ARAŞTIRMA MAKALESİ
Research Article

Relatives of Turkish: Altaic Languages Theory and Genes

Türkçenin Akrabaları: Altay Dilleri Teorisi ve Genler

Caner Kerimoğlu
İzmir/Türkiye
canerkerimoglu@yahoo.com
orcid: 0000-0002-8514-8578

Atıf Citation
Kerimoğlu, Caner (2024). Relatives of Turkish: Altaic Languages Theory and Genes. *BABUR Research, 3* (1), 99-137.

Gönderim Submitted
12.03.2024

Revizyon Revision
09.04.2024

Kabul Accepted
15.04.2024

Yayın Tarihi Publication Date
07.06.2024

Abstract

Genetics as a branch of science did not exist when Charles Darwin claimed that there might be a link between the spread of languages and human kinship. Genetic research has made it possible to examine this claim scientifically. Research on language and gene parallelism is progressing at a remarkable pace. Studies on language and gene overlap largely confirm Darwin. However, there are also a considerable number of cases where there is no overlap. In this paper, gene research on the Altaic languages and Turkish language in particular is reviewed. Recent research shows that speakers of Turkish language are genetically close to their neighbors. Studies on the theory of Altaic languages also emphasize the neighborly relationship. However, there are studies that report a genetic link between speakers of Altaic languages. The theory of Altaic languages, which is based on language features such as basic words and inflections, tends to expand with information obtained from archaeology and genetics.

Key words: Turkish, genetics, language family, Altaic languages theory, languages and genes.

Öz

Charles Darwin dillerin yayılımı ile insanların akrabalıkları arasında bir bağ olabileceğini iddia ettiğinde genetik diye bir bilim yoktu. Genetik araştırmaları bu iddiayı bilimsel olarak inceleme imkânı verdi. Dil ve gen paralelliği ile ilgili araştırmalar dikkat çekici bir hızla ilerlemektedir. Dil ve gen örtüşmesiyle ilgili çalışmalar büyük oranda Darwin'i doğrulamaktadır. Bununla birlikte azımsanmayacak sayıda örtüşme bulunmayan örnekler de vardır. Bu çalışmada Altay dilleri ve özellikle de Türkçe ile ilgili gen araştırmalarını değerlendiriyorum. Son araştırmalar Türk dili konuşurlarının genetik olarak komşularına yakın olduğunu göstermektedir. Altay dilleri teorisi ile ilgili çalışmalarda da komşuluk ilişkisi öne çıkmaktadır. Ancak Altay dillerini konuşanlar arasında genetik bir bağ olduğunu raporlayan araştırmalar bulunmaktadır. Temel sözcükler, çekim ögeleri gibi dil özellikleri üzerinden ilerleyen Altay dilleri teorisi arkeoloji ve genetikten gelen bilgilerle de genişleme eğilimdedir.

Anahtar Sözcükler: Türkçe, genetik, dil ailesi, Altay dilleri teorisi, diller ve genler.

BABURRESEARCH

Introduction

The study of language families is a field that has made significant progress with historical comparative language research. Today, there are nearly 7,000 languages in the world and a rich literature on the families to which these languages belong (Campbell & Poser, 2008; Eberhard et al., 2023; Pereltsvaig, 2021). Many measures are used to classify languages or determine the family to which they belong; these measures are usually based on linguistic features such as sound correspondences and basic lexical similarities. In this article, I will discuss kinship based on another phenomenon: Genes.

Genetics is the first field that comes to mind when we think of heredity today. In recent years, many surprising conclusions have been reached on the relationship between genes and language (Kerimoglu, 2017). Moreover, research centers, projects and textbooks on the relationship between genes and language are more popular than ever (Barbieri & Widmer, 2024; Dediu, 2015; A. McMahon & McMahon, 2013; Stoneking, 2017). Colin Renfrew even calls the tripod of archaeology, linguistics and genetics the "new synthesis", reminiscent of the renowned modern synthesis of natural selection theory[1] (Renfrew, 2000, 2010). While it is estimated that we can trace back at most 8-10 thousand years about the history of languages based on word comparisons compiled from written products (Greenhill et al., 2010; Nichols, 1992; Pagel, 2000), there have also been studies suggesting that it is possible to trace back even further with methods based on the typological features of languages (Dunn et al., 2005; Gray, 2005) but thousands of expressions can only partially provide data about what is going on a 4.5 billion- year-old planet. At this point, data from archaeology and genetics may make much more ancient contributions to our understanding of language, a complex ability that has survived millions of years. Just as Mendelian genetics "validated" Darwin's theory of natural selection, the idea that genetics could make a similar contribution to language is being discussed in the scientific world, and the methods to be used to ensure that the data obtained from the collaborative work of genetics, linguistics and archaeology experts are questioned (Greenhill, 2021; R. McMahon, 2004).

Today, genes are used in language research in two contexts. The first is the identification of genes involved in the "production" of human language. For example, a question such as whether there is a gene that allows us to process the meaning of words is the subject of this kind of research. The second context in which the relationship between genes and language is considered is the connection between the relatedness of languages and the relatedness of speakers of those languages. For example, the question of whether speakers of Indo-European family languages are genetically related to each other requires research into the genetic relatedness of speakers of those languages. In a separate article, I have elaborated on this topic and provided more comprehensive information about current discussions (Kerimoglu, 2023).

---

[1] Modern synthesis: A version of the theory of natural selection made even more powerful by the combination of Darwin and Mendel's major discoveries.

In this article, I will evaluate the current research in the scientific world on the parallels between the place of Turkish language in terms of its relatedness to other languages in the language family and the Turkish speakers' relatedness to the speakers of languages belonging to the same language family. First, let us see which languages Turkish is related to.

"Family" of Turkish According to the Data of Historical Comparative Language Studies

The question of whether Turkish language has a language family is usually answered by the "Altaic languages theory". This theory holds that Turkish, Mongolian and Manchu-Tungusic derive from the same "ancestral language". Over time, the theory has expanded to include Korean and Japanese in the family (For history, see. Blažek, 2019; de Rachewiltz & Rybatzki, 2010; Demir & Yılmaz, 2002; Janhunen, 2023; Robbeets & Savelyev, 2020; Tekin, 2003; Tuna, 1992). While the term "Altaic Languages" was introduced to the world of science by M. A. Castrén, the sound equivalences were determined by G. Ramstedt, who is considered the pioneer of the theory (Ercilasun, 2004). Castrén used the term "Altaic languages" for the Ural-Altaic languages, Castrén used the term "Altaic languages" for the Ural-Altaic languages but believed that the Uralic languages (Finno-Ugoric, Samoyed) were not related to the Turko-Tatar, Mongolian, Tungus languages. American researchers such as S. E. Martin, R. A. Miller, J. Street did important work on the inclusion of Korean and Japanese in the theory (Martin, 1966; Miller, 1971; Miller & Street, 1975; Street, 1980). Sergei Starostin is another name that must be mentioned in the theory of Altaic languages. The dictionary he published with his team A. Dybo and O. Mudrak stands out as an etymological dictionary of the Altaic languages and is a significant accomplishment (Starostin et al., 2003). Even after Sergei Starostin's early and unexpected death, his team (including his son Georgy) continued the research on origins based on lexical statistics (Egorov et al., 2022; Kassian et al., 2021). From Turkey, T. Tekin is one of the names that should be mentioned in Altaic studies (Tekin, 2003). In recent years, Martine Robbeets is the most prominent name who has brought the issue of Altaic languages being related to each other to the forefront by using the term "Trans-Eurasia". Using data from different fields such as agricultural culture, genetics and archaeology, he argues that these languages have a common origin (Robbeets, 2005; Robbeets et al., 2021; Robbeets & Bouckaert, 2018; Robbeets & Savelyev, 2020; Savelyev & Robbeets, 2020).

The theory of the Altaic languages followed a path similar to the one followed in the foundation of the Indo-European language family. Some common grammatical features of these languages (vowel harmony, lack of grammatical gender, etc.) had been known since the work of F. Wiedemann (Wiedemann, 1838). Phonetic correspondences, basic words and structural similarities were used as the main features in the construction of the theory. W. Schott was the first to find /ş/:/l/, /z/:/r/ correspondences within Turkish between General Turkish and Chuvash (Schott, 1841); G. Ramstedt later extended these equivalences to other Altaic languages and became the founder of the theory (Ramstedt, 1922, 1952, 1957, 1966). An intensive publication activity on the relatedness of these languages started. However, the theory of Altaic languages was not generally accepted in the scientific world, because important researchers such as G. Clauson and G. Doerfer argued that these languages were not related and that common features emerged as a result of the influence of Turkish on these languages.

One of the most frequently used tools for determining language relatedness is M. Swadesh's word lists (Swadesh, 1952, 1955; Swadesh et al., 1971). These lists are used to determine the degree of relatedness of the languages being compared by compiling a list of the most difficult words to borrow in a way that excludes influence. Words such as numbers, organ names, basic verbs are included in these lists. However, the opponents of the theory of Altaic languages have concluded that the claimed languages cannot be related, based on the tests they conducted on the word lists. Clauson argues that there are not enough "basic words" to show that the Altaic languages share a common ancestor and claims that the commonalities are the result of Turkish influencing these languages (Clauson, 2004, 2017). Numbers and organ names are words that are frequently used in language family studies and are assumed to be less prone to borrowing. There is no common number word among the Altaic languages, except for Turkish and Mongolian (*dört* and *dörben*). G. Doerfer, like Gerard Clauson, argues that these common words are borrowed from Turkish into Mongolian. Doerfer conducted a research on 11 main basic words (*head, eye, ear, nose, mouth, tongue, tooth, hair, heart, hand, foot*) and 5 intermediate basic words (*lip, finger, knee, beard, neck*) he identified from organ names, taking into account other language families. Stating that the borrowing of these basic words is more difficult than the borrowing of other words, Doerfer stated that in cases where the commonalities are intense, the theory of the same origin may be correct. In his research, he found that all 16 words in Semitic languages, 13 in Indo-European languages, 10 in Dravidian languages, 8 in Uralic languages are the same, but there is not a single commonality in Altaic languages. Based on these results, Doerfer considered the theory of Altaic languages as a "destroyed" theory[2] (Doerfer, 1983). Today, A. Vovin and S. Georg are two prominent scholars whose publications

---

[2] Against this interpretation based on organ names, Manaster Ramer et al. conducted another study and concluded that there are common organ names among Altaic languages (Manaster Ramer et al., 1998). Organ names have always attracted the attention of researchers as one of the criteria used for language relatedness. A recent study of 1028 languages found that cultural influences as well as some universal patterns play a role in organ naming (Tjuka et al., 2024). Cultural differences provide data for theories that center on "influence" in questions of linguistic diversity, but when universality is identified, questions of common origin arise: Are these universals the result of human biology and cognition, or do they stem from a single common ancestral language, as Eurasiatic and Nostratic theories claim? Can the identification of common features between an island language that has never come into contact with other languages and another language thousands of kilometers away from that island be explained by influence? Chomsky and his followers see "biology" as the cause of these commonalities, while the opposing camp hypothesizes a common "ancestral language" (Berwick & Chomsky, 2016; Johansson, 2005; Ruhlen, 2006).

are critical of the theory of Altaic languages (Georg, 2004, 2005, 2011, 2013; Vovin, 2005, 2009, 2011). [Vovin initially published in favor of the theory and Georg - while remaining skeptical - found some criticisms of the theory problematic (Georg et al., 1999; Manaster Ramer et al., 1998)].

There are also experts who examine the Altaic languages in the context of language evolution research in general and try new methods outside the classical Altaicist debates. As an example, I would like to mention two studies by N. Hübler. In one of his articles, N. Hübler took the Altaic languages as a sample group in the debate on whether structural features can provide data on evolutionary speed. In this study, Hübler used the Hidden Markov Model to calculate the evolutionary rate of structural features coded for 12 Japanese, 2 Korean, 14 Mongolian, 11 Tungus and 21 Turkic languages, and found that there is a correlation between phylogenetic signal and evolutionary rate, that overall two-thirds of the features have a high phylogenetic signal, and that more than half of the features evolve at a slow rate. Accordingly, argument marking, derivation and valency appear to be the most stable functional categories; pronouns and nouns the most stable lexical types; and phonological and morphological levels the most stable language domains (Hübler, 2022). These results may provide data for the debate on the use of stable items as a measure of language relatedness. However, I should remind you that other researchers have also raised objections to the allegedly stable items. I will address G. Longobardi's objection below.

N. Hübler conducted another study with S. Greenhill, one of the world's foremost experts on linguistic diversity, comparing the Altaic languages in terms of morphological and syntactic features (Hübler & Greenhill, 2023). In this article, the authors identified 60 languages from Turkic, Mongolian, Tungus, Korean and Japanese language families and coded 224 language features from their phonology, morphology and syntax. The findings of the study related to Turkic languages are as follows:

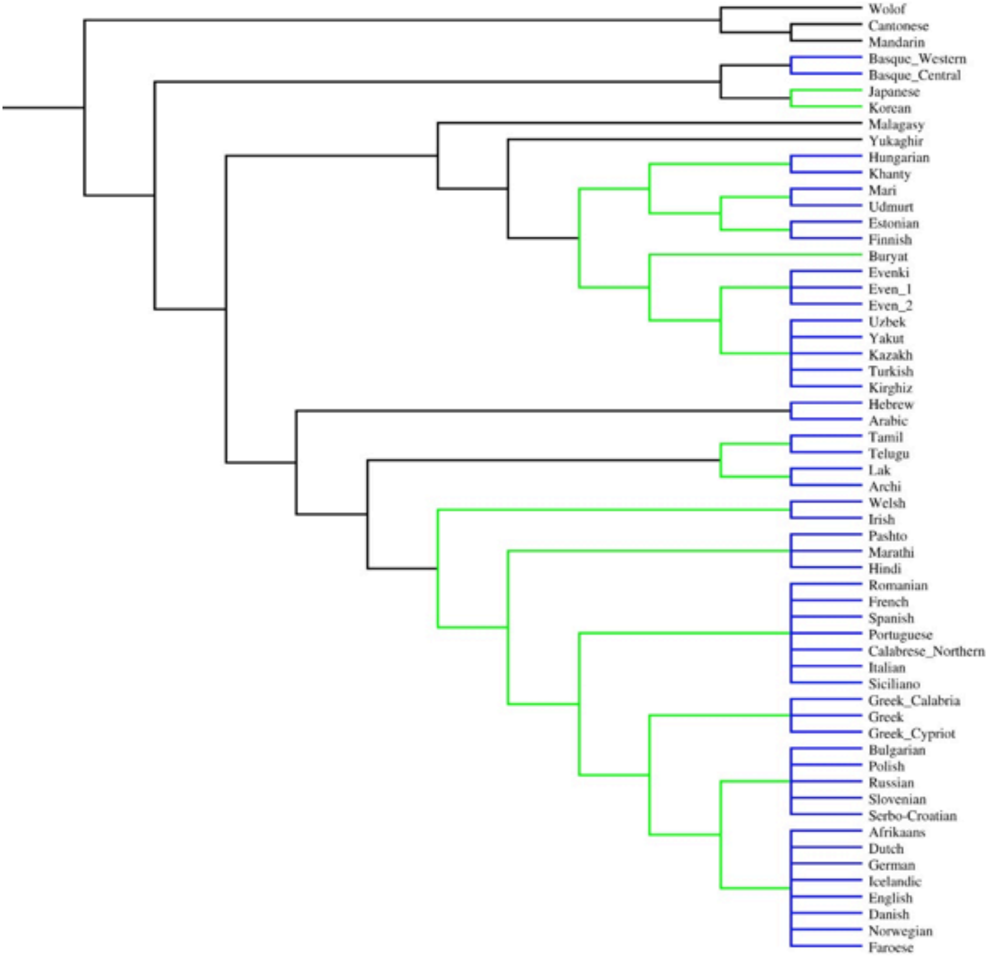
> *The Turkic languages stand out as a cluster with the same dominant ancestry, apart from several exceptions, on all language levels. All Turkic languages, except for Chuvash (30% of 'Mongolic' and 18% of 'Tungusic' ancestry), show the lowest levels of admixture at the morphological level. At the phonological level, several Turkic languages show the highest proportions of 'Mongolo-Koreanic' ancestry among all Turkic languages (in descending order: Chagatai 51%, Northern Uzbek 43%, Chuvash 29%, Tuvan 27%, etc.). At the syntactic level, Northern Siberian languages, Dolgan and Yakut, and a South Siberian language, Tuvan, are the languages with the highest admixture levels (more than 65%). In particular, Dolgan and Yakut have a high proportion of 'Mongolic' (47% and 49%, respectively) and 'Tungusic' (12% and 24% respectively) ancestries, Tuvan has a high*

> *proportion of 'Mongolic' (29%), 'Tungusic' (16%), and 'Japono-Koreanic' (28%) ancestries* (Hübler & Greenhill, 2023).

