may have experienced distinct admixture histories due to prolonged contact with Levantine and Anatolian populations. Future research should prioritize sampling from this region, as well as expanding ancient DNA coverage of the Iron Age Zagros, to further resolve the timing and scale of Indo-Iranian linguistic spread.

Increased collaboration with Kurdish academic institutions and greater investment in population-

---

* - Corresponding Author: p632129155@gmail.com



specific medical genomics —particularly regarding HLA-associated disease profiles— will be essential to translating these phylogenetic insights into clinical benefit for one of the world's largest stateless populations.

"The Kurdish genome is a high-resolution map of the Northern Fertile Crescent. The Mede influence was a linguistic and political overlay atop a massive, pre-existing population of mountain farmers who never left." — *Summary of findings from bioRxiv 2025.*

REFERENCES

[1] Hennerbichler F: The Origin of Kurds. Advances in Anthropology. 2012;2:64-79.

[2] Arnaiz-Villena A, Martinez-Laso J and Alonso-Garcia J: The correlation between languages and genes: the Usko-Mediterranean peoples. Human Immunology. 2001.

[3] Sellier J and Sellier A: Atlas des peuples d'Orient : Moyen-Orient, Caucase, Asie centrale. La Découverte; 1993. 206 p.

[4] Arnaiz-Villena A, Palacio-Grueber J, Muñiz E, Campos C and Alonso-Rubio J: Genetic HLA Study of Kurds in Iraq, Iran and Tbilisi (Caucasus, Georgia): Relatedness and Medical Implications. PLOS ONE. 2017;12. DOI: 10.1371/journal.pone.0169929

[5] Gugni V, Battaglia V, Hooshiar Kashiani B, Parolo S, Al-Zaheri N, Adhilli A, Olivieri A, Gandini F, Houshmand M, Hossein Sanati M, Torroni A and Semino O: Ancient Migratory Events in the Middle East: New Clues from the Y-Chromosome Variation of Modern Iranians. PLOS ONE. 2012. DOI: 10.1371/journal.pone.0041252

[6] Amirzargar A, Rey D, Muñiz E, Palacio-Grueber J, Nikbin B, Nicknam H, Khosravi F, Joshghan H, Areces C and Arnaiz-Villena A: Kurds HLA Genes: Its Implications in Transplantation and Pharmacogenomics. Open Medicine Journal. 2015. DOI: 10.2174/1874220301401010043

[7] Mellaart J: The Neolithic of the Near East. 1975.

[8] YouTube HF: DNA of the Kurds: A Journey Through History, Identity, and Genetic Legacy. 2025.

[9] Dogan S: A Genetic Context for the Levant and the Near East. Annals of Human Biology. 2017.

[10] Arnaiz-Villena A, Gómez-Casado E and Martinez-Laso J: Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. Tissue Antigens. 2002;6:111-121.

[11] Nei M: Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Science. 1973.

[12] Zarei F and Rajabi-Maham H: Phylogeography, genetic diversity and demographic history of the Iranian Kurdish groups based on mtDNA sequences. Journal of Genetics. 2016;95:767-776. DOI: 10.1007/s12041-016-0693-5

[13] Martinez-Laso J, Sartakova M and Allende L: HLA molecular markers in Tuvinians: a population with both Oriental and Caucasoid characteristics. Annal of Human Genetics. 2001;65:245-261.

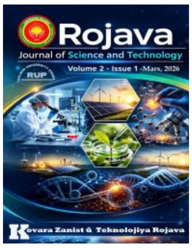
[14] Becquemont L: HLA: a pharmacogenomics success story. Pharmacogenomics. 2010;11:277-281.

[15] Broushaki F, Thomas MG, Link V, López S, van Dorp L, Kirsanow K, Hofmanová Z, Diekmann Y, Cassidy LM and Burger J: Early Neolithic genomes from the eastern Fertile Crescent. Science. 2016;353:499-503. DOI: 10.1126/science.aaf7943

[16] Lazaridis I, Nadel D, Rollefson G and Reich D: Genomic insights into the origin of farming in the ancient Near East. Nature. 2016;536:419-424.

---

* - Corresponding Author: p632129155@gmail.com



[17] Skoglund P and Mathieson I: Ancient Genomics of Modern Humans: The First Decade. Annual Review of Genomics and Human Genetics. 2018;19. DOI: 10.1146/annurev-genom-083117-021749

[18] Itoh Y, Mizuki N and Shimada T: High-throughput DNA typing of HLA-A, -B, -C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese population. Immunogenetics. 2005;57:717-729.

[19] Arguello R, Avakian H, Goldman JM and Madrigal JA: A novel method for simultaneous high resolution identification of HLA-A, HLA-B, and HLA-Cw alleles. PNAS. 1996;93:10961-10965. DOI: 10.1073/pnas.93.20.10961

[20] Schneider S: ARLEQUIN: a software for population genetics database analysis 2.0. 2000.

[21] Imanishi T, Akaza T, Kimura A, Tokunaga K and Gojobori T: Estimation of allele and haplotype frequencies for HLA and complement loci. 1991.

[22] Imanishi T, Akaza T, Kimura A, Tokunaga K and Gojobori T: Genetic relationships among various human populations indicated by MHC polymorphisms. 1991.

[23] Nei M, Tajima F and Tateno Y: Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. Molecular Evolution. 1983.

[24] Saitou N and Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406-425.

[25] Nei M: Genetic distances between populations. Am Nat. 1972;106.

[26] Young FW and Bann CM: A visual statistics system. Statistical computing environments for social researches. 1996:207-236.

[27] Nasidze I, Quinque D, Ozturk M, Bendukizde N and Stoneking M: mtDNA and Y-chromosome Variation in Kurdish Groups. Annals of Human Genetics. 2005. DOI: 10.1046/j.1529-8817.2005.00174.x

[28] Arnaiz-Villena A and Alonso J: Caucásicos, turcos, mesopotámicos y vascos. Madrid: Editorial Complutense; 2001. Available from: https://books.google.gr/books?hl=es&lr=&id=HpyFjLuiWswC&oi=fnd&pg=PA3&dq=Caucasicos,+Turcos,+Mesopotamicos+y+Vascos&ots=jBTeqAW27e&sig=rzb_P6e8heiucO-k856cS2uISG8&redir_esc=y#v=onepage&q=Caucasicos%2C%20Turcos%2C%20Mesopotamicos%20y%20Vascos&f=false

[29] Comas D, Calafell F, Bendukizde N, Fañanás L and Bertranpetit J: Georgian and Kurd mtDNA sequence analysis shows a lack of correlation between languages and female genetic lineages. American Journal of Biological Anthropology. 2000. DOI: 10.1002/(SICI)1096-8644(200005)112:1<5::AID-AJPA2>3.0.CO;2-Z

[30] Khalid SK and Fattah YM: Y-chromosomal Short Tandem Repeat Variation in Kurd, Assyrian, and Armenian populations in Iraq Kurdistan. Polytechnic Journal. 2023;12. DOI: 10.25156/ptj.v12n2y2022.pp148-157

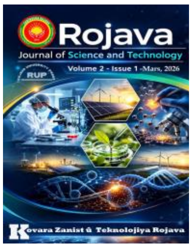
[31] Yunusbayev B: The Genetic Legacy of the Expansion of Turkic-Speaking Nomads across Eurasia. PLOS Genetics. 2015.

[32] Pailler J-M: Jean-Paul Demoule, Mais où sont passés les Indo-Européens ? Mythe d'origine de l'Occident. Paris: La librairie du xxie siècle; 2014. Available from: https://journals.openedition.org/pallas/3774 DOI: 10.4000/pallas.3774

[33] Amjadi MA, Ôzdemir YC, Ramezani M, Jakab K, Tavallaie M and Szécsényi-Nagy A: Ancient DNA indicates 3,000 years of genetic continuity in the Northern Iranian Plateau, from the Copper Age to the Sassanid Empire. Scientific Reports. 2025;15.

---

* - Corresponding Author: p632129155@gmail.com



[34] Haber M, Gauguier o, Youhanna S, Patterson N, Moorjani P, Botigué LR, Bertranpetit J, Comas D and Zalloua PA: Genome-Wide Diversity in the Levant Reveals Recent Structuring by Culture. PLOS Genetics. 2013. DOI: 10.1371/journal.pgen.1003316

[35] Dogan Sea: A glimpse at the intricate mosaic of ethnicities from Mesopotamia: Paternal lineages of the Northern Iraqi Arabs, Kurds, Syriacs, Turkmens and Yazidis. PLOS One. 2017.

[36] Albarzinji BM, Hadi S, Ismael B, Barqee A, Hadi A and Lazim H: An X-STRs analysis of the Iraqi Sorani Kurds. PLOS ONE. 2023. DOI: 10.1371/journal.pone.0294973

[37] Oppenheim A: The Y Chromosome Pool of Jews as Part of the Genetic Landscape of the Middle East. American Journal of Human Genetics. 2001;69:1095-1112.

[38] Rund D, Cohen T and Oppenheim A: Evolution of a genetic disease in an ethnic isolate: beta-thalassemia in the Jews of Kurdistan. PNAS. 1991;88. DOI: 10.1073/pnas.88.1.310

[39] Arnaiz-Villena A, Karin M, Bendikuze N, Gómez-Casado E, Moscoso J, Silvera C, Oguz FS, Sarper Diler A, de Pacho A, Allende L, Guillen J and Martinez-Laso J: HLA alleles and haplotypes in the Turkish population: relatedness to Kurds, Armenians and other Mediterraneans. Wiley Online Library. 2001.

[40] Suarez-Trujillo F, Juarez I, Palacio-Grueber J, Martin-Villa JM and Amirzargar A: HLA genetic study in Iran Saqqez-Baneh Kurds: no genetic trace of Aryan invasions in Anatolian Turks and Kurds is found. 2022. DOI: 10.1016/j.humimm.2022.07.005

[41] Gnecchi-Ruscone GA: Ancient DNA from the Iranian Plateau: Genomic perspectives on the Sassanid Empire. Nature Communications. 2024.

[42] Naderi S: Genomic stability in the Zagros Mountains over three millennia. Journal of Archaeological Science. 2023.

[43] Ghorani V: Genetic diversity and genome-wide association studies in Iranian populations. Human Heredity. 2020.

---

* - Corresponding Author: p632129155@gmail.com