The study shows that morphological features have the strongest lineage signal and syntactic features propagate most easily. "Our analysis shows that morphological features have the strongest genealogical signal and syntactic features diffuse most easily. When using only morphological structural data, the model is able to correctly identify three language families: Turkic, Mongolic, and Tungusic, whereas there are not enough structural dissimilarities between Japonic and Koreanic languages to assign them to different ancestries." According to the authors, morphological features have a precise historical signal, which can be used to establish relationships between other language families with no known relatives due to the time limitations of the comparative method. However, no definitive conclusion has been reached that the languages considered in this study have the same origin.

The discussion has largely focused on common lexical and morphological elements. Although there are commonalities, the question remains as to why usage varies widely (de Rachewiltz & Rybatzki, 2010: 351-353). Janhunen et al. conducted a study on the Wutun language in China and found that due to influences, this language has taken on a structure whose family cannot be identified (Janhunen et al., 2008). This finding has shown that different languages (Turkish, Mongolian, Chinese and Tibetan languages, etc.) can influence each other to such an extent that they can completely change their old characteristics in a short period of time, providing support for the view of the opponents of the Altaic languages theory that emphasizes the interaction between languages. Nevertheless, it is possible to say that the mainstream of the theory of Altaic languages is still dominated by phonological and especially lexical comparisons [For a few recent examples, see. (Ceolin, 2019; Egorov et al., 2022; Kassian et al., 2021; Turchin et al., 2010)].

On the other hand, the inadequacies of the method based on sound and lexical comparisons have been questioned more and more in recent years, and it can be argued that different methods should be used for the Altaic languages. For example, there is a hypothesis that syntax may be more decisive in determining language relatedness, and G. Longobardi and his team have done remarkable work on this subject (Ceolin et al., 2021; Colonna et al., 2011; Guardiano & Longobardi, 2005; Longobardi, 2003; Longobardi et al., 2013, 2015; Longobardi & Guardiano, 2009, 2017; Santos et al., 2020). Another study analyzing the syntactic features of Altaic languages from a phylogenetic point of view was conducted by Longobardi et al. *At the Boundaries of Syntactic Prehistory* used a statistical method based on syntax.

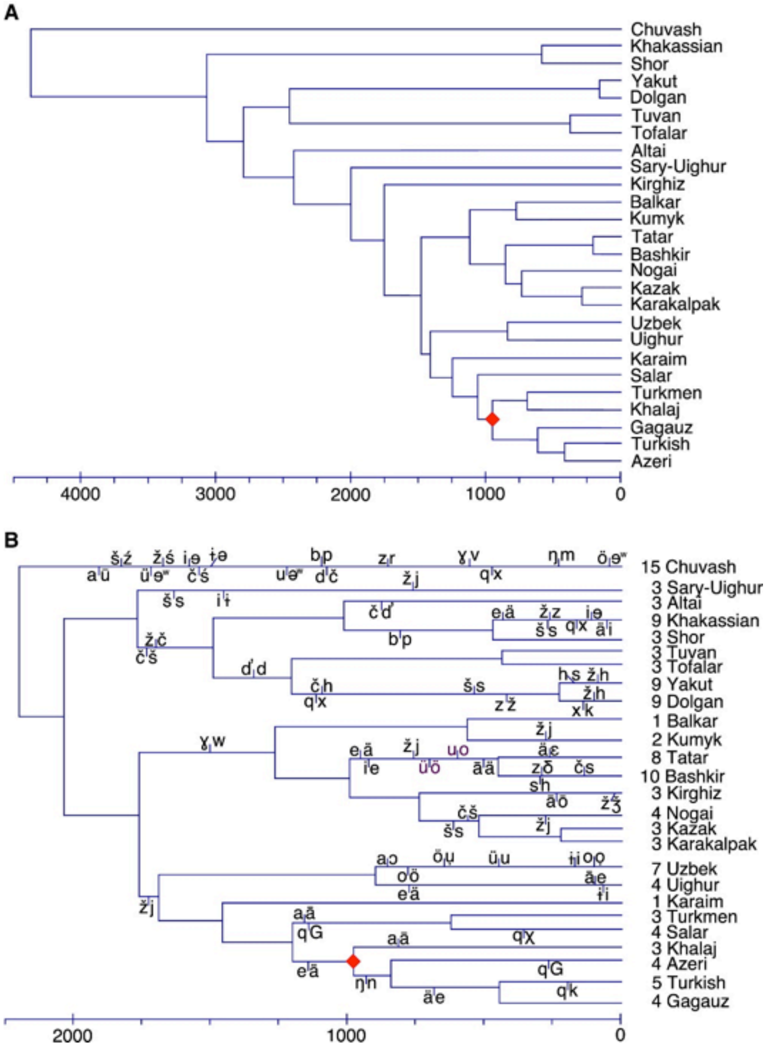


Longobardi et al. show the closeness of the language families they analyzed from a syntactic point of view with this drawing.

(Ceolin et al., 2021).

Accordingly, statistical tests have revealed that the syntax of known (Indo-European, Uralic) or suspected (Altaic) related languages exhibit some similarities that cannot be explained by chance. The tests also support a deeper relationship between the Uralic (Finno-Ugric) and Altaic languages, as suggested in the literature. However, "the tests so far do not provide any support for an Indo-European- Uralic or macro-Altaic unit "(Ceolin et al., 2021). The lack of syntactic support for the hypothesis of a macro-Altaic language family including Japanese and Korean is significant. As can be seen in the tree diagram above, Japanese and Korean are close to Basque, while the micro-Altaic group is closer to each other.

The last review I will discuss is based on Hruschka et al.'s adaptation of a model based on evolutionary variation in genetics to sound variation in Turkic languages (Hruschka et al., 2015). Accordingly, sound transformations in a language family (/p/ > /f/, /d/ > /y/, etc.) are similar to genetic variation and therefore sound changes can be studied by statistical genetic modeling.



(Hruschka et al., 2015).

The results obtained by analyzing the sound changes according to this model are summarized in the tables above. Accordingly, Model A represents a structure that allows irregular sound changes, while Model B represents a structure that allows regular sound changes. Regular sound changes are highlighted at the top and bottom of the tree branches: Events in black indicate the direction of change of the initial sound, while events in purple indicate two sounds that replace each other. The model also identifies the position of each regular sound change. According to this study, the common ancestor of Chuvash and other Turkic languages dates back about 4500 years: "The model additionally estimates the position of each regular sound change along the branch. Mean estimated age of root between Chuvash and other Turkic languages: sporadic model (A) = 2408 BCE, with 95% credible intervals of 3993–1279 BCE; regular model (B) = 204 BCE, with 95% credible intervals of 605 BCE–81 CE. The posterior date of the calibration node (red dot; [18, 19]) is 1017 ± 20 CE" (p. 4). As can be seen, the date estimates for the common ancestor of Turkic languages go back 4-5 thousand years, and we will see similar estimates in the studies I will discuss in the next section.

In conclusion, I can summarize this chapter as follows: Discussions on the theory of Altaic languages have focused on sound correspondences, morphology and lexical similarities. In recent years, syntax has also been used in comparisons. Although this research based on linguistic features has made great progress, it does not yet seem to have ended the debate. Opponents of the theory continue to argue that the proposed linguistic features do not firmly support the hypothesis of common descent. Those who believe that these languages share a common origin, on the other hand, argue that there is an undeniable commonality of structure and vocabulary between the languages.

What Does Genetic Research Tell Us About the Kinship of Turkic Speakers?

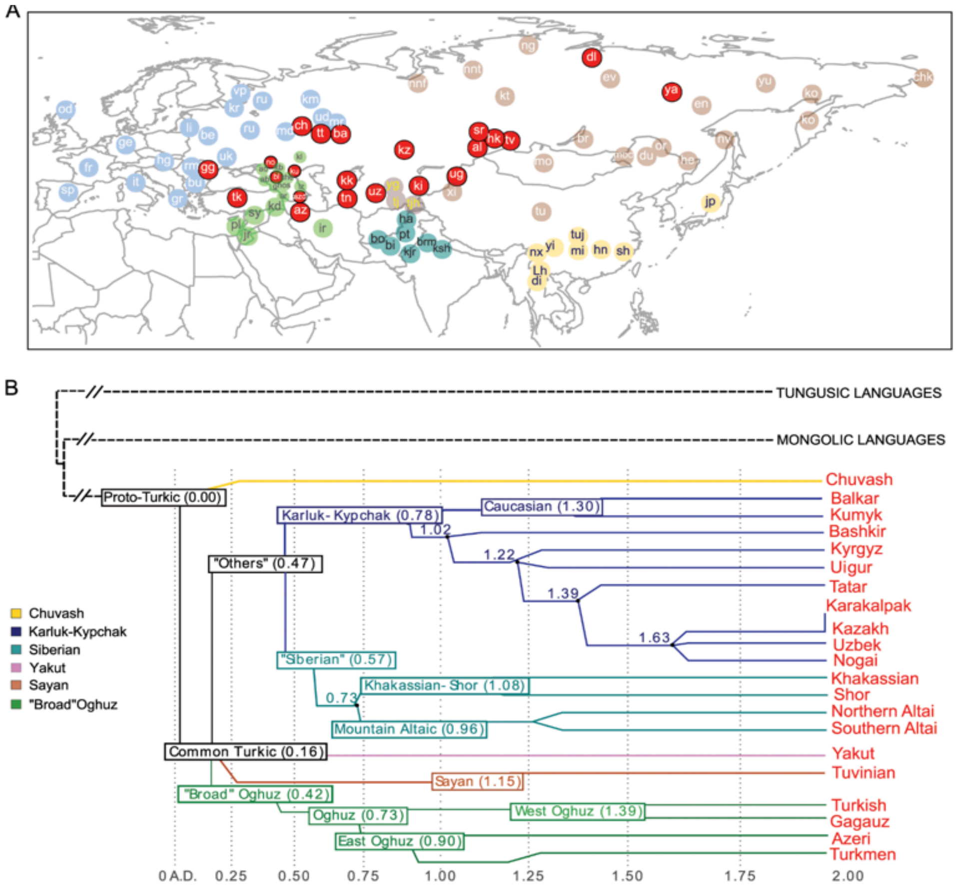
I have not been able to identify any scientific publications in Turkey that are a collaboration between geneticists, archaeologists and language researchers. There are mostly language-independent publications by geneticists. Most of these studies have dealt with gene distribution in present-day Turkey, e.g. (Arnaiz-Villena et al., 2001; Berkman et al., 2008; Cinnioğlu et al., 2004; Di Benedetto et al., 2001; Gökçümen et al., 2011; Hodoğlugil & Mahley, 2012; Kars et al., 2021). The genetics of Anatolia's first settlers, the first farmers, have also been revealed [For the genetics of the first farmers 15 thousand years ago, see. (Feldman et al., 2019)]. Especially the contributions of Mehmet Somel and colleagues are important (Kılınç et al., 2016, 2017; Yaka et al., 2021). However, I have not come across a Turkish publication specialized on the parallelism of language and genes[3], and there are not enough studies written in foreign languages that deal with Turkish in this respect. We

---

[3] A study titled genetics and Altaic languages belongs to Faruk Gökçe: *Genetik Dilbilime Giriş: Altay Dilleri Sorunu* "Introduction to Genetic Linguistics: The Problem of Altaic Languages". In this detailed analysis, Gökçe discusses the dynamics of cultural diversification in historical-comparative linguistic studies through the Altaic languages, not genetic studies based on DNA (Gökçe, 2015). However, this study does not include any data on the genetics of speakers.

can attribute this to the new and interdisciplinary nature of the field.

The Genetic Legacy of the Expansion of Turkic-Speaking Nomads across Eurasia by Yunusbayev et al. is one of the most important studies in recent years that examines the Turkish language through gene and language parallels. Yunusbayev and colleagues took 22 DNA samples from the Turkish-speaking population and aimed to determine whether there is an overlap between the relatedness of these individuals and the relatedness of the Turkic languages they speak. The divergence chart used in the study, which I present below, is adapted from A. Dybo's historical divergence chart based on lexical comparisons.[4]



On map A, Turkic languages of the same lineage are presented in the same colors and the division scheme of these languages is presented in the same colors on map B. The dates of branching are also shown on the map from BC to today. For example 1.63 = 1630, 0.16 = 160.

(Yunusbayev et al., 2015).

---

[4] Anna Dybo first used this genealogical drawing in her book published in Moscow in 2004 (Хронология тюркских языков и лингвистические контакты ранних тюрков, Moscow: Akademiya), and then posted an English version on the website linked below.
"Chronology of Turkic languages and linguistic contacts of early Turks" http://s155239215.onlinehome.us/turkic/40_Language/Dybo_2007LingivistContactsOfEarlyTurksEn.htm (accessed 6.12.2023)

As it can be seen from the map, the Turkic languages spoken mostly in Western Eurasia, such as Anatolia, the Balkans, and the Caucasus, and the Turkic languages spoken in Central Asia exhibit remarkable diversity in terms of geographical distance and the influence of neighbors who speak different languages. Researchers have found that Turkic speakers often exhibit genetic affinity with their neighbors, highlighting the unsurprising role of geographical proximity in genetic admixture. Despite this distance, Turkic speakers have also been found to have a genetic connection:

> *The Turkic peoples represent a diverse collection of ethnic groups defined by the Turkic languages. These groups have dispersed across a vast area, including Siberia, Northwest China, Central Asia, East Europe, the Caucasus, Anatolia, the Middle East, and Afghanistan. The origin and early dispersal history of the Turkic peoples is disputed, with candidates for their ancient homeland ranging from the Transcaspian steppe to Manchuria in Northeast Asia. Previous genetic studies have not identified a clear-cut unifying genetic signal for the Turkic peoples, which lends support for language replacement rather than demic diffusion as the model for the Turkic language's expansion. We addressed the genetic origin of 373 individuals from 22 Turkic-speaking populations, representing their current geographic range, by analyzing genome-wide high-density genotype data. In agreement with the elite dominance model[5] of language expansion most of the Turkic peoples studied genetically resemble their geographic neighbors. However, western Turkic peoples sampled across West Eurasia shared an excess of long chromosomal tracts that are identical by descent (IBD) with populations from present-day South Siberia and Mongolia (SSM), an area where historians center a series of early Turkic and non-Turkic steppe polities. While SSM matching IBD tracts (> 1cM) are also observed in non-Turkic populations, Turkic peoples demonstrate a higher percentage of such tracts (p-values 0.01) compared to their nonTurkic neighbors. Finally, we used the ALDER method and inferred admixture dates (~9th–17th centuries) that overlap with the Turkic migrations of the 5th–16th centuries. Thus, our results indicate historical admixture among Turkic peoples, and the recent shared ancestry with modern populations in SSM supports one of the hypothesized homelands for their nomadic Turkic and related Mongolic ancestors.* (Yunusbayev et al., 2015).

---

[5] The elite dominance model is based on the active presence of a particular group within a larger population for military or other reasons. When this small group imposes its language on the larger population, language shift occurs and the genetic relation between the spoken language and the speakers is severed. Examples of language change based on this model are common, especially in regions with imperial and colonial experience. Renfrew states that the initial spread of the Altaic languages about 10,000 years ago was based on farming and agriculture, and later the spread of the Altaic languages followed the model of elite domination (Renfrew, 1992, pp. 457-459).

This finding of Yunusbayev et al. can be considered as a contribution from the science of genetics to the issues of the ties of today's Turkic speakers and migration from Central Asia. The emphasis on the genetic resemblance of Turkic speakers to geographically close neighboring nations is to be expected as a natural consequence of all neighborly relations. Another conclusion to be drawn from the research is that the regions of present-day Mongolia and Southern Siberia were the starting point of Turkic and Mongolian expansion.

> *Our admixture analysis revealed that Turkic-speaking populations scattered across Eurasia tend to share most of their genetic ancestry with their current geographic non-Turkic neighbors. This is particularly obvious for Turkic peoples in Anatolia, Iran, the Caucasus, and Eastern Europe, but more difficult to determine for northeastern Siberian Turkic speakers, Yakuts and Dolgans, for which non-Turkic reference populations are absent. We also found that a higher proportion of Asian genetic components distinguishes the Turkic speakers all over West Eurasia from their immediate non-Turkic neighbors. These results support the model that expansion of the Turkic language family outside its presumed East Eurasian core area occurred primarily through language replacement, perhaps by the elite dominance scenario, that is, intrusive Turkic nomads imposed their language on indigenous peoples due to advantages in military and/or social organization* (Yunusbayev et al., 2015).

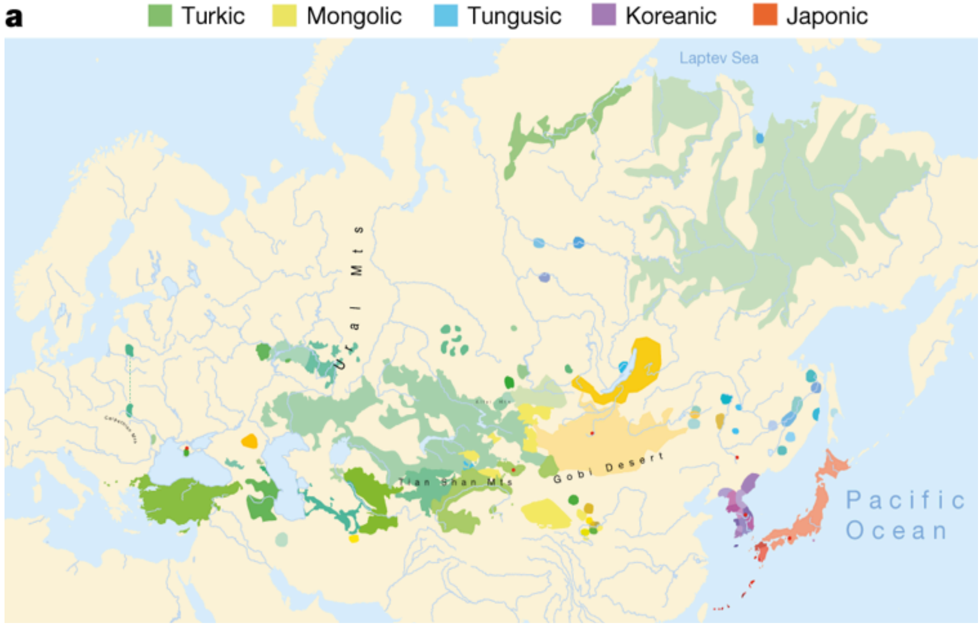
Another noteworthy emphasis of the researchers is on the genetic mixture of Turkic and Mongolian. Yunusbayev and his team, who especially mention the influence of Genghis Khan expansionism as a possibility here, state that their own views are in favor of a common ancestor or migration from a common geography. Their preference for the common ancestor view can undoubtedly be considered as a support in favor of the proponents of the Altaic languages' theory.

Another important feature of this study is the genetic admixture dates of Turkic speakers. For example, the date of admixture of the Chuvash with East Asian/Siberian populations was determined as 816 AD, and the date of admixture of the Nogai as 1657 AD. The findings regarding the Oghuz are as follows:

> *Differences in admixture dates for the three Oghuz speaking populations (Azeris, Turks, and Turkmens) were notable and their geographical locations suggest a possible explanation. Anatolian Turks and Azeris, whose Central Asian ancestors crossed the Iranian plateau and became largely inaccessible to subsequent gene flow with other Turkic speakers, both have evidence of earlier admixture events (12th and 9th centuries, respectively) than Turkmens. Turkmens, remaining in Central Asia, showed considerably more recent admixture dating to the 14th century, consistent with other Central Asian Turkic populations and most likely due to admixture with more recent, perhaps recurrent, waves of migrants in the region from SSM* (Yunusbayev et al., 2015).

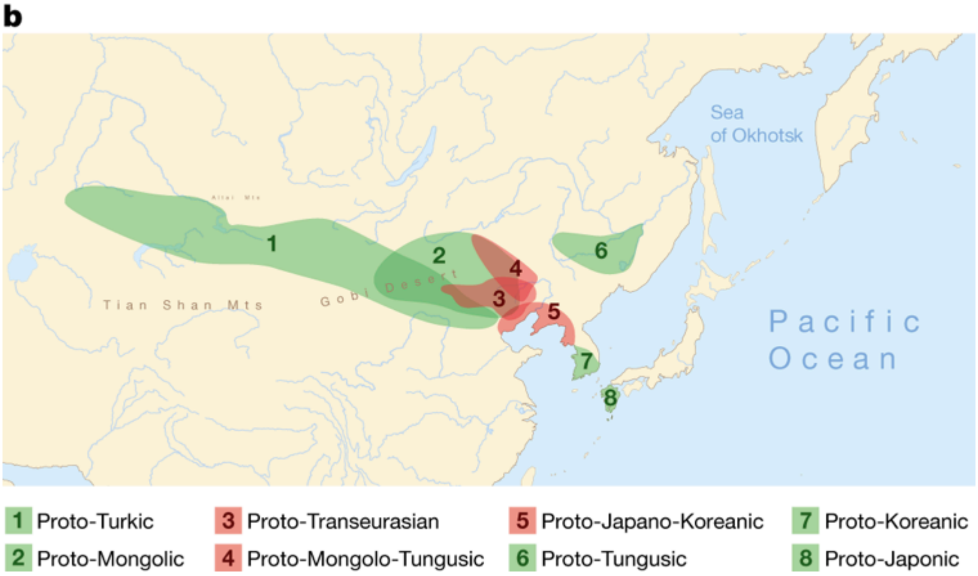
It seems that the Turkmen being closer to Central Asia in the Oghuz group also created a genetic difference. While the genetic mixing of other members of the Oghuz group with Central Asia slowed down over time due to geographical distance, this relationship continued in Turkmens. This genetic data is also compatible with language features. Some features of Turkmen Turkish are different from other Oghuz group languages and closer to Kipchak and Karluk group Turkic languages. For example, while other Oghuz group languages use the verb *ol-*, Turkmen Turkic uses the verb *bol-*, or in examples where the word-initial /b-/ phoneme in Old Turkic in the Oghuz group turns to /v-/ (*bar-* > *var-*), Turkmen Turkic preserves the word-initial /b-/ phoneme: *bar-*. These differences in Turkmen Turkish can be explained by the continued connection with Central Asian Turkic languages.

Triangulation Supports Agricultural Spread of the Transeurasian[6] *Languages*, written by a large team led by Martine Robbeets and published in the journal *Nature,* is one of the most acclaimed and controversial studies in recent years (Robbeets et al., 2021). I mentioned above that data from the triad of genetics, archaeology and linguistics was presented by C. Renfrew under the label of "new synthesis". Robbeets and his team use the term *triangulation* for the method that incorporates data from these three fields. In the study, different data such as the genetic characteristics of speakers and agricultural culture are blended with linguistic features to argue that Altaic languages have a common origin. The map below shows the current (a) and past (b) distribution of languages of common origin.



---

[6] Martine Robbeets uses the term *Transeurasian* instead of Altaic for the quintet group consisting of Turkish, Mongolian, Manchu-Tunguz, Korean and Japanese. She uses the term *Altaic* for the group of three consisting only of Turkish, Mongolian, Manchu-Tungusic.



(Robbeets et al., 2021, p. 617)

The researchers used Bayesian computation to evaluate a dataset of 3,193 root sets containing 254 basic lexical concepts from 98 different Altaic languages and dialects. They calculated that the root of the Altaic language family dates back 9181 years.[7] They used lexical data sets and Bayesian phylogeography methods to model the spatial expansion of the language and found that the Altaic language family originated in the Western Liaonin River region, from the Altai Mountains to the Yellow River. This expansion began in the Early Neolithic period and continued into the Late Neolithic and Bronze Age. In addition, the researchers tried to understand the prehistoric lifestyle of this language family by analyzing words related to agriculture and animal husbandry (p. 617). Accordingly, the findings support the agricultural hypothesis and exclude the animal husbandry hypothesis.

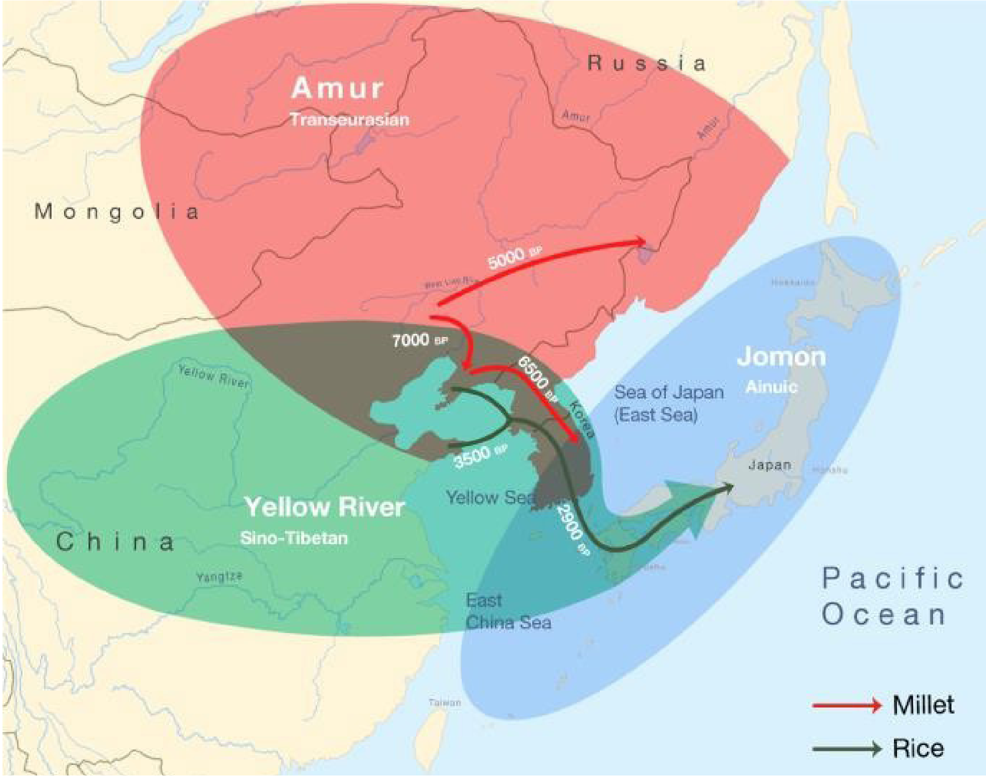
In archaeology, Robbeets et al. extracted data from published literature and identified 172 archaeological features for 255 Neolithic and Bronze Age sites and added carbon-14 dated early crop remains from northern China, Primorye, Korea and Japan to create a set of 169 items. Using this set, they found that millet cultivation was practiced in the region about 9,000 years ago and concluded that there was a correspondence between the spread of agricultural culture (millet and rice) and the spread of Altaic communities (pp. 617-618).

---

[7] "*Our results indicate a time-depth of 9181 BP (5595–12793 95% highest probability density (95% HPD)) for the Proto-Transeurasian root of the family; 6811 BP (4404–10166 95% HPD) for Proto-Altaic, the unity of Turkic, Mongolic and Tungusic languages; 4491 BP (2599–6373 95% HPD) for Mongolo-Tungusic; and 5458 BP (3335–8024 95% HPD) for Japano-Koreanic. These dates estimate the time-depth of the initial break-up of a given language family into more than one foundational subgroup*" (p. 617).

In the genetic dimension of the research, genome analyses of 19 confirmed ancient individuals from the Amur, Korea, Kyushu, and Ryukyu regions were reported; these were combined with published genomes going back 9500 to 300 years and covering the Eastern steppes, Western Liao, Amur and Yellow River regions, Liaodong, Shandong, Primorye and Japan. This historical genetic dataset was then subjected to primary component analysis with 149 present-day Eurasian populations and 45 East Asian populations. As a result, traces of lineage preserved in the Amur region can be traced in all Altai- speaking populations, suggesting a common genetic basis (pp. 618-619). In other words, according to this study, Altaic speakers are genetically related.

Based on the findings of the research, the following diffusion hypothesis is presented:



The Amur lineage is marked in red, the Yellow River lineage in green and the Jomon lineage in blue. The red arrows indicate the eastward migration of millet farmers during the Neolithic, bringing Korean and Tungusic languages to the indicated regions. Green arrows mark the integration of rice farming in the Late Neolithic and Bronze Age, bringing the Japanese language to Japan via Korea.

(Robbeets et al., 2021, p. 620)

9 thousand years ago, a group engaged in millet production in the Western Liao River region started to migrate to different regions about 7 thousand years ago. As shown in the map above, the group that migrated further north and west (to the Amur region) formed the ancestor of the Turkic-Mongolian- Manchu-Tungus union. During the Late Bronze Age (1600-1200 BC), Altaic languages witnessed extensive cultural exchange in the Eurasian steppes, with populations from the Western Liao region and the Eastern steppes mixing with Western Eurasian genetic lineages. Linguistically, this interaction led to the borrowing of agro-pastoral vocabulary by Proto-Mongolian and Proto-Turkic speakers, especially related to wheat and barley cultivation, animal husbandry, dairy farming and horses. Around 6500 years ago, another group speaking a language that is the ancestor of today's Korean and Japanese migrated further east, diversifying millet cultivation with rice. This branch gave rise to Korean and Japanese [It is also noteworthy that Robbeets et al. have written a book on the relationship between Altaic languages and agrarian culture (Robbeets & Savelyev, 2017)].

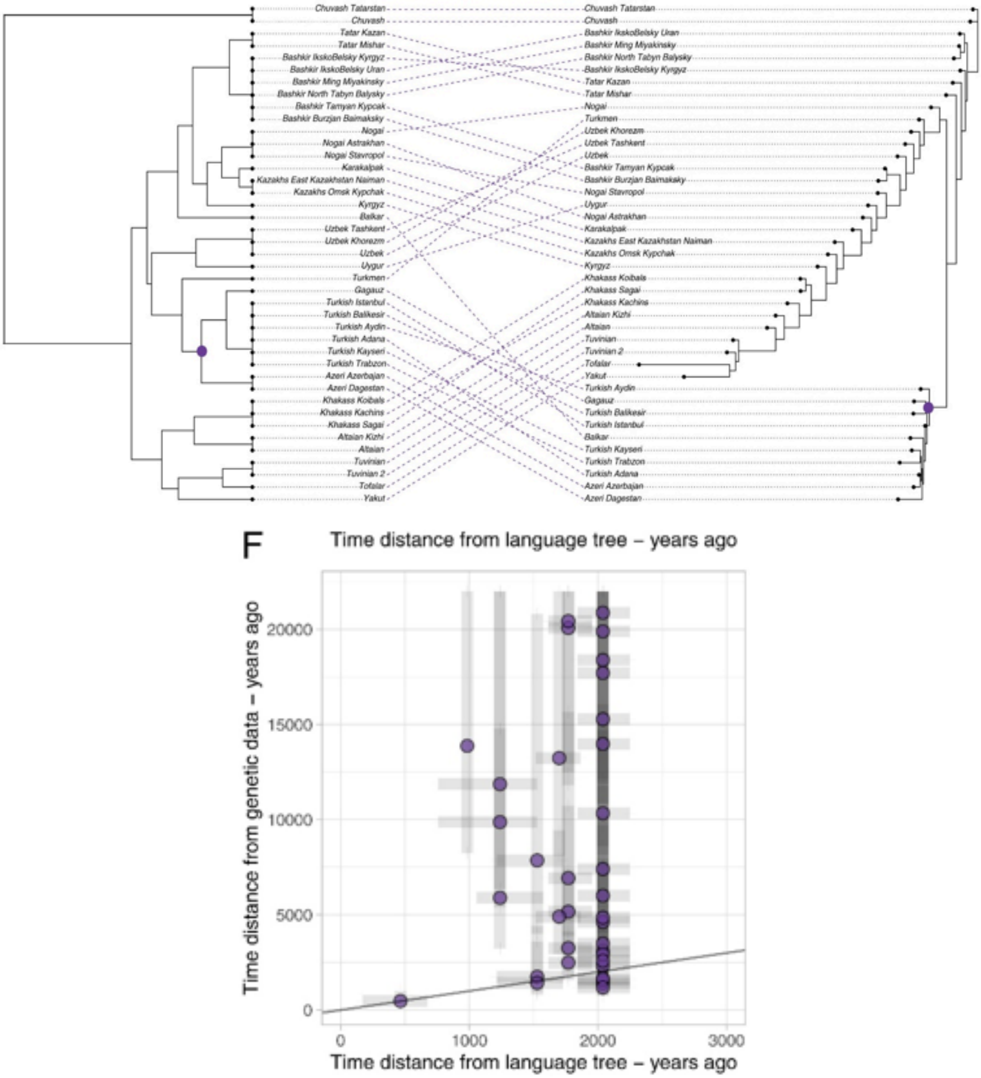
This study is pioneering in terms of using linguistics, archaeology, and genetic data together. No study of this depth has been done yet. However, it was not immediately accepted in the scientific world and received criticism for both the word lists used and the interpretation of archaeological and genetic data [For an example, see. (Tian et al., 2022)].

Another important study was published in 2022 by a team of researchers including S. Greenhill, R. D. Gray and C. Barbieri, who are leading figures in the field of linguistic diversity: *A Global Analysis of Matches and Mismatches between Human Genetic and Linguistic Histories*. This study does not only focus on Turkish but compares the genetic heritage of all languages and their speakers.

The GeLaTo dataset was used in the study. It contains genetic and linguistic information on more than 4,000 individuals representing 397 genetic populations speaking 295 languages. The first striking result is that "for most populations in the entire dataset, the closest genetic neighbor belongs to the same language family". However, a non-negligible 18% of discrepancies were found. In other words, language and genetic commonality overlap to a large extent, but there are mismatches in about one-fifth of the languages studied. According to the researchers, these mismatches are not rare outliers, but a regular consequence of language history. The Hungarians are cited for mismatches between their language and genes, and there are also similar cases in the Caucasus. Accordingly, Hungarians, "probably one of the most studied cases of incompatibility", are genetically similar to their Indo-European-speaking neighbors but maintain a separate linguistic identity as a member of the Uralic language family. "The Hungarian population preserved the language brought by the Magyars, who conquered the Carpathian Basin in the ninth century CE (21–23), while becoming genetically assimilated to their Indo-European–speaking neighbors through time. In our dataset, they are the only case of a linguistic enclave". (Barbieri et al., 2022). According to the study, the situation in Malta and the

Caucasus are also examples of mismatches. Accordingly, the Maltese are the only Afro-Asiatic speakers in Europe and represent an example of incompatibility not addressed by the genetic literature. The majority of Maltese people speak an Afro-Asiatic language with Italian and English influences and therefore their language does not match with the Afro-Asiatic language family. Their genetic profile shows close genetic similarity to Eastern Sicilian, while sharing genetic similarity with Indo-European speakers in the Balkans and geographically distant Turkish and Middle Eastern Afro-Asiatic speakers. This is also the case for Armenians in the Caucasus. The mitochondrial genetic profiles of Armenian (Indo-European speakers) and Azerbaijani (speakers of Turkic languages) show a closer resemblance to surrounding Caucasian populations (grammatically unrelated). In our dataset, the Azerbaijan-speaking population is indeed misaligned with other Turkic speakers, confirming the proposed mismatch. By contrast, Armenians show an FST distribution aligned with that of other Indo-European speakers of Anatolia, challenging the idea that they shifted their linguistic affiliation." (Barbieri et al., 2022).

Stating that it is still unclear whether these inconsistencies are the norm or the exception within language families or whether language families in general show genetic diversity, the researchers compared genetic profiles at the level of language families to answer this question. It was emphasized through the examples of different language families that discrepancies may arise for different reasons. One of the striking points in the study is the research on genetic divergence and the divergence of languages. In genetics, historical divergence is also determined by genetic methods. The researchers remind that the methods of language dating in linguistics are controversial and that two methods are generally used by linguists. The first one is the method based on the quantitative analysis of linguistic, archaeological and historical data, which I have discussed above. The second method, which is also used by researchers in their studies, is the Bayesian language dating method [For this method, see. (Rama & Wichmann, 2020)]. A detailed analysis of the three major language families is given in order to examine the congruence and incongruence between the dates of the genetic separation of speakers from their relatives and the dates of the divergence of languages: Indo-European, Austronesian and Turkic. (Barbieri et al., 2022). The gene and language overlap within these three language families is presented in tables. The table for Turkish is as follows:

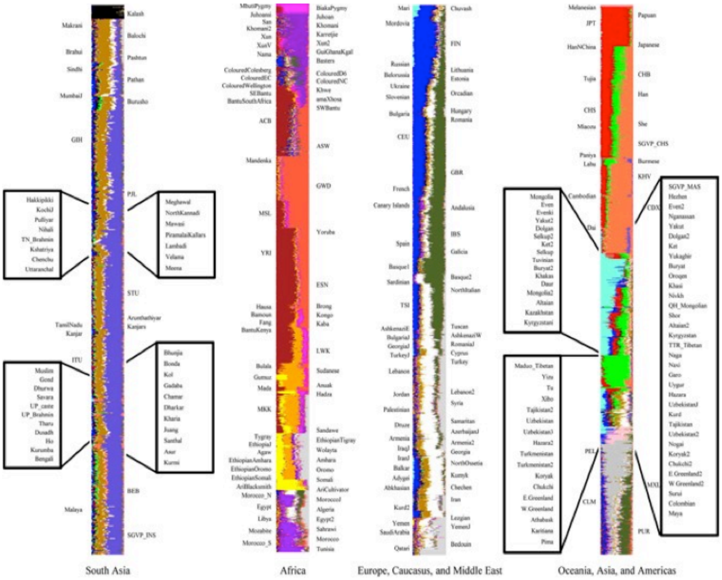


(Barbieri et al., 2022).

Comparisons within these three language families show that "The highest similarity metric is found for the Indo-European trees (0.68), followed by the Austronesian (0.65) and Turkic (0.57) trees". In other words, the dates of the genetic divergence of Indo-Europeans and the dates of the divergence of the languages they speak are more similar. Here is the following finding for Turkish: "The Turkic tree does not show relevant correspondences in the divergence times, with genetic divergence time much older than the linguistic divergence time". According to this result, the genetic divergence of speakers of Turkic languages is older, while the divergence time of Turkic languages is more recent.

In a detailed review entitled *Human Ancestry Correlates with Language and Reveals*

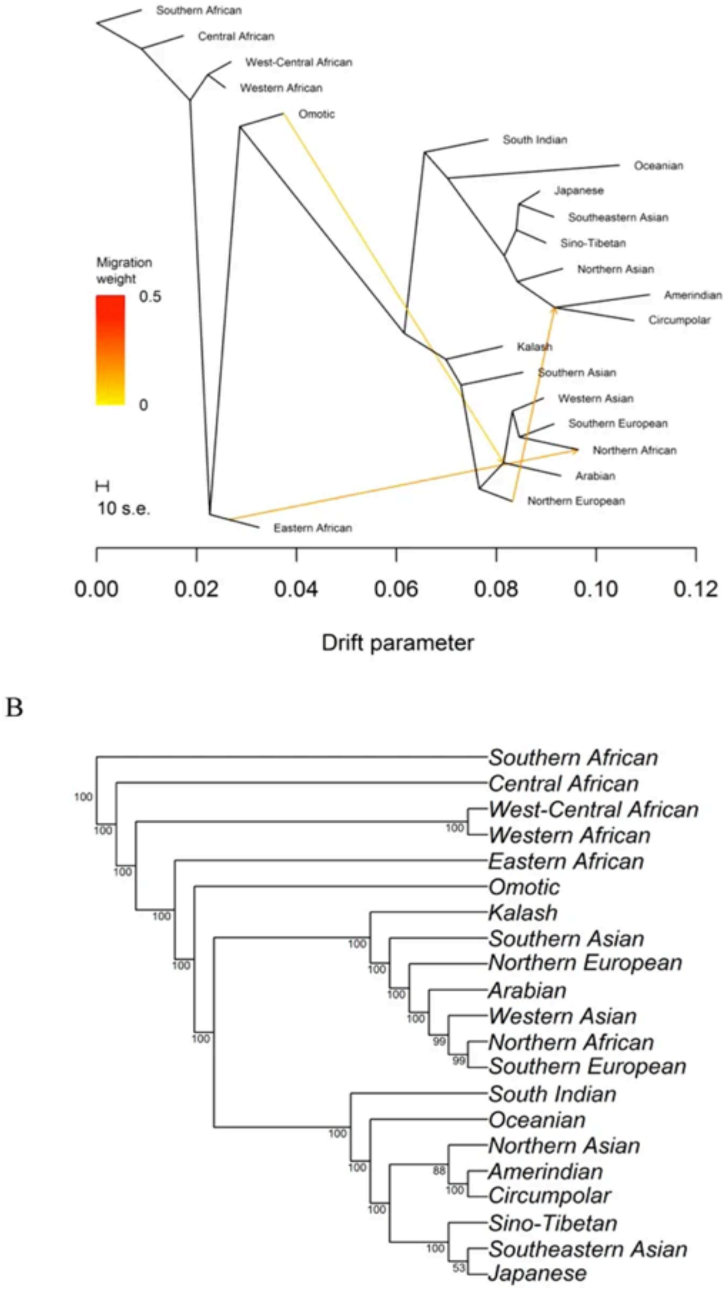
That Race Is Not an Objective Genomic Classifier, J. L. Baker et al. used a large dataset to investigate the genetic differentiation of early modern humans, human admixture and migration events, and the relationships between ancestry and language groups. After compiling genotype data on 5,966 individuals from 282 global samples representing 30 major language families, they concluded: “The best evidence supports 21 ancestries that delineate genetic structure of present-day human populations. Independent of self-identified ethno-linguistic labels, the vast majority (97.3%) of individuals have mixed ancestry, with evidence of multiple ancestries in 96.8% of samples and on all continents.”. The data indicate that continents, ethno-linguistic groups, races, ethnicities, and individuals all show substantial ancestral heterogeneity. We estimated correlation coefficients ranging from 0.522 to 0.962 between ancestries and language families or branches. Ancestry data support the grouping of Kwadi-Khoe, Kx’a, and Tuu languages, support the exclusion of Omotic languages from the Afroasiatic language family, and do not support the proposed Dené Yeniseian language family as a genetically valid grouping. Ancestry data yield insight into a deeper past than linguistic data can, while linguistic data provide clarity to ancestry data. The data show significant ancestral heterogeneity. Correlation coefficients ranging from 0.522 to 0.962 were found between ancestry and language families or branches.



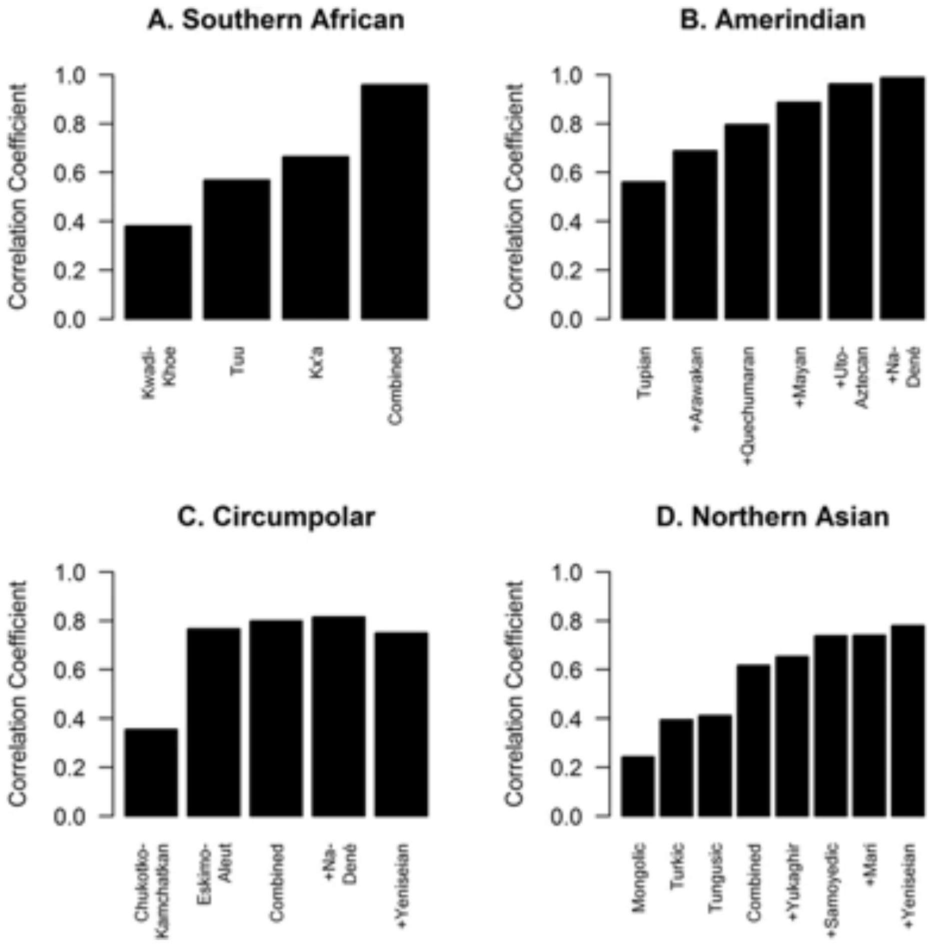
The 21 ancestral lines identified in the study are: Kalash (black), South Asian (ochre), South Indian (slate blue), Central African (magenta), South African (dark orchid), West-Central African (brown), West African (tomato), East African (orange), Omotic (yellow), North African (purple), Northern European (blue), Southern European (dark olive green), West Asian (white), Arabian (light gray), Oceanic (salmon), Japanese (red), Southeast Asian (coral), North Asian (aquarium blue), Sino-Tibetan (green), Circumpolar (pink) and Native American (gray).

(Baker et al., 2017)

The graph of the migrations started by the Homo sapiens species 150-200 thousand years ago from Africa, including these 21 ancestral lineages, is presented as follows:



(Baker et al., 2017).



(Baker et al., 2017).

(D) Mongolian, referred to as Altaic, refers to Turkish and Tungusic.

The article states the following conclusion about the Altaic languages: “Northern Asian ancestry correlates with Mongolic, Turkic, and Tungusic languages ($r = 0.617$, $p = 1.53 \times 10^{-27}$), which have been grouped into the Altaic language family. Additionally, Northern Asian ancestry correlates with the Samoyedic branch of the Uralic family, Yukaghir languages, the Mari language isolate, and Yeniseian languages ($r = 0.781$, $p = 2.53 \times 10^{-52}$”. The researchers reconstructed the present distribution of Homo sapiens, a species originating from Africa, backwards through genetic links and reached 21 ancestors. Although the study does not comment on single-ancestry language theories such as Eurasiatic and Nostratic, the result supports multi-ancestry theories. However, it should be remembered that 21 ancestral lineages have the potential to be associated with a single origin with further research. In this study, Turkish is seen to be related to the North Asian ancestral lineage; a conclusion is reached in favor of the common origin of the so-called core Altaic group (Turkish, Mongolian, Manchu-Tungusic, etc.). According to the researchers, ancestry data provide insights into a deeper past than linguistic data can provide, while linguistic data provide clarity to ancestry data (Baker et al., 2017).

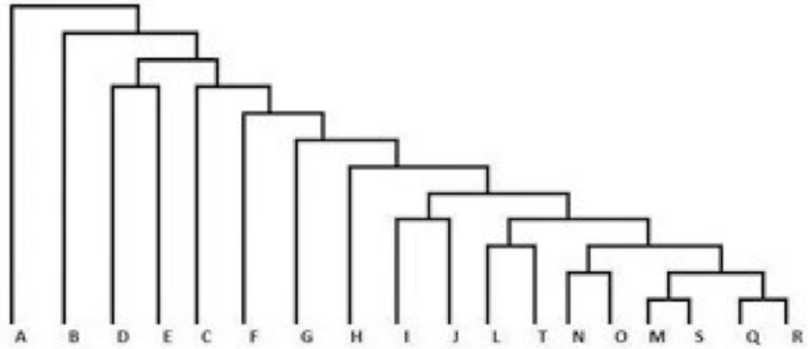
An article entitled *Transeurasian Unity from a Population-Genetic Perspective,* which reviews linguistic and genetic studies of the Altaic languages, was published in *The Oxford Guide to the Transeurasian Languages* edited by M. Robbeets and A. Savalyev (Jeong, Wang, & Ning, 2020). Choongwon Jeong, Chuan-Chao Wang and Chao Ning review the literature and highlight studies showing that Mongolian, Turkic, Manchu-Tungus, Korean and Japanese populations have a heterogeneous genetic population structure. However, they draw attention to publications that argue for the existence of a genetic unity beyond admixture in different time periods due to the influence of Genghis Khan and migrations. According to Jeong, Wang and Ning, these studies show an underlying common genetic substrate among Transeurasian populations, best represented by ancient populations from Northeast China and the Russian Far East, as well as present-day Tungus-speaking populations. The authors characterize Turkish-speaking populations as having the highest genetic heterogeneity as a consequence of having the widest geographical spread (pp. 787-788). Although the view that Altaic-speaking populations are mostly genetically mixed with surrounding populations is repeated, one point is important. Accordingly, despite genetic admixture, there is an underlying commonality in the genetics of the populations speaking Altaic languages. This interpretation is an important consideration for proponents of the Altaic theory.

Genetic studies were also conducted on the Turkic, Mongolian and Tungus languages spoken in China (Xu & Li, 2017). Here, a parallel was found between ancestry and languages. For example, the genetics of Mongolian speakers and the classification of the Mongolian language family overlap. However, comparative gene analysis of the speakers of these three different language families revealed significant differences, with a large genetic distance between Tungusic and Turkish speakers, whereas Mongolian and Tungusic speakers are relatively close (Wen et al., 2017).

In another study conducted on the Silk Road, it was found that there was no parallelism between languages and genes. Speakers of Turkic and Mongolian languages were compared genetically by focusing on the Silk Road geography. As a result, in line with the "dominance of elites model", it was determined that the language spoken by the societies and their ancestral lineages differed, which was explained by the imposition of languages by dominant groups in addition to the cultural diversity of the Silk Road (Xu & Wen, 2017).

Another notable research on Turkic is based on ancient DNA analysis: *A Comparative Analysis of Chinese Historical Sources and Y-DNA Studies with Regard to the Early and*

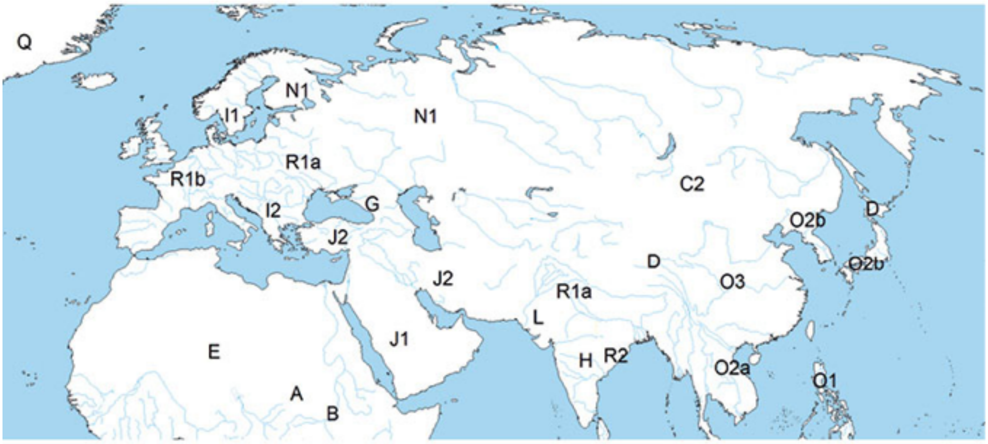
Medieval Turkic Peoples (Lee & Kuang, 2017). In this study, J. Y. Lee and S. Kuang do not comment on linguistic relatedness, but compile descriptions of the Turks in historical texts and compare them with contemporary ancient Y-DNA studies. The Y chromosome is one of the chromosomes that determine male sex. It is passed from father to son and is usually passed down from generation to generation without mutation. When a mutation occurs, this change is passed on to the entire male lineage. These mutations accumulate over time and allow us to trace the relatedness of people. Males who share a certain mutation form a Y- DNA haplogroup. Haplogroups refer to a group of people who have the same mutation and inherited it from a common ancestor. Population geneticists divide humans into more than 20 main groups and many subgroups, 39 of which belong to the main group shared by all males. Communities around the world are classified in this respect as follows:



Simplified y-dna haplogroup tree

(Lee & Kuang, 2017)

The distribution of haplogroups in the world is as follows:



(Lee & Kuang, 2017)

Stating that these haplogroups exhibit diversity in Turkish-speaking communities, Lee and Kuang visualize the haplogroups of Turkish-speaking groups with the map below:

(Lee & Kuang, 2017)

As a result of genetic studies of Turkic communities, researchers have noted that Turkic communities have different Y chromosome haplogroups - although they share certain haplogroups - and provide the following list of groups:

N1c1 between sites,

N (N1b and N1c1), C2, Q and R1a1 among Tuvans, R1a1 and C2 among the Southern Altaians,

R1a1 and C2 among the Kyrgyz, N and R1a1 among the Khakas,

R1a1, J, O3 and C2 among Xinjiang Uyghurs and Uzbeks,

C2, O3 and G1 among the Kazakhs, Turkmen groups include Q, J and R1a1,

Haplogroups J, R1b (R1b-M269), E and G2 are common among Turks and Azeris. This diversity implies that Turkic peoples living in different regions have heterogeneous paternal origins and harbor linguistically Turkicized indigenous elements. It also suggests that the Turkicization of many parts of Eurasia did not necessarily involve the mass migration of Turkic peoples (Lee & Kuang, 2017).

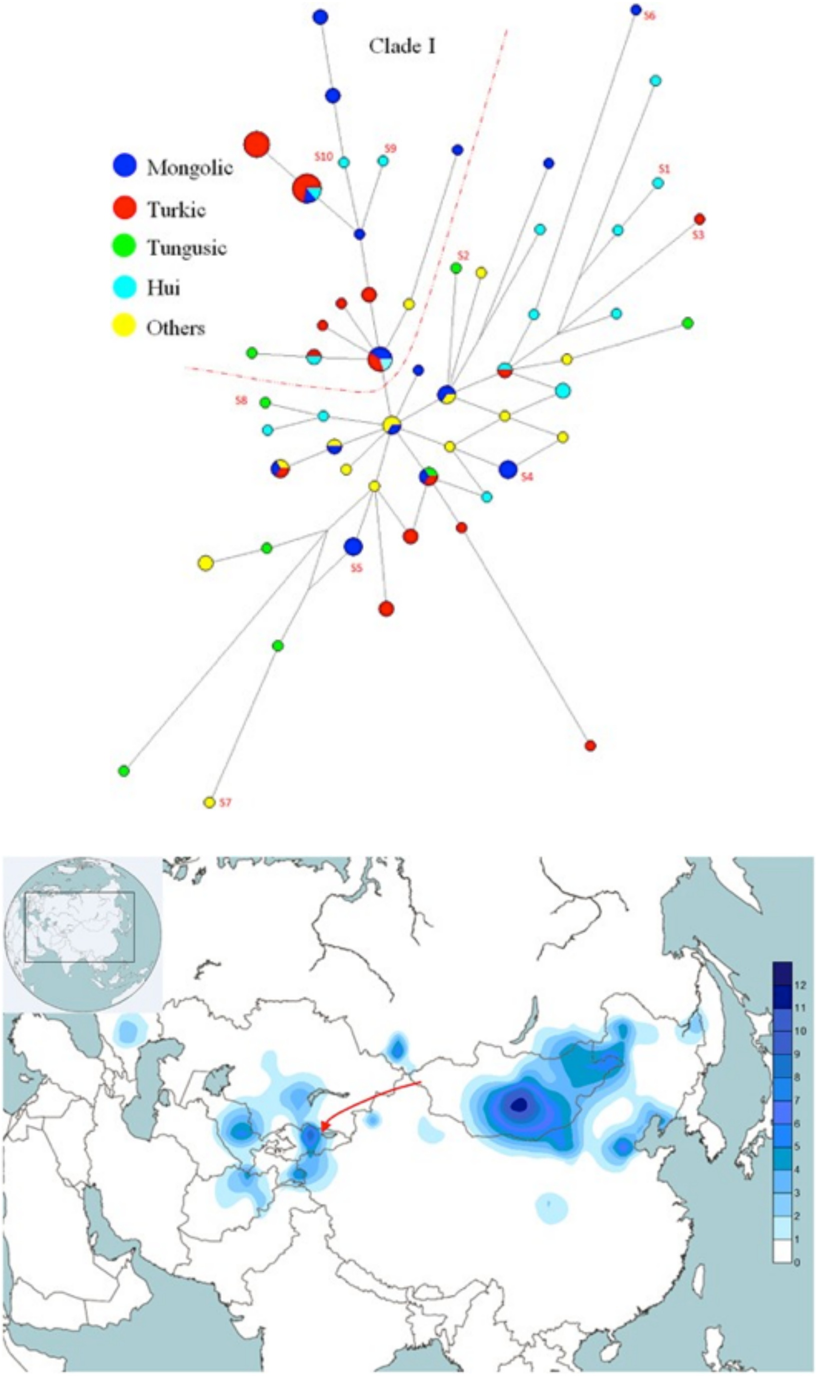
The study also includes DNA analyzes on historical Turkic communities and it is stated that there is a diversity there as well:

Community	***Ancient DNA***	***Possible lineage today***
Xiongnu	C2, N, Q, R1a1 (among others)	
Sakha	N	
Dingling (Tiele) Uighurs		C2, N, O3, Q, among others (if modern descendants are Western Uighurs, Naymans, Tuvans, Sahas and Buryats)
Kök Turks		R1a1 (if related to the Yenisei Kyrgyz), Q, inter alia (if their lineage is Önggüt) N and C2, among others (if their ancestry is Tuvan)
Yenisei Kyrgyz	R1a1 (Yenisei Indo-European pastoralists)	R1a1 (Tien Shan Kyrgyz)
Önggüt	Q (among others)	
Naiman		C3 and O3, among others
Kipchaks		R1b, among others (if their modern descendants is Kazakh Kipchak tribe), C2, among others (if their modern descendants is Kazakh)
Turkmen		Q, J, R1a1 and N, among others

(Lee & Kuang, 2017).

According to the researchers, it is possible that the early and medieval Turkic peoples did not form a homogeneous entity and that some of the non-Turkic peoples were Turkicized at some point in history. It could also be argued that many of the modern Turkic-speaking populations, which exhibit more diverse haplogroup compositions, are not direct descendants of early Turkic peoples. According to Lee and Kuang, this data suggests that Turkish-speaking populations cannot be restricted to a single haplogroup but exhibit diversity resulting from hybridization with different groups.

Recent studies focusing on the Y chromosome, i.e., paternal lineage, have also identified genetic connections among Altaic-speaking populations in Eurasia Malyarchuk vd., 2010; Z. Wang vd., 2024; Wei vd., 2017). For example, in the study *Phylogeny of Y-Chromosome Haplogroup C3b-F1756, An Important Paternal Lineage in Altaic-Speaking Populations* conducted by Lan-Hai Wei and colleagues, this genetic distribution is visualized as follows:



Distribution of the Y chromosome lineage C3-DYS448del across Eurasia* (Wei vd., 2017)

A new Y chromosome study by Z. Wang and colleagues also found that genetic scenarios for paternal lineage and genome-wide diversity in East Asia are strongly linked to language family, and that there is a strong parallel between the Sino-Tibetan, Altaic and South Chinese multilingual families and human expansion. The researchers note that North East Asia has a complex demographic history, with the expansion of nomadic populations and regime changes on the Eurasian steppes strongly determining the patrilineal landscape. The following findings on genetic linkage with Altaic speakers are important:

> *The Altaic language family can generally be divided into Tungusic, Mongolic, and Turkic groups. East Asian populations who speak these languages are mainly distributed in the northern regions with a distinct paternal lineage composition. First, C2-M231 is the dominant haplogroup in Altaic-speaking populations, but there is a stratification of the sublineage composition of C2-M231 within the Altaic-speaking populations distributed in different regions of northern East Asia [44,96]. In general, haplogroup C2-M231 can be divided into two subhaplogroups, C2a-F1396 and C2b-F1067, and the distribution pattern of these two subhaplogroups is region-specific, where C2a-F1396 has a more northern distribution than C2b-F1067 in present-day populations [15,97–99]. Haplogroups C2a-F1396 include three predominant subhaplogroups, C2a1a3-M504, C2a1a1-F3918, and C2a1a2-M48, which are widely distributed in distinct subgroups of Altaic-speaking populations* (Z. Wang vd., 2024, s. 7).

As can be seen, a genetic relationship that cannot be ignored has been detected in the Y chromosome examinations of the speakers of the three core Altaic languages, namely Turkish, Mongolian and Manchu-Tungusic.

What Do These Studies Tell Us?

In the previous sections, I have evaluated the studies on the theory of Altaic languages, which is the most prominent explanation in the debate on the cognates of Turkish. I would like to remind you that there are points to be considered for the question of what we can learn from these studies.

First of all, I have already mentioned the analysis based on linguistic features. Here it is useful to keep in mind a limitation concerning all languages. Today, going to the roots through language features can be interpreted as a near-impossible endeavor. Because the oldest data on these languages is limited to writing. For example, the earliest written products of Turkish date back to the 8th century AD. We do not have a rich linguistic material on the “dark” periods before Christ. Therefore, researchers can only act on the limited data they compile from historical texts. They also make inferences about the past with comparative data from today’s languages. One may wonder if Darwin and Wallace did not come up with the idea of evolution by natural selection using the same method. However, comparing the morphology of living things and cognitive features such as language require different methods. The famous saying reminds us of a difficulty: Language leaves no fossils.

In the above publications we have also seen different attempts at new methods. Darwin's theory later on, with support from genetics, became one of the most robust explanations in the scientific world, and perhaps the closest to the "theory of everything" debate. Of course, the boundaries of science are expanding every day. Perhaps we will reach clearer explanations with new methods that will exceed the limits of sound and word comparisons centered on cultural diversification. However, it is important to support the data based on sound and word comparisons with data from different fields such as archaeology and genetics. Because the Wutun example shows that even the words we call basic words can be transferred to other languages through interaction. Methods based on syntax and other structural features have not been sufficiently applied to Altaic languages. Moreover, when we talk about the "origin of a language", it is not easy to give a single date, because even if we identify the "ancestor language" from which today's languages originated, it is not difficult to guess that this ancestor language originated from another language. Just as it is very difficult to identify the first species in biology, it is also very difficult to find the "first", the "origin" of languages. We cannot answer questions such as who was the first Sapiens or who was the first individual that we would recognize as Sapiens with our current findings. According to the fossil record, we make a conclusion based on the time period in which the first individuals showing the characteristics of the Sapiens species are concentrated. Today, this date is approximately 300 thousand years ago (Hublin et al., 2017; Richter et al., 2017). But there is also the possibility that languages are older than this date. According to a group led by Chomsky, human languages should have emerged between Sapiens and 200,000 years ago, and in his recent publications, Chomsky has even moved this date closer to the last 50,000 years (For discussions, see Kerimoğlu, 2021). 50 thousand years ago is not a depth of history that we can reach with today's linguistic data. Therefore, it must be remembered that there are limits to the data that can be accessed through sound, lexical or syntactic features. Moreover, 50 thousand years is the "closest" estimate. Today, the state of the art in the study of the evolution of language is that this date goes back much further. The majority of researchers believe that the roots of language go back millions of years to more archaic human species such as Neanderthals, Heidelbergensis and Erectus. In fact, there is increasing evidence that there is a common history of language with other primates (I have discussed the current data on the origin of language in the following series of articles, and these discussions can be found there: Kerimoğlu, 2016, 2017, 2018, 2019a, 2019b, 2022). Another claim is that language begins with gestures and facial expressions before words, that is, the roots of language should be sought in gestures. (Arbib & Rizzolatti, 1996; Corballis, 2003). In conclusion, it is important what we mean by finding the roots of today's languages. We will probably never find the roots of any of today's languages through sound or lexical comparisons. Because they will only give us the previous "ancestor language". But the ancestral language we find as an ancestor also originated from another language. For evolutionary progress in language, we need a "philological" tool similar to what genetics does in biology. Current tools, at best, take us back 8-10 thousand years, which, as I mentioned above, is a long way from the roots of the language tree, and can only take us to a few branches at the top of the tree.

There are also problems with the use of genetic tools. The first of these is that examples of language switching break the overlap between language and genes. "Language switching" events, especially in periods and regions that we cannot trace in writing, are the biggest cause of the mismatches in the relatedness of today's language families and speakers. Although the above studies tell us that this incompatibility is approximately 1/5 for all world languages, a study covering all world languages, i.e. 7 thousand languages, has not yet been conducted. This rate may increase even more. People have changed their languages for many different reasons. Although this phenomenon seems unusual to us today, it was much more natural for prehistoric people:

Throughout history, human societies have donned new languages like new cloaks. The linguistic metamorphosis always went unnoticed – until there was writing (Fischer, 2013, p. 79).

In his wonderful book *The History of Language*, S. R. Fisher expresses this in a very striking way. Genetics gives us information about kinship that no other science can give us. Our history is written in our genes. Unquestionably, the neighborhoods, affinities, and unions that gave rise to us can be revealed through genes. However, we cannot say today that our DNA or other biological characteristics provide definitive results for language. Human societies have mostly lived side by side. The borders of the modern world are new. To the extent that humans have been able to overcome natural boundaries, they have reached other places, other people, and have become a "hybrid" species in every field from biology to culture. Human beings are essentially hybrids. This is an undeniable fact for Sapiens as a mixed species. As I mentioned in my book on Neanderthals, today we all have the genes of another species, Neanderthals, and there are even theories that today's Sapiens languages have traces of Neanderthal language or languages (Kerimoglu, 2020). It is emphasized in almost every genetic study that the communities speaking Altaic languages mixed with neighboring peoples. We can say that there is a consensus that these languages, which are spoken over a wide geography, spread especially through elite domination. However, there are also studies suggesting that there is a genetic link between the speakers of these languages. In fact, both of these views may not be wrong.

Speakers of Altaic languages who traveled to different regions mixed genetically with the people of that region. Genetic mixing has already been identified through DNA examinations. Today, it may be difficult to find a genetic commonality between an Evenki speaker and a Gagauz speaker. As a result of generations of mixing, the traces of the ancestral lineage have diminished even more. This hybridization may have been reflected in the language, reducing the features of the common ancestral language to an imperceptible degree. However, interaction may not be a phenomenon that completely eliminates common ancestry. The Hungarian example is interesting. Hungarians today are largely genetically close to their Indo-European neighbors. This is often explained by the model of elite dominance. Despite this genetic closeness, the linguistic features of Hungarian - albeit with some influences - are not close to Indo-European languages, but to Uralic languages such as Finnish.[8] When claiming kinship for a language, interaction and common origin explanations should not be considered as alternatives to each other. The mixing of languages with their neighbors and the mixing of humans with their neighbors may not go hand in hand. We have seen above that there are examples that show that the assertion that languages have a common origin, therefore there has been no interaction in genes and languages is not true. Therefore, interaction should always be taken into account in language and gene relationships. For example, the claim that the Altaic languages Turkish and Mongolian do not come from the same root is not true for the dynamics of diversification. There may be interaction between neighboring languages with the same root, or a language influenced by a completely different language in one area may preserve its root features in another area. Here, it is imperative to take pre-writing periods into account and get support from archaeology and genetics.

In spite of these difficulties, it is also true that genetic research offers important possibilities that we cannot reach through classical language studies. I conclude this section with M. Pagel's commentary on what Darwinian perspectives can bring to the study of the evolution of languages:

> *Where biological bodies are the temporary repositories of genes, human minds are the temporary repositories of words. Both genes and words increase their probability of being transmitted—one into a new body, the other into a new mind—by adopting forms that are fitter than their competitors. The last 50 years or so of evolutionary studies has documented countless instances of the adaptation of genes. Now, in the last 10 to 20 years, the increasing use of evolutionary perspectives in combination with phylogenetic-statistical methods is documenting patterns in the evolution of languages, words and sound systems that are consistent with language adapting to the minds and habits of its speakers (Christiansen & Chater, 2008). These new methods bring an explicit hypothesis testing rigor and make possible inferences, analyses, and tests not available to traditional studies* (Pagel, 2017).

---

[8] A new study based on syntax and DNA analysis has identified a genetic link between Finns and Hungarians, concluding that Uralic languages support Darwin's claim of language and gene overlap (Santos et al., 2020).

Conclusion

The journey of the human species on the planet Earth continues today at a speed and breadth never before seen in history. This borderless mobility brings with it the mixing of genes as well as cultures. The study *A Genetic Atlas of Human Admixture History* revealed that major historical events such as the Mongol invasion, European colonialism and Arab slavery were major factors in today's gene mixing (Hellenthal et al., 2014). When Darwin first proposed the idea that languages are related to lineages, there was not yet a science called genetics. Today, genetics stands out as a very important tool that allows us to investigate this idea.

In this article, I tried to evaluate the publications that investigate the relatedness of Turkish with genetic data. I would like to point out that there are also studies on other languages neighboring Turkish, such as Ural[9] , Asian[10] , Siberian[11] languages. Investigating whether relatedness of languages and their speakers overlap is a new field. The number of studies on Turkish is also not sufficient. Our knowledge on the subject will undoubtedly increase in the coming years. I think it is possible to reach a few conclusions based on the research conducted so far:

a. Languages and genes show parallelism to a great extent.

b. However, the number of examples that do not fit this generalization is also considerable. Mismatched examples due to language change or genetic admixture are also common as a result of human species interaction.

c. Although Turkish is close to languages such as Mongolian and Manchu-Tungus in terms of linguistic features, no consensus has yet been reached on the theory called the Altaic language family in the scientific world. Genetic similarities have also been identified in these communities. These studies can be considered as a genetic support for the Altaic theory. However, more data is needed to claim with certainty that both linguistic and genetic commonalities stem from the same origin. Because the speakers of these languages have both cultural and genetic intermingling with their neighbors.

d. There is a lack of collaboration between genetics, archaeology and linguistics researchers in Turkey. Interdisciplinary studies should be planned in this direction.

---

[9] For prominent gene studies on Uralic-speaking communities, see (Santos et al., 2020; Tambets et al., 2018; Tömöry et al., 2007).

[10] For studies on communities in Asia, see. (Arnaiz-Villena et al., 2003; Chen et al., 2021; Gavashelishvili et al., 2023; He et al., 2022; Jeong et al., 2016, 2019; Jeong, Wang, Wilkin, et al., 2020; Jones et al., 2015; Mallory et al., 2019; Matsumae et al., 2021; Siska et al., 2017; Wang et al., 2021; H. Zhang et al., 2020; M. Zhang et al., 2019).

[11] For studies on the communities in the Siberian region, see. (Novgorodov et al., 2015; Pakendorf et al., 2003; Pugach et al., 2016).

References

Arbib, M. A., & Rizzolatti, G. (1996). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition, 29*, 393–424.

Arnaiz-Villena, A., Karin, M., Bendikuze, N., Gomez-Casado, E., Moscoso, J., Silvera, C., Oguz, F. S., Sarper Diler, A., De Pacho, A., Allende, L., Guillen, J., & Martinez Laso, J. (2001). HLA alleles and haplotypes in the Turkish population: Relatedness to Kurds, Armenians and other Mediterraneans. *Tissue Antigens*, *57*(4), 308–317.

Arnaiz-Villena, A., Martinez-Laso, J., Moscoso, J., Livshits, G., Zamora, J., Gomez-Casado, E., Silvera-Redondo, C., Melvin, K., & Crawford, M. H. (2003). HLA genes in the Chuvashian population from European Russia: Admixture of Central European and Mediterranean populations. *Human Biology*, *75*(3), 375–392.

Baker, J. L., Rotimi, C. N., & Shriner, D. (2017). Human ancestry correlates with language and reveals that race is not an objective genomic classifier. *Scientific Reports 2017 7:1*, *7*(1), 1–10.

Barbieri, C., Blasi, D. E., Arango-Isaza, E., Sotiropoulos, A. G., Hammarström, H., Wichmann, S., Greenhill, S. J., Gray, R. D., Forkel, R., Bickel, B., & Shimizu, K. K. (2022). A global analysis of matches and mismatches between human genetic and linguistic histories. *Proceedings of the National Academy of Sciences*, *119*(47), e2122084119.

Barbieri, C., & Widmer, P. (2024). Advances in population genetics and language history: How large datasets and ancient DNA changed the picture. Içinde M. Robbeets & M. Hudson (Ed.), *Oxford Handbook of Archaeology and Language*. Oxford University Press.

Berkman, C. C., Dinc, H., Sekeryapan, C., & Togan, I. (2008). Alu insertion polymorphisms and an assessment of the genetic contribution of Central Asia to Anatolia with respect to the Balkans. *American journal of physical anthropology*, *136*(1), 11–18.

Berwick, R. C., & Chomsky, N. (2016). *Why only us: Language and evolution.* The MIT Press.

Blažek, V. (2019). *Altaic languages. History of research, survey, classification and a sketch of comparative grammar*. Masaryk University Press.

Campbell, L., & Poser, W. J. (2008). *Language classification. History and method.* Cambridge University Press.

Ceolin, A. (2019). Significance testing of the Altaic family. *Diachronica. International Journal for Historical Linguistics*, *36*(3), 299–336.

Ceolin, A., Guardiano, C., Longobardi, G., Irimia, M. A., Bortolussi, L., & Sgarro, A. (2021). At the boundaries of syntactic prehistory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1824), rstb.2020.0197.

Chen, J., He, G., Ren, Z., Wang, Q., Liu, Y., Zhang, H., Yang, M., Zhang, H., Ji, J., Zhao, J., Guo, J., Zhu, K., Yang, X., Wang, R., Ma, H., Wang, C. C., & Huang, J. (2021). Genomic insights into the admixture history of Mongolic- and Tungusic-speaking populations from southwestern east Asia. *Frontiers in Genetics*, *12*, 685285.

Cinnioğlu, C., King, R., Kivisild, T., Kalfoğlu, E., Atasoy, S., Cavalleri, G. L., Lillie, A. S., Roseman, C. C., Lin, A. A., Prince, K., Oefner, P. J., Shen, P., Semino, O., Cavalli-Sforza, L. L., & Underhill, P. A. (2004). Excavating

Y-chromosome haplotype strata in Anatolia. *Human genetics*, *114*(2), 127–148.

Clauson, G. (2004). Altay teorisinin leksikoistatistiksel bir değerlendirmesi (çev. İ. Ulutaş). *Süleyman Demirel Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi*, *13*(10), 153–174.

Clauson, G. (2017). *Türkçe-Moğolca çalışmaları (çev. F. Kömürcü)*. Türk Dil Kurumu.

Colonna, V., Boattini, A., Guardiano, C., Dall'Ara, I., Pettener, D., Longobardi, G., & Barbujani, G. (2011). Long-range comparison between genes and languages based on syntactic distances. *Human Heredity*, *70*(4), 245–254.

Corballis, M. C. (2003). *İşaretten konuşmaya dilin kökeni ve gelişimi (çev. A. Görey)*. Kitabevi.

de Rachewiltz, I., & Rybatzki, V. (2010). *Introduction to Altaic philology: Turkic, Mongolian, Manchu*. Brill.

Dediu, D. (2015). *An introduction to genetics for language scientists: Current concepts, methods and findings*. Cambridge University Press.

Demir, N., & Yılmaz, E. (2002). Ural-Altay ve Altay Dil Teorisi. Içinde *Türkler* (C. 1, ss. 394–402). Yeni Türkiye Yay.

Di Benedetto, G., Ergüven, A., Stenico, M., Castrì, L., Bertorelle, G., Togan, I., & Barbujani, G. (2001). DNA diversity and population admixture in Anatolia. *American journal of physical anthropology*, *115*(2), 144–156.

Doerfer, G. (1983). Temel sözcükler ve Altay dilleri sorunu. *Türk Dili Araştımaları Yıllığı-Belleten*, *28–29*, 1–16.

Dunn, M., Terrill, A., Reesink, G., Foley, R. A., & Levinson, S. C. (2005). Structural phylogenetics and the reconstruction of ancient language history. *Science*, *309*(5743), 2072–2075.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Ed.). (2023). *Ethnologue: Languages of the world* (26. baskı). SIL International.

Egorov, I. M., Dybo, A. V, & Kassian, A. S. (2022). Phylogeny of the Turkic languages inferred from basic vocabulary: Limitations of the lexicostatistical methods in an intensive contact situation. *Journal of Language Evolution*, *7*(1), 16–39.

Ercilasun, A. B. (2004). *Türk dili tarihi*. Akçağ.

Feldman, M., Fernández-Domínguez, E., Reynolds, L., Baird, D., Pearson, J., Hershkovitz, I., May, H., Goring-Morris, N., Benz, M., Gresky, J., Bianco, R. A., Fairbairn, A., Mustafaoğlu, G., Stockhammer, P. W., Posth, C., Haak, W., Jeong, C., & Krause, J. (2019). Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nature Communications 2019 10:1*, *10*(1), 1–10.

Fischer, S. R. (2013). *Dilin tarihi (çev. M. Güvenç)*. Türkiye İş Bankası Kültür Yayınları.

Gavashelishvili, A., Chukhua, M., Sakhltkhutsishvili, K., Koptekin, D., & Somel, M. (2023). The time and place of origin of South Caucasian languages: insights into past human societies, ecosystems and human population genetics. *Scientific Reports 2023 13:1*, *13*(1), 1–14.

Georg, S. (2004). Sergei Starostin, Anna Dybo, and Oleg Mudrak (eds.): Etymological Dictionary of the Altaic Languages (2003). *Diachronica*, *21*(2), 445–450.

Georg, S. (2005). Reply [to Starostin response, 2005]. *Diachronica*, *22*(2), 455–457.

Georg, S. (2011). The poverty of Altaicism. *Altaique ou pas? (organized by Guillaume Jacques and Anton Antonov on 10 Dec 2011 in Paris).*

Georg, S. (2013). *Remarks on anonymous remarks on my "Poverty of Altaicism"*. https://www.academia.edu/2560087/Some_remarks_on_anonymous_remarks_on_my_Poverty_of_Altaicism_.

Georg, S., Michalove, P. A., Ramer, A. M., & Sidwell, P. J. (1999). Telling general linguists about Altaic. *Journal of Linguistics*, *35*(1), 65–98.

Gökçe, F. (2015). *Genetik dilbilime giriş: Altay dilleri sorunu*. Sonçağ Yay.

Gökçümen, Ö., Gultekin, T., Alakoc, Y. D., Tug, A., Gulec, E., & Schurr, T. G. (2011). Biological ancestries, kinship connections, and projected identities in four Central Anatolian settlements: Insights from culturally contextualized genetic anthropology. *American Anthropologist*, *113*(1), 116–131.

Gray, R. (2005). Pushing the time barrier in the quest for language roots. *Science*, *309*(5743), 2007–2008.

Greenhill, S. J. (2021). Do languages and genes share cultural evolutionary history? *Science Advances*, *7*(41).

Greenhill, S. J., Atkinson, Q. D., Meade, A., & Gray, R. D. (2010). The shape and tempo of language evolution. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1693), 2443–2450.

Guardiano, C., & Longobardi, G. (2005). Parametric comparison and language taxonomy. Içinde M. vd. Batllori (Ed.), *Grammaticalization and Parametric Variation* (ss. 149–174). Oxford University Press.

He, G., Wang, M., Zou, X., Yeh, H., Liu, C., Liu, C., Chen, G., & Wang, C. (2022). Extensive ethnolinguistic diversity at the crossroads of North China and South Siberia reflects multiple sources of genetic diversity. *Journal of Systematics and Evolution*.

Hellenthal, G., Busby, G. B. J., Band, G., Wilson, J. F., Capelli, C., Falush, D., & Myers, S. (2014). A genetic atlas of human admixture history. *Science*, *343*(6172), 747–751.

Hodoğlugil, U., & Mahley, R. W. (2012). Turkish population structure and genetic ancestry reveal relatedness among Eurasian populations. *Annals of Human Genetics*, *76*(2), 128–141.

Hruschka, D. J., Branford, S., Smith, E. D., Wilkins, J., Meade, A., Pagel, M., & Bhattacharya, T. (2015). Detecting regular sound changes in linguistics as events of concerted evolution. *Current Biology*, *25*(1), 1–9.

Hübler, N. (2022). Phylogenetic signal and rate of evolutionary change in language structures. *Royal Society Open Science*, *9*(3).

Hübler, N., & Greenhill, S. J. (2023). Modelling admixture across language levels to evaluate deep history claims. *Journal of Language Evolution*, *XX*, 1–18.

Hublin, J. J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., & Gunz, P. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature 2017 546:7657*, *546*(7657), 289–292.

Janhunen, J. (2023). The unity and diversity of Altaic. *https://doi.org/10.1146/annurev-linguistics-030521-042356*, *9*(1), 135–154.

Janhunen, J., Peltomaa, M., Sandman, E., & Dongzhou, X. (2008). *Wutun (Languages of the World Materials)*. Lincom Europa.

Jeong, C., Balanovsky, O., Lukianova, E., Kahbatkyzy, N., Flegontov, P., Zaporozhchenko, V., Immel, A., Wang, C.-C., Ixan, O., Khussainova, E., Bekmanov, B., Zaibert, V., Lavryashina, M., Pocheshkhova, E., Yusupov, Y., Agdzhoyan, A., Koshel, S., Bukin, A., Nymadawa, P., … Krause, J. (2019). The genetic history of admixture across inner Eurasia. *Nature Ecology & Evolution*, *3*(6), 966–976.

Jeong, C., Ozga, A. T., Witonsky, D. B., Malmström, H., Edlund, H., Hofman, C. A., Hagan, R. W., Jakobsson, M., Lewis, C. M., Aldenderfer, M. S., Rienzo, A. Di, & Warinner, C. (2016). Long-term genetic stability and a high-altitude East Asian origin for the peoples of the high valleys of the Himalayan arc. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7485–7490.

Jeong, C., Wang, C. C., & Ning, C. (2020). Transeurasian unity from a population-genetic perspective. Içinde M. Robbeets & A. Savelyev (Ed.), *The Oxford Guide to the Transeurasian Languages* (ss. 784–793). Oxford University Press.

Jeong, C., Wang, K., Wilkin, S., Taylor, W. T. T., Miller, B. K., Bemmann, J. H., Stahl, R., Chiovelli, C., Knolle, F., Ulziibayar, S., Khatanbaatar, D., Erdenebaatar, D., Erdenebat, U., Ochir, A., Ankhsanaa, G., Vanchigdash, C., Ochir, B., Munkhbayar, C., Tumen, D., … Warinner, C. (2020). A dynamic 6,000-year genetic history of Eurasia's eastern steppe. *Cell*, *183*(4), 890-904.e29.

Johansson, S. (2005). *Origins of language. Constraints on hypotheses*. John Benjamins Publishing Company.

Jones, E. R., Gonzalez-Fortes, G., Connell, S., Siska, V., Eriksson, A., Martiniano, R., McLaughlin, R. L., Gallego Llorente, M., Cassidy, L. M., Gamba, C., Meshveliani, T., Bar-Yosef, O., Müller, W., Belfer-Cohen, A., Matskevich, Z., Jakeli, N., Higham, T. F. G., Currat, M., Lordkipanidze, D., … Bradley, D. G. (2015). Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications*, *6*, 8912.

Kars, M. E., Başak, A. N., Onat, O. E., Bilguvar, K., Choi, J., Itan, Y., Çaglar, C., Palvadeau, R., Casanova, J. L., Cooper, D. N., Stenson, P. D., Yavuz, A., Buluş, H., Günel, M., Friedman, J. M., & Özçelik, T. (2021). The genetic structure of the Turkish population reveals high levels of variation and admixture. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(36), e2026076118.

Kassian, A. S., Starostin, G., Egorov, I. M., Logunova, E. S., & Dybo, A. (2021). Permutation test applied to lexical reconstructions partially supports the Altaic linguistic macrofamily. *Evolutionary Human Sciences*, *3*.

Kerimoğlu, C. (2016). Dilin kökeni arayışları I: Dilin kökeniyle ilgili akademik tartışmalar. *Dil Araştırmaları*, *18*, 47–84.

Kerimoğlu, C. (2017). Dilin kökeni arayışları II: FOXP2 geni. *Dil Araştırmaları*, *21*, 35–50.

Kerimoğlu, C. (2018). Dilin kökeni arayışları 3: Hayvan iletişimi-I. *Dil Araştırmaları*, *23*, 23–56.

Kerimoğlu, C. (2019a). Dilin kökeni arayışları 3: Hayvan iletişimi-II. *Dil Araştırmaları*, *24*, 43–73.

Kerimoğlu, C. (2019b). Dilin kökeni arayışları 4: Neandertallerin dili var mıydı? *Dil Araştırmaları*, *25*, 7–53.

Kerimoğlu, C. (2020). *Neandertaller konuşur muydu? İnsanın ve dilin kökenine bir yolculuk*. Varyant Yay.

Kerimoğlu, C. (2021). *Chomsky Darwin'e karşı: Evrensel dilbilgisi, dilin kökeni ve evrim*. Varyant Yay.

Kerimoğlu, C. (2022). Dilin kökeni arayışları 5: Beyin ve dil. *Dil Araştırmaları*, *16*(30).

Kerimoğlu, C. (2023). Gen araştırmaları ve dil aileleri. Içinde C. Kerimoğlu, İ. Şahin, & E. Ayan (Ed.), *Prof. Dr. Zeki Kaymaz Armağanı* (ss. 349–364). Pegem Akademi.

Kılınç, G. M., Koptekin, D., Atakuman, Ç., Sümer, A. P., Dönertas, H. M., Yaka, R., Bilgin, C. C., Büyükkarakaya, A. M., Baird, D., Altinişik, E., Flegontov, P., Götherström, A., Togan, İ., & Somel, M. (2017). Archaeogenomic analysis of the first steps of Neolithization in Anatolia and the Aegean. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1867).

Kılınç, G. M., Omrak, A., Özer, F., Günther, T., Büyükkarakaya, A. M., Bıçakçı, E., Baird, D., Dönertaş, H. M., Ghalichi, A., Yaka, R., Koptekin, D., Açan, S. C., Parvizi, P., Krzewińska, M., Daskalaki, E. A., Yüncü, E., Dağtaş, N. D., Fairbairn, A., Pearson, J., … Götherström, A. (2016). The demographic development of the first farmers in Anatolia. *Current Biology*, *26*(19), 2659–2666.

Lee, J. Y., & Kuang, S. (2017). A comparative analysis of Chinese historical sources and y-dna studies with regard to the early and medieval Turkic peoples. *Inner Asia*, *19*(2), 197–239.

Longobardi, G. (2003). Methods in parametric linguistics and cognitive history. *Linguistic Variation Yearbook*, *3*(1), 101–138.

Longobardi, G., Ghirotto, S., Guardiano, C., Tassi, F., Benazzo, A., Ceolin, A., & Barbujani, G. (2015). Across language families: Genome diversity mirrors linguistic variation within Europe. *American Journal of Physical Anthropology*, *157*(4), 630–640.

Longobardi, G., & Guardiano, C. (2009). Evidence for syntax as a signal of historical relatedness. *Lingua*, *119*(11), 1679–1706.

Longobardi, G., & Guardiano, C. (2017). Phylogenetic reconstruction in syntax: The parametric comparison method. Içinde *The Cambridge Handbook of Historical Syntax* (ss. 241–272). Cambridge University Press.

Longobardi, G., Guardiano, C., Silvestri, G., Boattini, A., & Ceolin, A. (2013). Toward a syntactic phylogeny of modern Indo-European languages. *Journal of Historical Linguistics*, *3*(1), 122–152.

Mallory, J., Dybo, A., & Balanovsky, O. (2019). The impact of genetics research on archaeology and linguistics in Eurasia. Içinde *Russian Journal of Genetics* (C. 55, Sayı 12, ss. 1472–1487). Pleiades Publishing.

Malyarchuk, B., Derenko, M., Denisova, G., Wozniak, M., Grzybowski, T., Dambueva, I., & Zakharov, I. (2010). Phylogeography of the Y-chromosome haplogroup C in northern Eurasia. *Annals of Human Genetics*, *74*(6), 539–546.

Manaster Ramer, A., Vovin, A., & Sidwell, P. (1998). On body part terms as evidence in favor of the Altaic hypothesis. *Ural-altaische Jahrbücher (Neue Folge)*, *15*, 116–138.

Martin, S. E. (1966). Lexical evidence relating Korean to Japanese. *Language*, *42*(2), 185.

Matsumae, H., Ranacher, P., Savage, P. E., Blasi, D. E., Currie, T. E., Koganebuchi, K., Nishida, N., Sato, T., Tanabe, H., Tajima, A., Brown, S., Stoneking, M., Shimizu, K. K., Oota, H., & Bickel, B. (2021). Exploring correlations in genetic and cultural variation across language families in northeast Asia. *Science Advances*, *7*(34), 9223–9241.

McMahon, A., & McMahon, R. (2013). *Evolutionary linguistics*. Cambridge University Press.

McMahon, R. (2004). Genes and languages. *Public Health Genomics*, *7*(1), 2–13.

Miller, R. A. (1971). *Japanese and the other Altaic languages*. The Chicago University Press.

Miller, R. A., & Street, J. (1975). *Altaic elements in Old Japanese*. [yazarların kendi basımı].

Nichols, J. (1992). *Linguistic diversity in space and time*. University of Chicago Press.

Novgorodov, I., Lemskaya, V., Tokmashev, D., & Aktaş, E. (2015). On a multidisciplinary study of South Siberian Turkic varieties (in Comparison with Yakut). Part I. *Procedia - Social and Behavioral Sciences*, *206*, 114–122.

Pagel, M. (2000). The history, rate and pattern of world linguistic evolution. Içinde C. Knight, M. Studdert-Kennedy, & J. Hurford (Ed.), *The Evolutionary Emergence of Language* (ss. 391–416). Cambridge University Press.

Pagel, M. (2017). Darwinian perspectives on the evolution of human languages. *Psychonomic Bulletin and Review*, *24*(1), 151–157.

Pakendorf, B., Wiebe, V., Tarskaia, L. A., Spitsyn, V. A., Soodyall, H., Rodewald, A., & Stoneking, M. (2003). Mitochondrial DNA evidence for admixed origins of central Siberian populations. *American Journal of Physical Anthropology*, *120*(3), 211–224.

Pereltsvaig, A. (2021). *Languages of the world. An introduction* (3. baskı). Cambridge University Press.

Pugach, I., Matveev, R., Spitsyn, V., Makarov, S., Novgorodov, I., Osakovsky, V., Stoneking, M., & Pakendorf, B. (2016). The complex admixture history and recent southern origins of Siberian populations. *Molecular Biology and Evolution*, *33*(7), 1777–1795.

Rama, T., & Wichmann, S. (2020). A test of Generalized Bayesian dating: A new linguistic dating method. *PLOS ONE*, *15*(8), e0236522.

Ramstedt, G. J. (1922). Zur frage nach der stellung des tschuwassischen. *Journal de la Societe Finno-Ougrienne*, *38*(3), 3–34.

Ramstedt, G. J. (1952). *Einfuhrung in die Altaische sprachwissenschaft 2: Formenlehre*. Suomalais Ugrilainen Seura.

Ramstedt, G. J. (1957). *Einführung in die altaische Sprachwissenschaft 1: Lautlehre*. Suomalais-ugrilainen seura.

Ramstedt, G. J. (1966). *Einführung in die altaische Sprachwissenschaft 3: Register*. Suomalais-Ugrilainen Seura.

Renfrew, C. (1992). Archaeology, genetics and linguistic diversity. *Man*, *27*(3), 445.

Renfrew, C. (2000). At the edge of knowability: Towards a prehistory of languages. *Cambridge Archaeological Journal*, *10*(1), 7–34.

Renfrew, C. (2010). Archaeogenetics - Towards a "new synthesis"? *Current Biology*, *20*(4), R162–R165.

Richter, D., Grün, R., Joannes-Boyau, R., Steele, T. E., Amani, F., Rué, M., Fernandes, P., Raynal, J. P., Geraads, D., Ben-Ncer, A., Hublin, J. J., & McPherron, S. P. (2017). The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. *Nature 2017 546:7657*, *546*(7657), 293–296.

Robbeets, M. (2005). *Is Japanese related to Korean, Tungusic, Mongolic and Turkic?* Harrassowitz Verlag.

Robbeets, M., & Bouckaert, R. (2018). Bayesian phylolinguistics reveals the internal structure of the Transeurasian family. *Journal of Language Evolution*, *3*(2), 145–162.

Robbeets, M., Bouckaert, R., Conte, M., Savelyev, A., Li, T., An, D.-I., Shinoda, K., Cui, Y., Kawashima, T., Kim, G., Uchiyama, J., Dolińska, J., Oskolskaya, S., Yamano, K.-Y., Seguchi, N., Tomita, H., Takamiya, H., Kanzawa-Kiriyama, H., Oota, H., … Ning, C. (2021). Triangulation supports agricultural spread of the Transeurasian languages. *Nature 2021*, 1–6.

Robbeets, M., & Savelyev, A. (Ed.). (2017). *Language dispersal beyond farming*. John Benjamins Publishing.

Robbeets, M., & Savelyev, A. (Ed.). (2020). *The Oxford guide to the Transeurasian languages*. Oxford University Press.

Ruhlen, M. (2006). *Dilin kökeni (çev. İsmail Ulutaş)*. Hece Yay.

Santos, P., Gonzàlez-Fortes, G., Trucchi, E., Ceolin, A., Cordoni, G., Guardiano, C., Longobardi, G., & Barbujani, G. (2020). More rule than exception: Parallel evidence of ancient migrations in grammars and genomes of Finno-Ugric speakers. *Genes 2020, Vol. 11, Page 1491*, *11*(12), 1491.

Savelyev, A., & Robbeets, M. (2020). Bayesian phylolinguistics infers the internal structure and the time-depth of the Turkic language family. *Journal of Language Evolution*, *5*(1), 39–53.

Schott, W. (1841). *De lingua Tschuwaschorum*. Veitii & socii sumpt.

Siska, V., Jones, E. R., Jeon, S., Bhak, Y., Kim, H. M., Cho, Y. S., Kim, H., Lee, K., Veselovskaya, E., Balueva, T., Gallego-Llorente, M., Hofreiter, M., Bradley, D. G., Eriksson, A., Pinhasi, R., Bhak, J., & Manica, A. (2017). Genome-wide data from two early Neolithic East Asian individuals dating to 7700 years ago. *Science Advances*, *3*(2).

Starostin, S. A., Dybo, A. V, & Mudrak, O. A. (2003). *An etymological dictionary of Altaic languages*. Brill.

Stoneking, M. (2017). *An introduction to molecular anthropology*. Wiley Blackwell.

Street, J. (1980). Japanese and the Altaic Languages. *Orientalistische Literaturzeitung*, *75*(1–6), 51–53.

Swadesh, M. (1952). Lexico-statistic dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society. Studies of Historical Documents in the Library of the American Philosophical Society* , *96*(4), 452–463.

Swadesh, M. (1955). Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics*, *21*(2), 121–137.

Swadesh, M., Sherzer, J., & Hymes, D. H. (1971). *The origin and diversification of language*. Aldine Atherton.

Tambets, K., Yunusbayev, B., Hudjashov, G., Ilumäe, A. M., Rootsi, S., Honkola, T., Vesakoski, O., Atkinson, Q., Skoglund, P., Kushniarevich, A., Litvinov, S., Reidla, M., Metspalu, E., Saag, L., Rantanen, T., Karmin, M., Parik, J., Zhadanov, S. I., Gubina, M., … Metspalu, M. (2018). Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biology*, *19*(1), 1–20.

Tekin, T. (2003). *Makaleler 1: Altayistik*. Grafiker Yayıncılık.

Tian, Z., Tao, Y., Zhu, K., Jacques, G., Ryder, R. J., Fuente, J. A. A. de la, Antonov, A., Xia, Z., Zhang, Y., Ji, X., Ren, X., He, G., Guo, J., Wang, R., Yang, X., Zhao, J., Xu, D., Gray, R. D., Zhang, M., … Pellard, T. (2022). Triangulation fails when neither linguistic, genetic, nor archaeological data support the Transeurasian narrative. *bioRxiv*, 2022.06.09.495471.

Tjuka, A., Forkel, R., & List, J.-M. (2024). Universal and cultural factors shape body part vocabularies. *Scientific Reports 2024 14:1*, *14*(1), 1–12.

Tömöry, G., Csányi, B., Bogácsi-Szabó, E., Kalmár, T., Czibula, Á., Csosz, A., Priskin, K., Mende, B., Langó, P., Downes, C. S., & Raskó, I. (2007). Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations. *American Journal of Physical Anthropology*, *134*(3), 354–368.

Tuna, O. N. (1992). Altay dilleri teorisi . Içinde *Türk Dünyası El Kitabı* (C. 2). TKAE Yay.

Turchin, P., Peiros, I. I., & Gell-Mann, M. (2010). Analyzing genetic connections between languages by matching consonant classes. *Journal of Language Relationship*, *3*, 117–126.

Vovin, A. (2005). The end of the Altaic controversy in memory of Gerhard Doerfer. *Source: Central Asiatic Journal*, *49*(1), 71–132.

Vovin, A. (2009). 'Japanese, Korean, and Other Non-Altaic Languages', a review article of Martine Robbeets. Is Japanese related to Korean, Tungusic, Mongolic, and Turkic? Wiesbaden: Harrassowitz, 2005. *Central Asiatic Journal*, *53*(1), 105–147.

Vovin, A. (2011). *Why Japonic is not demonstrably related to 'Altaic" or Korean*. Historical Linguistics in the Asia-Pacific region and the position of Japanese, ICHL XX, Osaka.

Wang, C. C., Yeh, H. Y., Popov, A. N., Zhang, H. Q., Matsumura, H., Sirak, K., Cheronet, O., Kovalev, A., Rohland, N., Kim, A. M., Mallick, S., Bernardos, R., Tumen, D., Zhao, J., Liu, Y. C., Liu, J. Y., Mah, M., Wang, K., Zhang, Z., … Reich, D. (2021). Genomic insights into the formation of human populations in East Asia. *Nature 2021 591:7850*, *591*(7850), 413–419.

Wang, Z., Wang, M., Hu, L., He, G., & Nie, S. (2024). Evolutionary profiles and complex admixture landscape in East Asia: New insights from modern and ancient Y chromosome variation perspectives. *Heliyon*, *10*(9), e30067.

Wei, L. H., Huang, Y. Z., Yan, S., Wen, S. Q., Wang, L. X., Du, P. X., Yao, D. L., Li, S. L., Yang, Y. J., Jin, L., & Li, H. (2017). Phylogeny of Y-chromosome haplogroup C3b-F1756, an important paternal lineage in Altaic-speaking populations. *Journal of Human Genetics 2017 62:10*, *62*(10), 915–918.

Wen, S., Xie, X., & Xu, D. (2017). Relationship between Altaic speaking populations and their languages viewed from Y chromosomes. *Languages and Genes in Northwestern China and Adjacent Regions*, 79–86.

Wiedemann, F. (1838). *Ueber die früheren Sitze der tschudischen Völker und ihre Sprachverwandtschaft mit den Völkern Mittelhochasiens : Einladung zur öffentlichen Prüfung im hiesigen Gymnasium am 27sten und*. Gedruckt bei Lindfors Erben.

Xu, D., & Li, H. (Ed.). (2017). *Languages and genes in Northwestern China and adjacent regions*. Springer Singapore.

Xu, D., & Wen, S. (2017). The silk road: Language and population admixture and replacement. *Languages and Genes in Northwestern China and Adjacent Regions*, 55–78.

Yaka, R., Mapelli, I., Kaptan, D., Doğu, A., Chyleński, M., Erdal, Ö. D., Koptekin, D., Vural, K. B., Bayliss, A., Mazzucato, C., Fer, E., Çokoğlu, S. S., Lagerholm, V. K., Krzewińska, M., Karamurat, C., Gemici, H. C., Sevkar, A., Dağtaş, N. D., Kılınç, G. M., … Somel, M. (2021). Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Current Biology*, *31*(11), P2455-2468.E18.

Yunusbayev, B., Metspalu, M., Metspalu, E., Valeev, A., Litvinov, S., Valiev, R., Akhmetova, V., Balanovska, E., Balanovsky, O., Turdikulova, S., Dalimova, D., Nymadawa, P., Bahmanimehr, A., Sahakyan, H., Tambets, K., Fedorova, S., Barashkov, N., Khidiyatova, I., Mihailov, E., … Villems, R. (2015). The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genetics*, *11*(4).

Zhang, H., Ji, T., Pagel, M., & Mace, R. (2020). Dated phylogeny suggests early Neolithic origin of Sino-Tibetan languages. *Scientific Reports 2020 10:1*, *10*(1), 1–8.

Zhang, M., Yan, S., Pan, W., & Jin, L. (2019). Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature 2019 569:7754*, *569*(7754), 112–115.